



UNIVERSITY OF OREGON



**23rd North American Conference on
Chinese Linguistics
Volume 1**

Edited by Zhuo Jing-Schmidt

June 17-29, 2011
Eugene, OR

©Cover design and front matters by NACCL-23

©All papers copyrighted by the authors

Proceedings of the 23rd North American Conference on Chinese
Linguistics (NACCL-23)

Volume 1/edited by Zhuo Jing-Schmidt

Published by:

University of Oregon, Eugene, OR.

Distributed by:

NACCL Proceedings Online

Department of East Asian Languages and Literatures

398 Hagerty Hall

1775 College Road

The Ohio State University

Columbus, Ohio

URL: <http://naccl.osu.edu>

CONTENTS OF VOLUME 1

	Page
Preface	v
Acknowledgements	vi

Part I Corpus Linguistics

<i>Tao Ming and Liang Chen</i>	1-17
The ordering of multiple relative clauses modifying the same head NP in Chinese follows information-flow principles	
<i>Shan Wang, Sophia Lee and Chu-ren Huang</i>	18-34
A corpus-based analysis of semantic type system of event nouns: a case study on <i>huiyi</i>	
<i>Wen Jin</i>	35-50
A statistical argument for the homophone avoidance approach to the disyllabification of Chinese	

Part II Historical Linguistics

<i>Scott DeLancey</i>	51-64
The origins of Sinitic	
<i>Jeeyoung Peck</i>	65-83
Analogy and reanalysis in the postposing of durative and iterative adverbials in the history of Chinese	

Part III Phonetics and Phonology

<i>Shawn Yung-hsiang Chang</i>	84-96
Distinction between Mandarin tones 2 and 3 for L1 and L2 listeners	
<i>Ping Jiang and Aishu Chen</i>	97-109
Representation of Mandarin intonations: boundary tone revisited	
<i>Yah-ting Shih and Eunjong Kong</i>	110-119
Perception of Mandarin fricatives by native speakers of Taiwan Mandarin and Taiwanese	

<i>Ying-Shing Li</i>	120-137
Investigating Taiwan Southern Min sub-syllabic structure using maximum entropy models and wordlikeness judgments	
<i>Chiung-Yao Wang and Yen-Hwei Lin</i>	138-155
Variation in tone 3 sandhi in Mandarin: the case of prepositions and pronouns	
<i>Seth Wiener</i>	156-172
Grass-mud horse to victory: the phonological constraints of subversive puns	
<i>Seth Wiener and Ya-ting Shih</i>	173-190
Divergent places of articulation: [ʋ] and [v] in modern spoken Mandarin.	

Part IV Psycholinguistics

<i>Chien-Jer Charles Lin</i>	191-199
Chinese and English relative clauses: processing constraints and typological consequences	
<i>Yowyu Lin</i>	200-214
Locality versus anti-locality effects in Mandarin sentence comprehension	
<i>Yu-an Lu</i>	215-226
The psychological reality of phonological representations: the case of Mandarin fricatives	

Part V Second Language Acquisition

<i>Ying Liu and Qian Du</i>	227-236
Chinese as a Foreign Language (CFL) learners' linguistic representation of voice in argumentative writing	
<i>Hongying Xu</i>	237-249
The acquisition of some properties of the BA construction by English-speaking learners of Chinese	
<i>Jin Zhang</i>	250-261
L2 acquisition of Chinese locative inversion	

PREFACE

The 23rd North American Conference on Chinese Linguistics (NACCL-23) was organized and hosted by the University of Oregon on June 17-19, 2011.

A total of 153 abstracts were submitted to the conference organizing committee. The abstracts were reviewed and rated by the NACCL-23 Scientific Committee (Susan G. Anderson, Marjorie K.M. Chan, Ying Chen, Scott DeLancey, Agnes W.Y. He, Zhuo Jing-Schmidt, Vsevolod Kapatsinski, Lizhen Peng, Chaofen Sun, Hongyin Tao, Liang Tao, and Janet Xing), and a total of 98 proposals were selected for presentation, and 83 presentations were actually made at the conference. The presentations represented 102 scholars from China, France, Germany, Hong Kong, Korea, Singapore, Taiwan, and the United States. Topics ranged from psycholinguistics and neurolinguistics to discourse analysis and corpus linguistics, from historical linguistics and Buddhist text translation studies to sociolinguistics and dialectology, from phonetics and phonology, syntax, semantics, and pragmatics to second language acquisition and language pedagogy.

Five internationally renowned scholars gave keynote speeches at the conference. They were Prof. Walter Bisang (Johannes Gutenberg University), Prof. Ina Bornkessel-Schlesewsky (University of Marburg), Prof. Chu-Ren Huang (The Hong Kong Polytechnic University), Prof. Agnes He (SUNY Stony Brook), and Prof. Fu-xiang Wu (Chinese Academy of Social Sciences).

Included in the NACCL-23 Conference Proceedings are 38 of the presented papers, divided into two volumes. Volume 1 consists of five parts: Part I Corpus Linguistics, Part II Historical Linguistics, Part III Phonetics and Phonology, Part IV Psycholinguistics, and Part V Second Language Acquisition. Volume 2 consists of five parts: Part I Semantics and Pragmatics, Part II Sociolinguistics and Socio-pragmatics, Part III Studies on Regional Varieties of Chinese, Part IV Syntax, and Part V Translation Study of Buddhist Sutras.

Zhuo Jing-Schmidt, Ph.D.

October 2011, Eugene

ACKNOWLEDGMENT

The 23rd North American Conference on Chinese Linguistics would not have been possible without the support and dedication of many. With gratitude I acknowledge the generosity of the following sponsors:

The Center of Asian and Pacific Studies, University of Oregon

The Confucius Institute, University of Oregon

The College of Arts and Sciences, University of Oregon

The Department of East Asian Languages and Literatures, University of Oregon

The conference would not have been successful without the help and hard work of many individuals. I am grateful to Marjorie K.M. Chan of Ohio State University, who was a great resource to me during the organizing process, sharing experience and patiently providing advice and encouragement.

I would like to thank colleagues on the NACCL-23 Scientific Committee who carefully reviewed abstracts despite their busy schedule. The committee members are Susan Guion Anderson (University of Oregon), Marjorie K.M. Chan (Ohio State University), Ying Chen (University of Oregon), Scott DeLancey (University of Oregon), Agnes Weiyun He (Stony Brook University), Zhuo Jing-Schmidt (University of Oregon), Vsevolod Kapatsinski (University of Oregon), Lizhen Peng (Zhejiang University), Chaofen Sun (Stanford University), Hongyin Tao (UCLA), Liang Tao (Ohio University), and Janet Zhiqun Xing (Western Washington University).

I owe a special *thank you* to Lori O'Hollaren, the assistant director of the Center of Asian and Pacific Studies at the University of Oregon, who managed the administration and logistics of the conference with great care and efficiency. Her professionalism, experience, attention to detail, and her positive energy are very much appreciated. I thank Yifang Zhang who joined Lori in the organizing process, working hard to ensure a successful conference. I also had a wonderful group of scholars and students volunteering for the conference. They are Dr. Lan Dai (Co-director of the UO Confucius Institute), Rong Hu, Galen Ettlin, Yingying Gu, Linda Konnerth, Ying Chen, Hideko Teruya, Katherine Thompson, and Tianqi Yang. Their hard work is very much appreciated.

Zhuo Jing-Schmidt, Ph.D.

October 2011, Eugene

The Ordering of Multiple Relative Clauses Modifying the Same Head NP in Chinese Follows Information-Flow Principles

Tao Ming
Concordia College

Liang Chen
The University of Georgia

This study investigates the ordering restriction of two relative clauses modifying the same head noun phrase in Chinese. We use both retrospective and corpus data to challenge Larson and Takahashi's (2007) account of the ordering of such multiple relative clauses in Chinese in terms of the distinction of individual-level and stage-level relative clauses. Instead, we offer an account based on the discourse function of the multiple relative clauses in Chinese. Specifically, we argue that relative clauses which serve to provide grounding information for the head noun tend to take precedence over other types of relative clauses. The proposed ordering restriction is shown to account for the ordering of multiple relative clauses in both Chinese and English.

Key Words: multiple relative clause, information flow, grounding, Chinese

1. Studies on Multiple Relative Clauses

In their seminal work on the grammar of relative clauses in conversational data, Fox & Thompson (1990) argue that the ordering of English relative clauses is determined by the information flow of the discourse. Specifically they find that non-human head nouns occurring in the subject position of the main clause tend to occur with an object relative clause (RC) in which the head noun is the object in the RC (e.g., *The book which I bought was very interesting.*). They suggest that the object RC can provide anchoring information for the non-human head noun whereas human head nouns prefer to occur with a subject RC in which the head noun is the subject in the RC (e.g., *The student who did not do his homework was my friend.*). The positioning of the head noun in the main clause also plays a role in determining the type of RCs to be used. For example, if a non-human head noun occurs in the object position of the main clause, it tends to occur with a subject RC instead of an object RC, thus precluding the dominance of O-O in which the head noun is the object in the RC as well as in the main clause (e.g., *I love the book which she wrote.*). Fox and Thompson, however, only explore human head nouns in

existential sentences. Whether human head nouns occurring in other sentence patterns behave the same way remains to be investigated.

More relevant to our present study is Fox and Thompson's discussion of utterances containing a sequence of two relative clauses that serve to modify the same head noun. They observe that the two RCs are sequenced in an orderly and predictable fashion in their conversational data. An object RC tends to precede other RCs (1a) when modifying a non-human head noun in the subject position of a main clause¹. However, if the head noun is human, subject RC tends to occur before other RCs (1b). Fox and Thompson (1990) argue that the order of the two relative clauses can be predicted from the principles governing information flow in discourse.

1. a. There was something [which we needed]_{RC1}[which was really obscure]_{RC2}
(Fox and Thompson: 313)
- b. There was a boy [that played the trombone]_{RC1} [that he kind of knew]_{RC2}
(Fox and Thompson: 314)

Chinese and English are both SVO languages, but in Chinese RCs precede their head NP². We may wonder whether information flow principles can also account for ordering of the multiple prenominal RCs in Chinese. According to Larson and Takahashi (2007), the answer is a resounding no. In fact, Larson and Takahashi challenge even Fox and Thompson's observations on English and report that postnominal multiple RCs do not show ordering preference and that multiple RCs can occur in any order. For example, the ordering of multiple RCs in (2a) and (2b) is said to be equally acceptable³.

2. a. The person [who I met]_{RC1} [who smokes]_{RC2}
- b. The person [who smokes]_{RC1}[who I met]_{RC2}

Larson and Takahashi (2007) argue that postnominal RCs differ from prenominal RCs in that the ordering of the former is free from any constraint whereas that of the latter hinges on whether they express stage-level versus individual level properties. Their study shows that stage-level RCs take precedence over individual RCs (3a) and that placement of individual-level RCs before stage-level RCs renders sentences (3b) unacceptable.

¹ As for a non-human head occurring in the object position of the main clause, it is expected that the first RC should be a subject RC. However, their data do not produce such an example.

² Among its other functions in Chinese, the word *de*, which we have notated as *de*, serves as a RC marker.

³ It is well-established that English restrictive RCs occur before non-restrictive ones and reduced ones such as modifying infinitivals.

3. a. [wo zuotian kanjian]_{RC1} de [hui shuo yidaliyu]_{RC2} de ren shi Lisi⁴.
 I yesterday see de can speak Italian de person is Lisi
 ‘The person whom I saw yesterday who can speak Italian is Lisi.’ (Good)
- b. [hui shuo yidaliyu]_{RC1} de [wo zuotian kanjian]_{RC2} de ren shi Lisi.
 can speak Italian de I yesterday see de person is Lisi
 ‘The person whom I saw yesterday who can speak Italian is Lisi.’ (Bad)

However, there is no ordering restriction if all the RCs exhibit the same property. For example, RCs expressing stage-level properties order freely among themselves.

4. a. [Cong Yidali huilai]_{RC1} de [wo zuotian kanjian]_{RC2} de ren shi Lisi.
 From Italy return de I yesterday see de person is Lisi
 ‘The person who returned from Italy whom I saw yesterday is Lisi.’ (Good)
- b. [Wo zuotian kanjian]_{RC2} de [cong Yidali huilai]_{RC1} de ren shi Lisi.
 I yesterday see de from Italy return de person is Lisi
 ‘The person whom I saw yesterday who returned from Italy is Lisi.’ (Good)

Similarly, RCs exhibiting individual-properties need not observe any ordering restriction either, as in (5a) and (5b).

5. a. [Hui shuo yidaliyu]_{RC1} de [xihuan qu yinyuehui]_{RC2} de ren shi Lisi.
 can speak Italian de like go concert de person is Lisi
 ‘The person who can speak Italian who likes to go to concert is Lisi.’ (Good)
- b. [Xihuan qu yinyuehui]_{RC2} de [hui shuo yidaliyu]_{RC1} de ren shi Lisi.
 like go concert de can speak Italian de person is Lisi
 ‘The person who can speak Italian who likes to go to concert is Lisi.’ (Good)

According to Larson and Takahashi (2007), the ordering of multiple postnominal RCs (i.e., stage-level RCs before individual-level RCs) follows a similar pattern with the ordering of multiple prenominal modifiers in English. If there is more than one prenominal adjective, the outside one tends to encode individual, permanent, and characteristic properties whereas the inside adjective tends to encode stage-level, episodic, transient properties (6a). The reversal of the order of the stage-level adjectival modifier and individual-level adjectival modifier generates such awkward expressions as (6b).

6. a. The nonvisible visible stars include Capella. (Coherent!)
- b. The visible nonvisible stars include Capella. (Odd!)

⁴ Following is a list of the abbreviations used in this paper following the convention in Li and Thompson (1981): BEI: passive marker; CL: classifier; PFV: perfective aspect.

Multiple prenominal RCs in Chinese behave similarly to multiple prenominal adjectives in English, according to Larson and Takahashi (2007), because both prenominal RCs in Chinese and prenominal adjectives in English do not have full clausal status because of “the reduced/participial status of prenominal relatives”⁵. On their account, individual-level prenominal RCs are closer to an NP for generic force quantification.

Now here is the issue. Whereas Fox and Thompson (1990) find that the ordering of multiple postnominal RCs is subject to the information flow principles in discourse, Larson and Takahashi (2007) believe that multiple postnominal RCs show no ordering preference, but multiple prenominal RCs are sequenced in terms of whether they exhibit stage-level or individual-level properties. This issue is addressed in the present study. In the next section, we show that the analysis by Larson and Takahashi (2007) is problematic and that a different analysis is needed to account for the ordering of Chinese multiple prenominal RCs.

2. The Ordering of Stage-Level RCs and Individual-level RCs

Larson and Takahashi (2007) propose that stage-level RCs always occur before individual-level RCs for generic force quantification and that multiple RCs expressing same properties do not obey any ordering restriction. However, counterexamples to their claim are abundant. The following examples in Chinese show that stage-level RCs can occur after individual-level RCs.

7. a. [Wo renshi]_{RC1}de [xie le yi-ben shu]_{RC2} de na-ge xuesheng hen congming.
I know de write PFV one-CL book de that-CL student very smart
‘The student whom I know who wrote a book is very smart.’ (Good)
- b. [Xie le yi-ben shu]_{RC1} de [wo renshi]_{RC2}de na-ge xuesheng hen congming.
write PFV one CL book de I know de that CL student very smart
‘The student who wrote a book whom I know is very smart.’ (Odd)

The first RC in (7a) expresses individual-level properties since knowing somebody is a stable knowledge. And the second RC in (7a) encodes stage-level properties because the completion of three books is a one-time event happening in the past. Contrary to the claim that stage-level RCs must precede individual-level RCs, the contrast in (7) shows that the opposite is true because the placement of the individual-level RC before the stage-level RC generates an awkward, if not ungrammatical, sentence in (7b).

⁵ It should be noted that Larson and Takahashi’s (2007) suggestion that a full-fledged Chinese RC should be treated as a participle or a reduced clause is questionable. Actually, the other way is said to be true. According to Sproat & Shih (1988, 1991) and Duanmu (1998), adjectives in A-de-N structures should be analyzed as full-fledged RCs in Chinese.

Secondly, the ordering of RCs expressing the same properties is not always random. Although Larson and Takahashi (2007) correctly report that both sentences (5a) and (5b) are equally acceptable, it is not difficult to find examples that exhibit ordering preference among the multiple RCs with the same properties. The following contrast shows that the word order in (8a) is favored over the one in (8b).

8. a. [Wo renshi]_{RC1} de [xihuan changge]_{RC2} de na ge xuesheng hen youhao.
 I know de like sing de that CL student very friendly
 ‘The student whom I know who likes to sing is very friendly.’ (Good)
- b. [Xihuan changge]_{RC2} de [wo renshi]_{RC1} de na ge xuesheng hen youhao.
 like sing de I know de that CL student very friendly
 ‘The student who likes to sing whom I know is very friendly.’ (Odd)

Although the two RCs in both (8a) and (8b) express the same properties, (e.g., individual-level properties), contrary to Larson and Takahashi’s account, their ordering is not random at all. The word order in (8a) generates a grammatically sound sentence whereas (8b) sounds pretty odd. Of course, there is no denying that sometimes RCs expressing individual-properties can be ordered freely as shown in (5a) and (5b). The same observation also extends to multiple RCs which encode stage-level properties.

9. a. [Wo zuazhu]_{RC1} de [ni fangzou]_{RC2} de na tiao yu hen da.
 I catch de you release de that CL fish very big
 ‘The fish which I caught which you released is very big.’ (Good)
- b. [Ni fangzou]_{RC1} de [wo zuazhu]_{RC2} de na tiao yu hen da.
 you release de I catch de that CL fish very big
 ‘The fish which you released which I caught is very big.’ (Good)

The two RCs in (9) exhibit stage-level properties since both of them describe a transient action happening in a certain past and their ordering is free. However, sometimes RCs expressing stage-level properties do manifest an ordering preference.

10. a. [Wo zuotian kanjian]_{RC1} de [bei daibu]_{RC2} de na ge xuesheng hen shuai
 I yesterday see de BEI arrest de that CL student very handsome
 ‘The student whom I saw who was arrested is very handsome.’ (Good)
- b. [Bei daibu]_{RC2} de [wo zuotian kanjian]_{RC1} de na ge xuesheng hen shuai
 BEI arrest de I yesterday see de that CL student very handsome
 ‘The student who was arrested whom I saw is very handsome.’ (Odd)

Examples such as those from (7) to (10) challenge Larson and Takahashi’s (2007) argument that the ordering of Chinese multiple RCs is related to the distinction of stage-level RCs and individual RCs.

Another place to examine the validity of Larson and Takahashi's (2007) argument is the ordering of prenominal adjectives in Chinese. It is a well-known fact that an adjective can be used to modify a noun with (11a) or without *de* (11b).

11. a. yi ge piaoliang de xuesheng
 one CL beautiful de student
 'a beautiful student'
 b. yi ge piaoliang xuesheng
 one CL beautiful student
 'a beautiful student'

It has been argued (Sproat & Shih 1988, 1991; Duanmu 1998) that 'A-de-N' structure should be analyzed as a RC⁶ since the same *de* is used in Chinese RCs whereas the *de*-less structure 'A-N' should be treated as a word. If the assumption that the *de*-modification 'A-de-N' should be treated as a RC is correct, the theory by Larson and Takahashi (2007) would predict that the stage-level adjectives takes precedence over the individual-level adjectives. Examples that do not follow this stringent pattern of ordering would cast serious doubt on their theory. Consider the examples in (12).

12. a. Gaogao de lüe xian pijuan de na ge xuesheng
 tallish de little appear tired de that CL student
 'The tall student who appeared a little bit tired' (Good)
 b. Lüe xian pijuan de gaogao de na ge xuesheng
 little appear tired de tallish de that CL student
 'The student who appeared a little bit tired who is pretty tall' (Good)

The examples in (12) indicate that shuffling the two modifying adjectives around does not affect the acceptability of the expression. They pose a challenge to the analysis of Larson and Takahashi. Actually, adjectives can be placed in any order with the help of *de* as shown in the following examples.

13. a. nenggan de congming de qinfen de xuesheng
 competent de smart de diligent de student
 'competent, smart, and diligent student' (Good)
 b. congming de nenggan de qinfen de xuesheng
 smart de competent de diligent de student
 'smart, competent, and diligent student' (Good)
 c. qinfen de congming de nenggan de xuesheng
 diligent de smart de competent de student

⁶See Paul (2005) for counterarguments.

‘diligent, smart, competent and student’ (Good)

14. a. feichang shengqi de hen pijuan de jichanglulu de xuesheng
 very angry de very tired de very-hungry de student
 ‘Very angry, tired, and hungry student’ (Good)
- b. hen pijuan de feichang shengqi de jichanglulu de xuesheng
 very tired de very angry de hungry de student
 ‘Very tired, angry, and hungry student’ (Good)
- c. jichanglulu de hen pijuan de feichang shengqi de xuesheng
 hungry de very tired de very angry de student
 ‘Very hungry, tired, and angry student’ (Good)

All the adjectives in (13) are individual-level ones while adjectives in (14) showcase stage-level adjectives. The reordering of them in any sequence does not diminish the acceptability of the expression at all.

Thus, the distinction between stage-level RCs and individual RCs cannot account for the ordering of Chinese multiple RCs, nor can it explain the ordering of multiple English RCs. In the next section, we will present our account of the ordering restriction on Chinese multiple RCs.

3. Grounding RCs Versus Non-grounding RCs

According to Lapolla (1995), when a referent is introduced into a discourse, its information status is either identifiable or unidentifiable. An unidentifiable referent can be in one of the following three activation statuses.

1) Active

The referent is in the current focus of consciousness

2) Accessible

The referent is not in the current focus of consciousness, but its identity can be derived from previous text, from situation, or through logics.

3) Inactive

The referent is currently not in the focus of consciousness of periphery of consciousness, but in the long term memory.

Identifiable referents are either brought back into the text after previous mention in the discourse or textually, situationally, inferentially derivable. Unidentifiable referents are introduced into the discourse for the first time and its identity cannot be established through a link with the previous text, situation or logical inference.

In effective communication, referents are supposed to be brought into the discourse in a way to make it relevant to the listener or speaker. Referents whose relevance is not established from its previous mention or situation must be grounded to justify their existence. According to Fox and Thompson (1990), a referent can be

grounded either by a main clause as in the English examples (15a-b) or by a RC as in Chinese examples 16(a-b).

15. a. He's got a spring-[that comes, way up]_{RC} (Fox and Thompson 1990: 301)
 b. There were two people there [who were constantly on stage]_{RC}

Sentence (15) instantiates main clause grounding. The unidentifiable referent *a spring* is grounded through the link with an identifiable referent *he*. A referent can also be grounded by a locative in the main clause. For example, the referent *two people* in (15b) is grounded by the locative *there*. In the same token, a referent can also be grounded by associating with an identifiable referent (16a) or locative (16b) in its modifying relative clause⁷.

16. a. [wo zuotian mai]_{RC} de na ben shu bei tou le
 I yesterday buy de that CL book BEI steal PFV
 'The book which I bought yesterday was stolen.'
 b. [fang zai zuozhi shang]_{RC} de na ben shu bei tou le
 put at table on de that CL book BEI steal PFV
 'The book which was put on the table was stolen.'

The referent *na ben shu* 'that book' in (16a) is grounded by connecting it with an identifiable discourse entity *wo* 'I' and the referent in (16b) is grounded by the locative *zai zuozhi shang* 'on the table'. Van Valin (1975) argues that the unmarked information status of locatives is accessible and that is the reason why they can serve to ground a referent. However, this does not mean any locative can be used to ground a referent. Generic locatives such as *anywhere*, *everywhere*, *the whole world*, and so on cannot be used to ground a referent since they are too broad to single out and identify a referent in the spatial world.

Besides grounding a referent, a RC can also serve to provide an assertion of a referent. In that case, the RC is used to characterize or describe a referent (Fox and Thompson 1990: 301).

17. a. The man [who I have for linguistics]_{RC} is really too much.
 b. There is a woman in my class [who is a nurse]_{RC}

Other than grounding, the RC in (17a) serves to ground the referent *the man* through the link with the accessible referent *wo* 'I'. The RC in (17b), on the other hand, does not supply any accessible referent or locative to ground the unaccessible referent *a woman*. Instead, it provides characterizing information revealing her profession.

⁷Grounding by a RC is called *anchoring* in Prince (1981).

For the purpose of this study, we only identify two groups of RCs: grounding RCs and non-grounding RCs. The former serves to provide an accessible referent or a locative expression to ground its head NP while the latter provides neither of them. The distinction of grounding RCs and non-grounding RCs allows us to make the following proposals regarding the ordering of Chinese multiple RCs:

18. a. Grounding RCs tend to take precedence over non-grounding RCs.
- b. If all the prenominal RCs are either grounding RCs or non-grounding RCs, they may order freely among themselves when modifying the same head NP.

With the two proposals in order, we now proceed to examine the data in Larson and Takahashi (2007) and the counterexamples mentioned in Section 2. Section 2 shows that the claim that stage-level RCs always precede individual-RCs is untenable. The relevant examples are repeated below.

19. a. [Wo renshi]_{RC1}de [xie le yi ben shu]_{RC2} de na ge xuesheng hen congming.
 I know de write PFV one CL book de that CL student very smart
 ‘The student whom I know who wrote a book is very smart.’ (Good)
- b. [Xie le yi ben shu]_{RC1} de [wo renshi]_{RC2} de na ge xuesheng hen congming.
 write PFV one CL book de I know de that CL student very smart
 ‘The student who wrote a book whom I know is very smart.’ (Odd)

The two RCs in (18) fulfill different discourse functions. The RC *worenshi* ‘I know’ helps to ground the head NP *na ge xuesheng* ‘that student’ by linking the head NP to an accessible referent *wo* ‘I’. By contrast, the RC *xie le yiben shu* ‘wrote a book’ does not provide any accessible referent to ground the head NP, but provides additional information regarding the achievement of the referent denoted by the head NP. According to our proposal in (18a), the grounding RC *wo renshi* ‘I know’ should precede a non-grounding RC *xie le yi-ben shu* ‘wrote a book’. In order to test the validity of proposal (18a), let’s examine the following sentences in Larson and Takahashi (2007).

20. a. [Wo zuotian kanjian]_{RC1} de [hui shuo yidaliyu]_{RC2} de ren shi Lisi.
 I yesterday see de can speak Italian de person is Lisi
 ‘The person whom I saw yesterday who can speak Italian is Lisi.’ (Good)
- b. [Hui shuo yidaliyu]_{RC1} de [wo zuotian kanjian]_{RC2} de ren shi Lisi.
 can speak Italian de I yesterday see de person is Lisi (Bad)
 ‘The person whom I saw yesterday who can speak Italian is Lisi.’

We agree with their judgment that (20a) is acceptable whereas (20b) is not. According to our proposal (18a), the contrast in (20) is expected. The first RC in (20a) is a grounding

RC which contains an accessible referent *wo* ‘I’ to ground the referent *ren* ‘person’. Thus, proposal (18a) correctly predicts that the grounding RC *wo zuotian kanjian* ‘(whom) I saw yesterday’ should be placed before the non-grounding RC *hui shuo yidaluyu* ‘(who) can speak Italian’.

After showing how proposal (18a) is superior to the analysis by Larson and Takahashi, we now turn to examine the validity of proposal (18b).

21. a. Na ge [meiyou lai]_{RC1} de [hen xihuan shang ke]_{RC2} de xuesheng shi Lisi.
 that CL not come de very like attend class de student is Lisi
 ‘The student who did not come who loves school is Lisi.’ (Good)
- b. Na ge [hen xihuan shang ke]_{RC2} de [meiyou lai]_{RC1} de xuesheng shi Lisi
 that CL very like attend class de not come de student is Lisi
 ‘The student who loves school who did not come is Zhagsan.’ (Good)

The RC *meiyou lai de* ‘(who) did not come’ is a stage-level RC whereas the RC *hen xihuan shang ke* ‘(who) loves school’ is an individual-level RC. Larson and Takahashi (2007) wrongly predict that the word order in (21a) is the only acceptable one. However, the 16 native speakers we have consulted all agreed that (21a) and (21b) are equally acceptable. Our proposal can easily account for the grammaticality of the sentences in (21). The stage-level RC *meiyou lai* ‘(who did not come)’ contains no grounding expression and the same is true with the individual-level RC *hen xihuan shangke* ‘(who) loves school’. Thus, proposal (18b) correctly predicts that the two non-grounding RCs show no ordering preference. Proposal (18b) can also easily explain the contrast in (22) in which the ordering of two stage-level RC cannot be shuffled.

22. a. [Wo zuotian kanjian]_{RC1} de [bei daibu]_{RC2} de na ge xuesheng hen piaoliang.
 I yesterday see de BEI arrest de that CL student very pretty
 ‘The student whom I saw yesterday who was arrested is very pretty.’ (Good)
- b. [Bei daibu]_{RC2} de [wo zuotian kanjian]_{RC1} de na ge xuesheng hen piaoliang.
 BEI arrest de I yesterday see de that CL student very pretty
 ‘The student who was arrested whom I saw yesterday is very pretty.’ (Odd)

Both RCs in (22) express stage-level properties which are predicted to order freely in Larson and Takahashi (2007). The unacceptability of the word order in (22b) again shows their analysis has to be refined to accommodate the contrast in (22). Our proposal (18b), in contrast, correctly predicts that the only acceptable order is (22a) because only the RC *wo zuotian kanjian* ‘(whom I saw yesterday)’ is a grounding RC because it contains a grounding accessible referent *wo* ‘I’. The other RC *bei daibu* ‘(who) was arrested’ should follow the grounding RC because it does not have any grounding information in it. After examining the ordering of two stage-level RCs in (22), we

proceed to discuss the ordering of two individual-level RCs in Larson and Takahashi (2007) to examine the validity of our proposals.

23. a. [Hui shuo yidaliyu]_{RC1} de [xihuan qu yinyuehui]_{RC2} de ren shi Lisi.
 can speak Italian de like go concert de person is Lisi
 ‘The person who can speak Italian who likes to go to concert is Lisi.’ (Good)
- b. [Xihuan qu yinyuehui]_{RC2} de [hui shuo yidaliyu]_{RC1} de ren shi Lisi.
 like go concert de can speak Italian de person is Lisi
 ‘The person who can speak Italian who likes to go to concert is Lisi.’ (Good)

Both RCs (23) are individual-level RCs because the RC *hui shuo yidaliyu* ‘(who) can speak Italian’ describes the ability of a person and the RC *xihuan qu yinyuehui* reveals a habitual activity of that person. Larson and Takahashi (2007) correctly predict their ordering is not subject to any constraint. Our proposal can also account for the ordering of the two RCs in (23). Examination of the two RCs in (23) shows that neither of them contains grounding element in them. In other words, both of them are non-grounding RCs and that is the reason why they do not show any ordering preference.

Lastly we show the ordering of adj-de-N structures. According to Sproat & Shih (1988, 1991) and Duanmu (1998), A-de-N structures should be analyzed as RCs. Interestingly, the ordering of the adjectives in the A-de-N structure is not subject to any restriction regardless whether they are individual-level adjectives or stage-level adjectives.

24. a. meili de congming de na ge xuesheng
 beautiful de smart de that CL student
 ‘that beautiful and smart student’
- b. congming de meili de na ge xuesheng
 smart de beautiful de that CL student
 ‘that smart and beautiful student’
25. a. pibei de jichanglulu de na ge xuesheng
 tired de hungry de that CL student
 ‘that tired and hungry student’
- b. jichanglulu de pibei de na ge xuesheng
 hungry de tired de that CL student
 ‘that hungry and tired student’
26. a. meili de jichanglulu de na ge xuesheng
 beautiful de hungry de that CL student
 ‘that beautiful and hungry student’
- b. jichanglulu de meili de na ge xuesheng
 hungry de beautiful de that CL student
 ‘that hungry and beautiful student’

The adjectives in (24) and (25) are individual-level adjectives and stage-level adjectives respectively whereas the two adjectives in (26) exhibit different properties. One adjective *meili* ‘beautiful’ is an individual-level adjective and the other one *jichanglulu* ‘hungry’ is a stage-level adjective. The acceptability of all the expressions from (24) to (26) indicates that adjectives in *de*-modification structures show no ordering preference. Obviously the theory by Larson and Takahashi (2007) cannot accommodate the ordering of the adjectives from (24) to (26). In contrast, our proposals can easily explain the ordering patterns of the adjectives. If the assumption that adjectives in *de*-modification structures should be analyzed as RCs is correct, our proposals correctly predict that the ordering of adjectives in *de*-modification structures is random since none of the adjectives contain any grounding element. As non-grounding RCs, the ordering of adjectives in the *de*-modification structures is predicted to be random.

To summarize, our proposals can not only explain the data in the study by Larson and Takahashi (2007), but also accommodate the data which challenges it. To further test the reliability of our proposals, in the next section we examine the ordering of multiple Chinese RCs in a large corpus, the Lancaster Corpus of Mandarin Chinese (LCMC).

4. Multiple RCs in LCMC

While the proposals in (18) can account for the full range of examples in both English and Chinese, one may complain about the reliance on purely linguistic intuitions and judgment. Anyway, different native speakers might have slightly different intuitions regarding the acceptability of a certain linguistic expression and it is not unusual to find that the same sentence might be judged as perfect by some native speakers and be frowned upon by others. So in this section, we will confront the proposals in (18) to the corpus data and if necessary go back to step one for possible refinement of the proposals regarding the ordering of multiple RCs.

The Chinese multiple RCs for analysis in this section are selected from a publicly available Chinese language corpus, the Lancaster Corpus of Modern Chinese (LCMC) (McEnery et al. 2003). LCMC is a one-million-word balanced corpus of written Mandarin Chinese, which consists of five hundred 2,000-word samples of written Chinese texts selected from fifteen text categories published in Mainland China around 1991. It provides a web-based concordance search functionality, which greatly facilitates this research. The concordance results from LCMC always come with a complete sentence where the searched word occurs. Careful examination of LCMC generates 31 relevant examples which are detailed in the following table, where the two or more RCs modifying the same head NPs are coded along the grounding/non-grounding and individual/stage-level dimensions.

The inspection of the ordering of multiple RCs shows that a total of 27 tokens observe the ordering pattern of grounding RC-non-grounding RC and only 4 tokens

follows the order of non-grounding RC-non-grounding RC. The ordering patterns thus fit nicely into our proposals in section 3.

Table 1: The ordering of Multiple RCs in LCMC

#	RC1	RC2	RC3	#	RC1	RC2
1	NGRC/STRC	NGRC/SRC		2	GRC/STRC	NGRC/IRC
3	GRC/STRC	NGRC/IRC		4	GRC/IRC	NGRC/IRC
5	GRC/STRC	NGRC/STRC		6	GRC/STRC	NGRC/STRC
7	GRC/STRC	NGRC/STRC		8	GRC/STRC	NGRC/STRC
9	GRC/STRC	NGRC/STRC		10	GRC/STRC	NGRC/IRC
11	GRC/STRC	NGRC/IRC		12	GRC/STRC	NGRC/IRC
13	GRC/STRC	GRC/STRC		14	GRC/STRC	NGRC/IRC
15	GRC/STRC	NGRC/STRC		16	GRC/STRC	GRC/STRC
17	GRC/STRC	NGRC/IRC	NGRC/IRC	18	GRC/STRC	NGRC/IRC
19	GRC/STRC	NGRC/IRC		20	GRC/STRC	NGRC/IRC
21	GRC/STRC	NGRC/IRC		22	NGRC/IRC	NGRC/IRC
23	GRC/IRC	NGRC/IRC		24	GRC/STRC	NGRC/STRC
25	GRC/STRC	NGRC/STRC		26	NGRC/IRC	NGRC/IRC
27	GRC/STRC	NGRC/IRC		28	GRC/STRC	NGRC/IRC
29	NGRC/IRC	NGRC/IRC		30	GRC/STRC	NGRC/IRC
31	GRC/IRC	NGRC/IRC				

STRC= Stage-level Relative Clause; IRC= Individual-level Relative Clause; NGRC=Non-grounding Relative Clause; GRC=Grounding Relative Clause

It should be noted that the ordering patterns observed in the above table also seem to be consistent with the Larson and Takahashi's (2007) account, as stage-level RCs always precede individual-level RCs. However, their account falls short on explaining the

examples involving two RCs of the same property (i.e., when they are both individual-level RCs or both stage-level RCs). According to Larson and Takahashi (2007), multiple pronominal RCs exhibiting the same properties order freely among them. That is, multiple pronominal stage-level RCs can be ordered in random sequence and the same is said to be true for multiple pronominal individual-level RCs. A careful inspection of the data shows that a grounding stage-level RC always take precedence before a non-grounding RC (e.g., #5, #6, #7, #8, #9, #13, #15, #16, #24, #25) and the same observation extends to multiple pronominal individual-level RCs (see #4, #23, #31). We provide one relevant example each from the LCMC to illustrate the ordering of stage-level RCs (27a) and that of individual-level RCs (27b) below. The first RCs in these two examples are both grounding RCs, because they provide the grounding referent *wo* ‘I’ in (27a) and a locative *zai qiangjiao* ‘at the corner’ in (27b), respectively. According to our proposal (18a), they occur before the second RCs which do not contain a grounding element.

27. a. [wo zai meiguo xie]_{RC1} de [gang chuban]_{RC2} de gongcheng kongzhi lun
 I in US write de just publish de engineer control theory
 ‘*Theory on Engineering Control*’ which I wrote in US which was just published’
 (LCMC A)
- b. [fang zai qiang jiao]_{RC1} de na ba [yong gaoliang gan kunza]_{RC2} de saozhou
 put at wall corner de that CL use sorghum stem tie de broom
 ‘The broom which is in the corner of the wall which is made of sorghum stem’
 (LCMC K)

Furthermore, if pronominal stage-level RCs could be freely ordered as Larson and Takahashi (2007) suggest, we would expect that some non-grounding stage-level RCs can occur before grounding stage-level RCs. However, not a single instance where a non-grounding stage-level RC occurs before a grounding stage-level RC can be found. The same observation also applies to the ordering of multiple pronominal individual-level RCs as shown in #4, #23, #31. In conclusion, our proposal provides a better account for the ordering of multiple pronominal RCs irrespective of whether they are stage-level RCs or individual-level RCs.

Our proposal can not only better predict the ordering of Chinese multiple RCs, but also can be extended to explain the ordering of English multiple RCs. As opposed to Larson and Takahashi (2007) who claim that there is no ordering restriction for multiple RCs, Fox and Thompson (1990) believe that the ordering of English multiple RCs follows a predictable fashion. Let’s examine all the multiple RCs in Fox and Thompson (1990) to test whether our analysis can be extended to English data.

28. a. And there was one thing [they said in article] [that was really intriguing]
 b. There was something [we needed] [which was really obscure]
 c. He claims that there’s kind of stuff [you make] [that has a pear in it]

- d. Well, this little, this other little atomic clock [that I have] [that used to be in the bathroom]
- e. Cause the one [I got in my office][I got for three]

An examination of the data in (28) shows that the first RC contains an accessible referent in the subject position of the RC to ground the head NP and that its ordering nicely conforms to our proposal (18a) that grounding RC precedes non-grounding RCs⁸.

Finally, our proposals in (18) can be extended to the well-established observation that restrictive RCs always precede non-restrictive ones in English (Larson and Takahashi 2007). A restrictive RC in English is to identify and single out the referent denoted by the head NP. In other words, a restrictive RC serves to ground the referent of the head NP. On the contrary, a non-restrictive in English helps to describe the head NP without providing any grounding information. Therefore, our proposal correctly predicts that a restrictive RC which fulfills the function of a grounding RC should be placed before a nonrestrictive RC which provides additional information of the head NP. It's not clear how the account by Larson and Takahashi (2007) could handle this pattern of ordering among restrictive and nonrestrictive RCs.

4. Concluding remarks

In this paper we show that grounding RCs always precede non-grounding RCs in Chinese as well as in English. In addition, we show that distinguishing grounding

⁸ All the head NP in (28) are non-human. As for human head nouns, Fox and Thompson (1990) believe that the ordering of multiple RCs is different. For example, the first RC in the following sentence does not have a grounding referent.

There was a boy [that played the trombone]_{RC1} [that he kind of knew]_{RC2}

It seems that the ordering of the two RCs here constitutes a counterexample to our proposal. Fox and Thompson (1990), however, argue that human heads are grounded differently from non-human heads. Non-human heads are typically grounded by humans who own them, use them, and manipulate them and the grounding humans are typically given. On the contrary, human heads are grounded by “their own activities, that is, to early predicates” (Fox and Thompson: 309). In other words, the first RC in the above sentence can serve as a grounding RC although it does not contain any grounding given referent. However, the Chinese data shows that human heads behave like nonhuman heads in the sense that RCs containing a grounding referent always occur before RCs which lack a grounding referent. Whether Chinese differs from English in terms of the animacy of the heads merits further research.

modifiers from non-ground modifiers allows us to offer an elegant explanation of the ordering of the adjectives in *de*-modification structures, e.g., A-*de*-N in Chinese. The question then arises as to why multiple RCs are ordered that way. Our tentative answer to that is that this ordering pattern pertains to the information flow in the discourse. Since given information tends to be deployed before new information (Clark & Clark, 1978; Ming & Chen, 2010), it is reasonable to assume that a linguistic unit containing a given referent should occur before a linguistic unit which lacks any given information. By analogy, a grounding RC should be placed before a non-grounding RC because the former contains a given referent but the latter lacks any given information, that is, any given referent. Future research is needed to test the validity of the above assumption.

REFERENCES

- Clark, Eve V.; and Herbert H Clark. 1978. Universals, relativity, and language processing. In J. H. Greenberg (Ed.), *Universals of human language*, Vol. I., 225-277. Stanford: Stanford University Press.
- Fox, Barbara. A.; and Sandra A. Thompson. 1990. A discourse explanation of the grammar of relative clauses in English conversation: *Language* 66.297-316.
- LaPolla, Randy J. 1995. Pragmatic relations and word order in Chinese. *Word Order in Discourse*, ed. by Pamela Dowing and Michale Noonan, 297-329. Amsterdam and Philadelphia: John Benjamins.
- Larson, Richard. K., and Naoko Takahashi. 2007. Order and interpretation in prenominal relative clauses. *Proceedings of the Workshop on Altaic Formal Linguistics II. MIT Working Papers in Linguistics Vol. 54.* eds. By M. Kelepir and B. Öztürk, 101-120. Cambridge, MA: MITWPL.
- McEnery, Anthony; Zhonghua Xiao; and Lili Mo. 2003. Aspect marking in English and Chinese: Using the Lancaster Corpus of Mandarin Chinese for contrastive language study. *Literary and Linguistic Computing* 18.361-378.
- Ming, Tao, and Liang Chen. 2010. A discourse-pragmatic study of word order variation in Chinese relative clauses. *Journal of Pragmatic* 42.168-189.
- Paul, Waltraud. 2005. Adjectival modification in Mandarin Chinese and related issues. *Journal of linguistics* 43.757-793
- Prince, Ellen. (1981). Toward a taxonomy of given-new information. *Radical Pragmatics*, ed. By Peter Cole, 223-255. New York and London: Academic Press.
- Sproat, Richard, and Chilin Shih. 1988. Prenominal adjectival ordering in English and Mandarin. *Proceedings of NELS 18*, 465-489. Amherst, MA: GLSA.
- Sproat, Richard, and Chilin Shih. 1991. The cross-linguistic distribution of adjective ordering restrictions. *Interdisciplinary approaches to language: Essays in honor*

- of S.-Y. Kuroda*, ed. by C. Georgopoulos and R. Ishihara, 565-592. Dordrecht: Kluwer.
- San, Duanmu. 1998. Wordhood in Chinese. *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*, ed. by Jerome Packard, 135-196. Berlin, New York: Mouton de Gruyter.
- Van Valin, Robert. 1975. *Existential locatives in Mandarin*. Presented at the Eighth International Conference on Sino-Tibetan Languages and Linguistics, Oct. 24-26, University of California, Berkeley.

A Corpus-based Analysis of Semantic Type System of Event Nouns: A Case Study on *huìyì*

Shan Wang, Sophia Lee, Chu-Ren Huang
Department of Chinese and Bilingual Studies
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

Event nouns are a special type of nouns which lexically encode eventive information, including event structure and time. In this paper, we discuss a typical event noun in Chinese, *huìyì*, to find out the evidences of *huìyì* as an event, and examine the three levels of semantic representation system, i.e. argument structure, event structure and qualia structure, based on the generative lexicon theory.

1. Introduction

Typical nouns are usually used to represent concrete or abstract objects, such as *chuānghu* ‘windows’ and *sīxiǎng* ‘thoughts’. However, there is a kind of nouns which lexically encode eventive information, including event structure and time. These nouns are called event nouns, temporal nouns or process nouns. They are a particular type in Mandarin Chinese. In this study, we use the term event nouns, because an “event” not only has temporal properties, but also includes activities happening in it.

The aims of this research are to find out the evidences of *huìyì* ‘conference’ as an event, and reveal its argument structure, event structure and qualia structure to establish its semantic representation system based on the generative lexicon theory (henceforth GLT).

GLT was first proposed in Pustejovsky (1991) and further developed in Pustejovsky (1995). Its goal is to capture the generative nature of lexical creativity and sense extension phenomena. GLT is a computational system composed of four levels of representation: (I) Argument Structure: the number and type of logical arguments; (II) Event Structure: event type and subeventual structure; (III) Qualia Structure: comprised of formal, constitutive, telic and agentive roles; (IV) Lexical Inheritance Structure: position of a lexical structure in the type lattice.

We adopt GLT as the theoretical basis of this study because it provides a comprehensive account of lexical semantic information of a lexical item. Qualia structure is the most important contribution of GLT which shows the semantic information of a lexical item.

Lexical inheritance structure is using a very shallow ontology in GTL, thus it will not be discussed in this paper.

The data in this study are collected from the tagged Chinese Gigaword Corpus (Second Edition)¹ using Chinese Word Sketch Engine². It includes three distinct international sources of Chinese newswire, e.g. Central News Agency, Xinhua News Agency, and Zaobao Newspaper, with a total of 2,803,632 documents.

2. Related Research

There are various studies that focus on the patterns and properties of events nouns. Ma (1995) points out that typical nominals and predicates are related to denotation meaning and predication meaning respectively, and they are located on both ends of a continuum. Atypical nominals and predicates locate in the middle of the continuum, including nouns with predication meaning and verbs with denotation meaning.

He investigates the combination of nouns and localizers. Localizers are divided into three types:

A(spatial): *zuǒ* ‘left’ / *yòu* ‘right’ / *dōng* ‘east’ / *xī* ‘west’ / *lǐmiàn* ‘inside’ / *yǐwài* ‘outside’

B(temporal): *yǐqián* ‘before’ / *yǐhòu* ‘after’ / *dāngér* ‘when’

C(spatial or temporal): *qián* ‘before’ / *hòu* ‘after’ / *shàng* ‘on’ / *xià* ‘under’ / *zhōng* ‘middle, in the course of’

- Common nouns are words that can be followed by type A, but can not be followed by type B: such as *qiáo* ‘bridges’, *mén* ‘doors’, *yǐzi* ‘chairs’
- Process nouns are words that can be followed by type B, but can not be followed by type A: such as *zhànzhēng* ‘wars’, *léiyǔ* ‘thunderstorms’, *dàhuì* ‘large conferences’
- Both common nouns and process nouns can be followed by type C, but the former shows spatial meaning, while the latter shows temporal meaning.

Process nouns can be further divided into continuous and non-continuous ones according to whether they can be followed by *zhōng* ‘in the course of’. Continuous process nouns are able to collocate with durative temporal expressions, e.g. *zhànzhēng* ‘wars’, *léiyǔ* ‘thunderstorms’, while non-continuous ones are in collocation with punctual temporal expression, e.g. *jīntiān* ‘today’, *yuándàn* ‘New Year's Day’.

Secondly, process nouns can combine with activity classifiers, for examples:

¹ <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T03>

² <http://wordsketch.ling.sinica.edu.tw/>

yī cì huìyì/zhànzhēng/chēhuò/shìgù/shǒushù (a conference/a war/a car accident/an accident/a surgery)

yī cháng yǔ/xuě/fēng/bīngbáo/bìng/xì/zhànzhēng (rain/snow/wind/hail/a disease /a drama/a war)

yī dùn fàn ‘a meal’

Chu (2000) proposes three formats in identifying event nouns as shown below:

(I) N + *qī jiān* ‘during’

(II) N + *qián* ‘before’/*hòu* ‘after’/*yǐqián* ‘before’/*yǐhòu* ‘after’/*zhīqián* ‘before’/*zhīhòu* ‘after’

(III) N + *shíqī/shídài* ‘epoch’, including N + *shíhou* ‘during’, N + *niándài* ‘era’, N + *jiēduàn* ‘period’/*shíguāng* ‘time’

Type I and type III look at the time flow of these nouns from the internal aspect, including starting point, continuity, and ending point. Type II looks at the time flow of these nouns from the external world. According to the above three formats that nouns can enter into, they are divided into five classes:

Class A can only enter type II, including:

- A1: monosyllabic words, e.g. *bìng* ‘diseases’, *fàn* ‘meals’, *huì* ‘conferences’, *kè* ‘lessons’, *shì* ‘things’, *shuāng* ‘frost’, *yǔ* ‘rain’, *xuě* ‘snow’
- A2: nouns whose last morpheme is *qī* ‘period’, e.g. *huìqī* ‘conference period’, *xùngqī* ‘flood period’, *wēixiǎnqī* ‘dangerous period’, *jiàqī* ‘holidays’
- A3: activities or events with a short duration, e.g. *shǎndiàn* ‘lightening’, *kōng nà* ‘aircrashes’, *bàizhàng* ‘lost battles’, *shìgù* ‘accidents’, *cǎn’àn* ‘tragedies’, *chēhuò* ‘car accidents’
- A4: participants in an event, e.g. *wǔfàn* ‘lunch’, *zǎochá* ‘morning tea’, *zǎodiǎn* ‘breakfast’, *yèxiāo* ‘late snacks’, *zuòwén kè* ‘writing classes’

Class B can only enter type III. Most words represent human beings or things, e.g. *tángtàizōng shíqī* ‘the period of Emperor Taizong of Tang’, *sānyèchóng shídài* ‘the era of trilobites’.

Class C can enter any two types, including:

- C1 can only enter I and II. One type is representing activities, and is longer than A3, e.g. *shǒushù* ‘operations’, *wǎnyàn* ‘dinners’, *dàhuì* ‘large conferences’, *línyù* ‘showers’. The other type is words with *jià* ‘holidays’ as a morpheme, e.g. *hánjiǎ* ‘winter vacation’, *hūnjiǎ* ‘marriage leave’, *bìngjià* ‘sick leave’.
- C2 can only enter I and III, e.g. *dàxué* ‘universities’, *chūzhōng* ‘junior high schools’.
- C3 can only enter II and III. These are nouns representing people, works and dynasties.

Class D can enter I, II, and III. They represent activities, events or natural phenomena, e.g. *gémìng* ‘revolutions’, *zhànzhēng* ‘wars’, *yùndòng* ‘sports’, *jiāoyì* ‘trades’, *shìbiàn* ‘incidents’, *bǐsài* ‘competition’, *léiyǔ* ‘thunderstorms’, *táifēng* ‘typhoons’, *rìshí* ‘eclipses’.

Class E cannot enter any of the three formats.

With the above analysis, Chu (2000) sum up the temporality of nouns with a sequence from strong to weak: ①according to classes: D > C > A > B > E; ②according to semantics: activities, events or phenomena nouns > human or object nouns, concrete nouns > non-concrete nouns, humanity nouns > natural nouns, distinctive feature nouns > non-distinctive feature nouns, internal temporal nouns > external temporal nouns.

Wang & Zhu (2000) propose a way of subcategorizing nouns according to the classifiers that nouns can be in collocation with. Process nouns can only be modified by activity classifiers or temporal classifiers, e.g. *yī chǎng yǒuyì sài* ‘a friendly competition’, *yī zhèn léiyǔ* ‘a burst of thunderstorms’, *yī dùn wǎncān* ‘a dinner’, and *shí nián nèizhàn* ‘a decade of civil wars’.

Han (2010) proposes three formats to identify event nouns: (a) numeral + verbal classifier + ___; (b) ___ + *hòu* ‘after’; (c) ___ + *zhōng* ‘in the course of’.

Format I is the best frame and Format II is a preferred frame in determining event nouns. They have general application to prototype event nouns. Format III is just an aiding frame in determining event nouns.

From the above mentioned research, we know that a consensus that a noun followed by *qián* ‘before’ or *hòu* ‘after’ expresses temporal meaning, the noun is an event noun. Also, activity classifiers can be used to detect event nouns. In section 3, we will follow the two formats to check *huìyì*.

However, there are some problems with previous research: Ma (1995) regards temporal nouns, e.g. *jīntiān* ‘today’ and *yuándàn* ‘New Year’s Day’, as event nouns. In our study, we hold that these nouns only have temporal properties, but do not have event properties. So we do not treat them as event nouns. For Chu (2000), we argue that A2(nouns whose last morpheme is *qī*), C1(nouns with *jià* as a morpheme) and C2(e.g. *dàxué* ‘universities’, *chūzhōng* ‘junior middle school’) should belong to the same category, because they all purely express time. For this reason, they should have strongest time adaptation, instead of words of class D.

Previous studies have given us a picture of the properties of event nouns. Some questions, however, remain unexplored. In the following sections, we will discuss the three problems: (1) Why *huìyì* can be treated as an event? (2) What is *huìyì*’s semantic

representation, e.g. argument structure, event structure and qualia structure? (3) How to represent *huìyì*'s semantic type system?

3. Evidences of *huìyì* as an Event

The word *huìyì* has two basic senses: one is a formal meeting for consultation or discussion, the other is an organization that usually discusses and deals with important issues, such as *zhōngguó rénmin zhèngzhì xiéshāng huìyì* (Chinese People's Political Consultative Conference). *huìyì* on the first sense is regarded as an event, while the second is not.

In this section, we will use four formats to prove that *huìyì* is an event noun.

3.1 *huìyì* + *qián* 'before' / *hòu* 'after'

(Chu 2000; Han 2010; Liu 2004; Ma 1995) all hold that when a noun followed by *qián* 'before' or *hòu* 'after' expresses temporal meaning, the noun is an event noun. The following examples show that *huìyì* can collocate with *qián* 'before' and *hòu* 'after', which indicate *huìyì*'s boundaries and distinguish it from other events. In (1), before the conference event, there is another event that the stock market is rising. In (2), after the conference event, there will be an event of studying an issue.

(1) jīngjì huìyì qián gǔshì jiāng chíxù shàngzhǎng 。

Economy conference before stock market will continue rise
'The stock market will continue rising before the economy conference.'

(2) zhè cì huìyì hòu , jiù yào zhuòshǒu yánjiū zhè gè wèntí 。

This classifier conference after, should will set about study this CL issue
'After this conference, (we) should set about studying this issue.'

3.2 *huìyì* + Durative Temporal Expression

huìyì can be modified by durative temporal expression as shown in (3). It means that *huìyì* is an event with process.

(3) jīntiān de huìyì jìnxíng le qī gè duō xiǎoshí 。

Today DE conference go on ASP seven CL more hour
'Today's conference went on for over seven hours.'

This process can be interrupted, for example:

(4) huìyì měi jìnxíng liǎng gè xiǎoshí jiù yǒu yī cì cháxiē 。

Conference every go on two classifier hour at once have one CL tea break
'There is a tea break every two hours during the conference.'

Also, this process has starting point and ending point:

(5) běn cì huìyì cóng liùyuè qīrì kāishǐ 。

this CL conference from June seventh start

‘This conference starts from June 7th.’

(6) zhōng fǎ gōngyè hézuò huìyì yuánmǎn jiéshù 。
China France industry cooperation conference satisfactory end

‘Sino-French industrial cooperation conference comes to a successful ending.’

3.3 Numeral/Demonstrative + Event Classifier + *huìyì*

Huang et al. (1997) distinguish event classifiers from activity classifiers. The former is used to enumerate the number of events, e.g. *yī bǐ shēngyì* ‘a deal of business’; the latter is used to enumerate the times of action, e.g. *qù le yī tàng* ‘went once’. Moreover, the former usually appear before the enumerated event; the latter usually appear after the enumerated action. Huang & Ahrens (2003) further point out that an event classifier can coerce an event reading from a noun.

Huang et al. (1997) list 35 event classifiers: *bō* ‘of staggered event’, *bān* ‘of shift, scheduled flight/bus etc.’, *bǐ* ‘of transaction’, *bù* ‘step (event procedures)’, *pào* ‘a brewing (of tea etc.)’, *pán* ‘a serving round (of a dish)’, *mù* ‘act (of a play)’, *fān* ‘times (of a repeated event)’, *dào* ‘of dishes of procedures’, *dàng* ‘duration of a run (of play, movie etc.)’, *duàn* ‘section (of play etc.)’, *dùn* ‘the process of a meal’, *tái* ‘a run of a traveling troupe’, *táng* ‘a class’, *tàng* ‘a journey’, *tōng* ‘a phone call’, *lún* ‘a round’, *huí* ‘a roundtrip’, *jié* ‘a class, a session’, *jiè* ‘an annual event’, *jiàn* ‘event’, *jú* ‘game’, *qī* ‘term’, *qǐ* ‘event (especially a happening, an accident)’, *quān* ‘round (of majong)’, *xí* ‘lecture’, *zhé* ‘an act (in a Chinese play)’, *zhèn* ‘one of a sporadic event(s)’, *zhuāng* ‘event’, *chǎng* ‘a (scheduled) event (with beginning and ending)’, *chū* ‘a play’, *rèn* ‘term (of a termed position)’, *zōng* ‘trade/transaction’, *cān* ‘a meal’, *cì* ‘once (re. frequency of event)’.

huìyì can combine with four of them: *lún*, *jiè*, *chǎng*, *cì*. Besides, *huìyì* can also combine with *dù* ‘frequency of an event’, which is an event classifier similar to *cì*.

dù can directly modify *huìyì* to enumerate the times of the conference, which shows that *huìyì* is an event noun, For example:

(7) měi nián yī dù de zhōng rì jīngjì mào yì huìyì
every year one CL DE China Japanese economy trade conference
‘the annual meeting of the Sino-Japanese economic and trade conference’

(8) yī nián yī dù de yà tài jīng hé huìyì
one year one CL DE Asia Pacific economy cooperation conference
‘the annual meeting of the Asia-Pacific Economic Cooperation conference’

Table 1 depicts the frequency and saliency when *huìyì* combines with event classifiers in tagged Chinese Gigaword (Second Edition). It shows that *cì* and *jiè* are the two most salient event classifiers of *huìyì*, followed by *chǎng*, *lún* and *dù*.

Table 1: Event Classifiers the modify *huìyì*

Event Classifiers	Frequency	Saliency	Examples
<i>cì</i> ‘once (re. frequency of event)’	94477	100.13	xiǎozǔ jiāng zài běnzhōu wǔ zài zhàokāi yī <u>cì</u> <i>huìyì</i> 。 ‘This group will convene another meeting on Friday.’
<i>jiè</i> ‘an annual event’	12471	65.05	běn <u>jiè</u> <i>huìyì</i> tāolùn zhǔtí wéi gōngyè jí kējì hézuò 。 ‘The discussion themes of this conference are industrial and technological cooperation.’
<i>chǎng</i> ‘a (scheduled) event (with beginning and ending)’	660	19.78	zuìhòu yī <u>chǎng</u> <i>huìyì</i> zhǔyào tàntǎo zázhi de fāzhǎn 。 ‘The final session of this conference is to probe into the development of journals.’
<i>lún</i> ‘a round’	185	17.27	sàimòdá de hǎigǎng gōngsī guòqù 10 gè yuè zhàokāi le jǐ <u>lún</u> gāocéng <i>huìyì</i> 。 ‘Seaport Terminal of Syed Mokhtar held several rounds of high-level meetings in the past 10 months.’
<i>dù</i> ‘frequency of an event’	50	1.48	shuāngfāng zhàokāi guò shù <u>dù</u> dāngjúzhě <i>huìyì</i> 。 ‘The two sides held authority meetings several times.’

3.4 *huìyì* + *hé* ‘and’ / *jí* ‘and’ / *huò* ‘or’

If *huìyì* is an event, then the word that coordinates with it should be an event as well. Such coordination is connected by *hé* ‘and’, *jí* ‘and’ or *huò* ‘or’. For examples, *huìyì hé qìngdiǎn* ‘conferences and celebrations’, *huìyì jí shāngzhǎn* ‘conferences and trade shows’, *huìyì huò jiǎngzuò* ‘conferences or seminars’. Since *qìngdiǎn* ‘celebrations’, *shāngzhǎn* ‘trade shows’, *jiǎngzuò* ‘seminars’ are all events, *huìyì* is correspondingly an event.

4. Argument Structure of *huìyì*

4.1 *huìyì* can Assign Argument Roles

In this section, we argue that *huìyì* can assign argument roles to its participants. The reasons are:

First, many verbs in collocation with *huìyì* are light verbs which do not contain much content and do not assign participant role, for examples, *zuò* ‘作, make’, *zuòchū* ‘作出, make’, *zuò* ‘做, make’, *zuòchū* ‘做出, make’. So we need to look for it somewhere else, for instance, from the event noun.

(9) jīntiān de huìyì méiyǒu zuòchū rènhé juéyì。
today DE conference not have make any resolution
'Today's meeting did not make any resolution.'

Second, typical verbs, such as *tōngguò* 'pass', *xuānbù* 'announce', *xiūgǎi* 'revise', *shēnchá* 'examine' can assign argument structure.

(10) huìyì tōngguò/xuānbù/xiūgǎi/shēnchá le juéyì
conference pass/announce/revise/examine ASP proposal
'The conference passed/announced/revised/examined the proposal.'

(11) zhōngdōng héping huìyì jǔxíng zàijí。
Middle East peace conference hold soon
'Middle East peace conference is to be held soon.'

However, *huìyì* itself does not meet the selectional requirement of these words. Instead, the participant role specification of these words comes from the qualia structure of *huìyì*. It is the people who attend the conference *pass/announce/revise/examine* the proposal. Thus these words are the telic role of *huìyì*.

4.2 Argument Types of *huìyì*

Pustejovsky (1995) distinct four types of arguments: true arguments, default arguments, shadow arguments, true adjuncts. True arguments are syntactically realized words, e.g. *Julia* came. Default arguments are words that participate in the logical expression in the qualia, but they do not need to be expressed syntactically, e.g. Bill made the table *out of rosewood*. Shadow arguments are words that are lexicalized into the lexical item and they can only be expressed by subtyping or discourse specification; otherwise, they will lead to pleonasm. For example, Evan kicked the door *with his left leg*. True adjuncts are words that modify the logical expression, but are not fastened to any lexical item's semantic representation, e.g. Lucy goes shopping *in Harbor City*. In the following section, we will show that *huìyì* takes two default arguments as core arguments, and takes a variety of true adjuncts as modification.

4.2.1 Default Arguments of *huìyì*

huìyì is a kind of event attended by some people with common interests on a particular topic. Thus interlocutors and a topic are the arguments of *huìyì*. They have the following characteristics:

First, both of them are default arguments, which are not necessarily expressed syntactically. However, even when they do not appear syntactically, they are presupposed by the *huìyì* event and can usually be found before or after the sentence that does not contain them.

(I) Interlocutors

(12) běiyuēzǔzhī zhàokāi tèbié shǒunǎo huìyì 。 (expressed syntactically)
NATO hold special head conference
‘NATO held a special summit.’

(13) huìyì yǐ yuánzé tōngguò zài wèilái shí nián nèi jiàn shè
conference already principle pass in future ten year within construct
[sānxiá gōngchéng] de juédìng 。 (not expressed syntactically)
Three Gorges Project DE resolution
‘The conference has passed in principle the resolution of constructing Three Gorges Project in the next ten years.’

(II) Topic

(14) zhōng jiā gāoděngjiàoyù huìyì zài jiānádà jǔxíng 。 (expressed syntactically)
China Canada high education conference in Canada hold
‘Sino-Canada higher education conference is held in Canada.’

(15) jīntiān de huìyì yuándìng zhōngwǔ jiéshù 。 (not expressed syntactically)
today DE conference scheduled noon finish
‘Today’s conference was scheduled to finish at noon.’

Second, they are encoded into the *huìyì* event and are indispensable from it. One can hardly imagine any conference without people or a topic. Therefore they are the core arguments of *huìyì*.

Thirdly, the interlocutors of *huìyì* have these properties:

(I) *huìyì*’s interlocutors must be more than one person. One person speaking without any audience is called monologue, not conference. *huìyì* has an encoded action of discussion, which need not less than two participants.

(II) position, identity, country, place, and organization are always used to represent the interlocutors.

(a) Position

(16) shíyóushūchūguóuzǔzhī bùzhǎng huìyì zài rìnèiwǎ jǔxíng 。
OPEC minister conference in Geneva hold
‘The minister conference of OPEC will be held in Geneva.’

(b) Identity

(17) yuànshì huìyì měi liǎng nián zhàokāi yī cì 。

academician conference every two year hold one CL
'The academician conference is held every two years.'

(c) Country or Place

(18) nánhán xīwàng jīnnián qiūtiān ānpái nánběihán jǔxíng lìshǐxíng de
South Korea hope this year fall arrange south and north Korea hold historic DE
gāofēng huìyì 。
summit conference
'South Korea hoped that this fall there can be an arrangement of a historic summit between
South and North Korea.'

(d) Organization

(19) xià cì dǒngshìhuì huìyì jiāng huì jìnsù jǔxíng 。
next CL board conference will should ASAP hold
'Next board conference will be held as soon as possible.'

4.2.2 True Adjuncts of *huìyì*

huìyì takes temporal and spatial modification, which are true adjuncts of *huìyì*.

(I) Time: *niánchū* 'the beginning of a year', *niánzhōng* 'the end of a year', *qùnián* 'last year', *zuìjìn* 'recently', *jīnrì* 'today', *rìqián* 'a few days ago', *bàngwǎn* 'in the evening', *yuèdǐ* 'the end of a month', *xiàjì* 'the summer'

(II) Place: *chǎngsuǒ* 'sites', *huìchǎng* 'conference venues', *chǎngdì* 'venues', *zhùdì* 'station', *huìzhǐ* 'the site of a conference', *jiùzhǐ* 'former site', *dìzhǐ* 'address', *dìdiǎn* 'location', *suǒzài* 'location'

5. Event Structure of *huìyì*

Pustejovsky (1995) distinguishes three situation types: state, activity and event, where the last class is divided into accomplishment, and achievement events. In this section, we introduce the situation types of *huìyì*, reveal the aspectual shift of it, and figure out some tests for *huìyì*'s situation types. We also notice that some sentences can express more than one situation type.

5.1 Basic Situation Types of *huìyì*

5.1.1 State

A state has no internal structure, and it can hold for certain time. *huìyì* itself is a bare noun. When it appears in the subject position, and followed by an adjectival predicate, it describes a state and has a generic reading. For example:

(20) *huìyì* hěn yǒuqù 。
conference very interesting
'This conference is very interesting.'

5.1.2 Activity

An activity describes an unbounded dynamic situation. When a sentence is describing the process of *huìyì*, then it expresses activity. For example:

- (21) yàzhōu fángkòng qínligǎn jǐnjí huìyì zhèngzài mǎngǔ jǔxíng 。
 Asia prevent bird flu emergent conference now Bangkok hold
 ‘Asia bird flu emergent prevention conference is being held in Bangkok.’

5.2 Aspectual Shift of *huìyì*

5.2.1 Words that Lead to Aspectual Shift of *huìyì*

It is well known that for verbs, their aspectual properties may change due to some factors. For instance, Pustejovsky (1995) finds that adverbial modification (durative and frame), the structure of the NP in an argument position (e.g., definite vs. bare plural), or a prepositional phrase can lead to aspectual shift. Smith (1990) notices that a situation type is affected by constellation of verb, complements and adverbial. Smith (1994) points out that in Mandarin Chinese certain resultative verbal complements, such as *wán*, *hǎo*, *zháo*, *chéng* convey the completion of an accomplishment.

In this section, we show that *huìyì* is always shifted to be (interpreted as) an accomplishment when it is delimited by demonstratives, localizers or quantifiers.

5.2.1.1 Demonstratives

- (22) nà cì huìyì de yī gè jiéguǒ jiùshì jiànli le guójì huòbì jījīn zǔzhī 。
 that CL conference DE one CL result is establish ASP international currency organization
 ‘One result of that meeting is the establishment of the International Monetary Fund.’

- (23) zhè cì huìyì hěn chénggōng 。
 this CL conference very successful
 ‘This conference is very successful.’

5.2.1.2 Localizer

- (24) shàng cì huìyì yú 2003 nián 7 yuè zài èguāduōěr shǒudū jīduō jǔxíng 。
 last CL conference on 2003 year 7 month in Ecuador capital Quito hold
 ‘Last conference is held in Quito, the capital city of Ecuador, in July 2003.’

- (25) xià cì huìyì jiāng zài míngnián chū jǔxíng 。
 next CL conference will in next year early hold
 ‘The next meeting will be held early next year.’

- (26) zhè cì huìyì zài qián cì huìyì zhǔtí de jīchǔ shàng, yòu
 this CL conference on last CL conference theme DE basis on, again
 zēngjiā le 3 gè fùzhǔtí 。
 add ASP 3 CL sub-theme
 ‘This conference added three sub-themes based on the theme of last conference.’

5.2.1.3 Quantifiers

(27) zài yī cì huìyì shàng, wǒ yùjiàn tā 。
on a CL conference on, I meet him
'On a conference, I met him.'

5.3 Tests for Different Situation Types

5.3.1 Test for States

When *huìyì* appears in the format “N+(很 *hěn* ‘very’)+adj”, it is a state.

5.3.2 Entailment of Activity Sentences

Activity is a homogenous event. Any part of it is the same as other parts. For instance:

(28) *huìyì kāi le wǔ gè xiǎoshí le* 。 → *huìyì kāi le sān gè xiǎoshí le* 。
conference hold ASP five CL hour ASP → conference hold ASP three CL hour ASP
'The conference has been on for five hours.' entails 'The conference has been on for three hours.'

5.4 Ambiguity

We noticed that a sentence or expression with *huìyì* is sometimes underspecified. In such a case, it expresses more than one situation types. For instance:

(29) zhōngguó zuòxié zhòngqìng *huìyì* hěn yǒu qù 。
China writer association Chongqing conference very interesting
'China Writer Association's conference in Chongqing is very interesting.'

There is no tense marking on verbs in Mandarin Chinese. When there is no temporal expression in a sentence, the event time is not expressed. In such a case, we can understand the above sentence as an event that is just finished, then *huìyì* here is accomplishment; or we can understand *hěn yǒuqù* 'very interesting' as a regular property of *zhōngguó zuòxié zhòngqìng huìyì* 'China Writer Association's conference in Chongqing', then *huìyì* here is state.

Another example is *zhè cì huìyì* 'this conference'. It can mean the conference has not started, is ongoing or was just finished. Thus it can be activity or accomplishment.

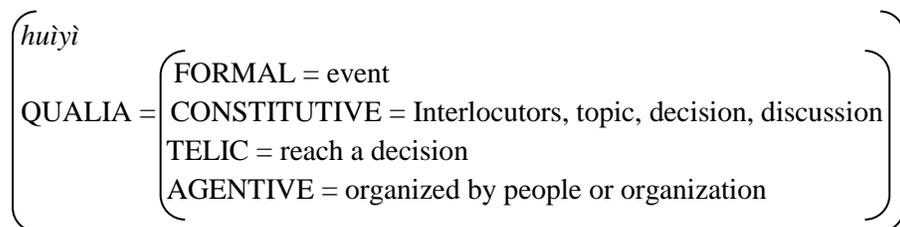
zhè jǐtiān de huìyì 'the several days' conference' can mean an ongoing conference or a just completed conference. Consequently, it is activity or accomplishment.

6. Qualia Structure

Pustejovsky (1995) shows how lexical items encode semantic information in the qualia structure. There are four roles in a qualia structure, and each is associated with some values. (I) The constitutive role is about the relation between an object and its constituents or parts. Its role values include material, weight, parts and component elements. (II) The formal role can distinguish the object within a larger domain. Orientation, magnitude, shape, dimensionality, color, and position are its role values. (III)

The telic role is about the purpose and function of the object. (IV) The agentive role describes factors involved in the origin of an object, such as creator, artifact, natural kind, causal chain.

huìyì is a formal event on which some people discuss a topic with the aim of reaching a decision. Consequently, the qualia structure of *huìyì* is illustrated below:



huìyì as a head and a modifier behaves differently. First, the frequency and saliency of nouns appear before and after *huìyì* is quite different as depicted in Table 2. About 2/3 more nouns tend to appear before *huìyì* than after it. Also, nouns before *huìyì* are more salient than after it.

Table 2. Frequency and Saliency of Nouns before and after *huìyì*

No.	Relation	Frequency		Saliency	Examples
		Tokens	percentage		
1	N_Modifier	358601	35.38%	1.6%	<i>shǒunǎo huìyì</i> ‘conferences of heads of a state or government’, <i>gāofēng huìyì</i> ‘summit meetings’
2	Modifies	129350	12.76%	0.6%	<i>huìyì qījiān</i> ‘during the conference’, <i>huìyì jìyào</i> ‘minutes of a conference’

Second, the same noun seldom appears both before and after *huìyì*. Out of all nouns modifying *huìyì* or as heads of *huìyì*, only two can appear in both pre- *huìyì* and post-*huìyì* position. The two nouns are *lǐngxiù* ‘leaders’ and *zhǔxítuán* ‘presidiums’. By intuition, there should be many such nouns, but according to Gigaword corpus data, it is not.

Third, even when the same noun appears both before and after *huìyì*, the meaning can be different. This can be shown in the contrast between 领袖会议 and 会议领袖. The former is ambiguous: it can refer to a conference about leaders or a conference attended

by hosts; the latter refers to the role of the participant. Also, the former must be more than one person while the latter can be only one person.

In the following section, we will explore the qualia modification to *huìyì* when it is the head noun, and *huìyì*'s modification to other nouns when it is the modifier.

6.1 *huìyì* as the Head

When *huìyì* is the head of a structure, its eventive information decides what kind of modification it gets. We find that *huìyì* usually gets formal and constitutive modification when it is the head.

6.1.1 Formal Modification

$$\left(\begin{array}{l} \text{dàxíng huìyì 'large scale conferences'} \\ \text{QUALIA = [FORMAL = scale]} \end{array} \right)$$

In *dàxíng huìyì* 'large scale conferences', *dàxíng* 'large scale' points to the scale of the *huìyì* event, which is the formal role of *huìyì*. Similarly, *dà* 'big', *xiǎo* 'small', *xiǎoxíng* 'small scale', *shèngdà* 'grand' all refer to the scale of a conference.

6.1.2 Constitutive Modification

According to Pustejovsky (1995), constitutive role expresses the relation between an object and its parts. *huìyì* is not a typical object, but a kind of event involving a series of activities with a particular topic and a group of people, which are inseparable parts of *huìyì*, just as a CPU is inseparable from a computer.

Here we extend constitutive role to include the relation between an event and its participant roles. Consequently, in *jīngjì huìyì* 'economy conferences' and *yuànrshì huìyì* 'academician conferences', economy and academicians are regarded as the constitutive role of *huìyì*.

$$\left(\begin{array}{l} \text{jīngjì huìyì 'economy conferences'} \\ \text{QUALIA = [CONSTITUTIVE = economy as the topic]} \end{array} \right)$$

$$\left(\begin{array}{l} \text{yuànrshì huìyì 'academician conferences'} \\ \text{QUALIA = [CONSTITUTIVE = academicians as Interlocutors]} \end{array} \right)$$

Topic: *shāngwù* 'business', *guówù* 'national affairs', *guóshì* 'national affairs', *tiānwén* 'astronomy', *qìhòu* 'climate', *hǎishì* 'maritime affairs'

Interlocutors: *qīnshǔ* ‘relatives’, *zōngqīn* ‘clans’, *zhǎnglǎo* ‘presbyteries’, *zhǔjiào* ‘bishops’, *shǔjī zhǔjiào* ‘cardinal-bishops’, *jiàoshī* ‘priests’, *zhàojí rén* ‘convenors’

There are some words which modify the encoded content aspect of the constitutive role of *huìyì* as shown from (a) to (e). (a)Duration: *rǒngcháng* ‘tediously long’, *mǎlāsōng shì* ‘marathon-style’, *jiǎnduǎn* ‘short’, *duǎnzàn* ‘brief’. *huìyì*, as an event, has duration aspect encoded in its meaning, so these words modifies the constitutive role of *huìyì*. (b)Manner: *zhòngdà* ‘significant’, *pángdà* ‘huge’, *kuòdà* ‘enlarged’, *biékāishēngmiàn* ‘break new ground’, *zhèngshì* ‘formal’, *fēizhèngshì* ‘informal’, *zhǔmù* ‘fix one's eyes upon’, *xīn* ‘new’, *yīliú* ‘first-class’, *yánsù* ‘serious’, *jǐnjí* ‘urgent’, *zhùmíng* ‘well-known’, *gōngkāi* ‘public’. (c)Frequency: *mìjí* ‘intensive’, *yìliánchuàn* ‘series of’. (d)Influence: *huà shídài* ‘epoch-making’, *jìwǎng kāilái* ‘to carry on the past and open a way for future’, *shēnyuǎn* ‘far-reaching’. (e)Result: *chénggōng* ‘successful’, *yuánmǎn* ‘satisfactory’.

6.2 *huìyì* as the Modifier

When *huìyì* is the modifier, it usually modifies the telic and agentive role of the head noun.

6.2.1 Telic Modification

$$\left(\begin{array}{l} \textit{huìyì jìniànpǐn} \text{ ‘conference souvenirs’} \\ \text{QUALIA} = \left[\text{TELIC} = \text{conference use} \right] \end{array} \right)$$

jìniànpǐn are souvenirs for memory. When it is modified by *huìyì*, its function is specified as an item for the memory of a conference. Similar examples are *zhǔxí tái* ‘podiums’, *dàlóu* ‘buildings’, *dàshà* ‘buildings’, *dàtīng* ‘halls’, *tīngtáng* ‘halls’, *bàngōngshì* ‘offices’, *jiǎngtán* ‘platforms’, *shāpán* ‘sand tables’, *zhuānxiàn* ‘direct lines’, *zhǔzhóu* ‘spindles’, *qǐngjiǎn* ‘invitations’, *shǒucè* ‘manuals’, *xūkězhèng* ‘licenses’, *jìniànpǐn* ‘souvenirs’, *cáiliào* ‘materials’

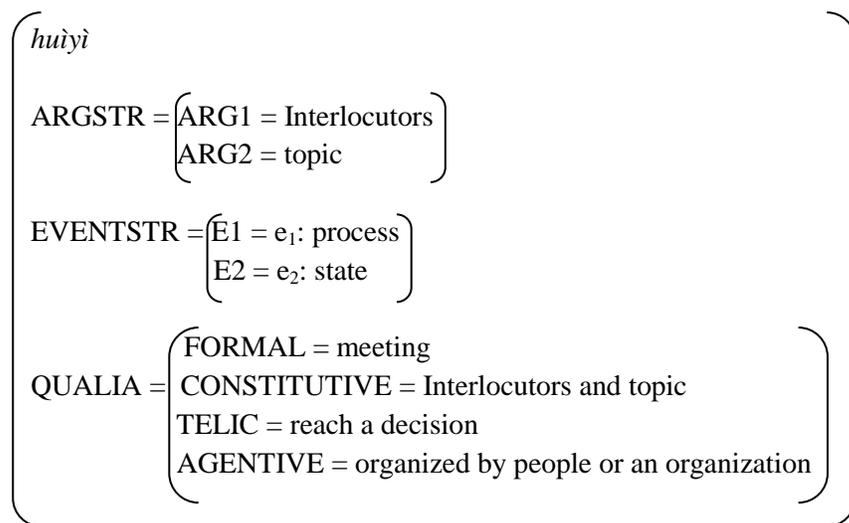
6.2.2 Agentive Modification

$$\left(\begin{array}{l} \textit{huìyì juéyì} \text{ ‘conference resolution’} \\ \text{QUALIA} = \left[\text{AGENTIVE} = \text{source} \right] \end{array} \right)$$

juéyì is typically a kind of resolution reached after discussion. In *huìyì juéyì* ‘conference resolution’, *huìyì* states the origin of the resolution.

Similar examples are *chǎngmiàn* ‘scenes’, *chéngguǒ* ‘achievements’, *shìfēi* ‘right and wrong’, *huǒshí* ‘food’, *wǎncān* ‘dinners’, *xiāoxi* ‘messages’, *xīnwén* ‘news’.

To conclude, through the above exploration, we have revealed the argument structure, event structure and qualia structure of *huìyì* and established the semantic type system of it, which is illustrated below.



7. Conclusion

In sum, this paper gives a detailed semantic analysis of the event noun *huìyì* based on GLT. First, it finds that *huìyì* takes two default arguments: Interlocutors and two default arguments: time and space. Second, the basic event types of *huìyì* are state and activity. Furthermore, *huìyì* can go aspectual shift and become accomplishment when delimited. Third, *huìyì*'s qualia structure and its qualia modification as the head and the modifier are explored. Finally, we expanded the constitutive qualia role of GLT to include an event and its activities.

Acknowledgements: The authors would like to thank Prof. Dingxu SHI, Prof. Haihua PAN, Jing JIN and Jing YANG for their constructive comments on *Linguistic Society of Hong Kong-Annual Research Forum 2010*. The remaining errors are ours.

REFERENCES

- Chu, Zexiang. 2000. *An Investigation on Nouns' Adaptation to Temporality (名词的时间适应性情况考察)*. Modern Chinese Grammar Studies that Face the Challenges of New Century: the International Conference on Modern Chinese Grammar, 1998

- (面临新世纪挑战的现代汉语语法研究：'98 现代汉语语法学国际学术会议论文集), ed. by J. Lu. Jinan: Shandong Education Press.
- Han, Lei. 2010. Analysing the Word Class Status of Event Nouns (试析事件名词的词类地位). *Journal of Ningxia University (Humanities & Social Sciences Edition)* (1).
- Huang, Chu-Ren & Kathleen Ahrens. 2003. Individuals, Kinds and Events: Classifier Coercion of Nouns. *Language Sciences* 25 (4).353-373.
- Huang, Chu-Ren, Keh-Jiann Chen & C.-X Lai. 1997. Mandarin Chinese Classifier and Noun-Classifier Collocation Dictionary. Taipei: Mandarin Daily Press.
- Liu, Shun. 2004. A Study of Temporality of Common Nouns (普通名词的时间性研究). *Language Teaching and Linguistic Studies* (4).
- Ma, Qingzhu. 1995. *Verbs with denotational Meaning and Nouns with Statement Meaning(指称义动词和陈述义名词)*. Research and Exploration of the Grammar (语法研究和探索). Beijing: The Commercial Press.
- Pustejovsky, James. 1991. The generative lexicon. *Computational Linguistics* 17 (4).409-441.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Smith, Carlota S. 1990. Event types in Mandarin. *Linguistics* 28 (2).309-336.
- Smith, Carlota S. 1994. Aspectual viewpoint and situation type in Mandarin Chinese. *Journal of East Asian Linguistics* 3 (2).107-146.
- Wang, Hui & Xuefeng Zhu. 2000. *Subcategorization and Quantitative Research on Modern Chinese Nouns (现代汉语名词的子类划分及定量研究)*. Modern Chinese Grammar Studies that Face the Challenges of New Century:the International Conference on Modern Chinese Grammar,1998 (面临新世纪挑战的现代汉语语法研究：'98 现代汉语语法学国际学术会议论文集). Jinan: Shandong Education Press.

A statistical argument for the homophony avoidance approach to the disyllabification of Chinese

Wen Jin

Department of Chinese Studies, National University of Singapore

Under the homophone avoidance (HA) theory (Guo 1928, Wang 1944, Karlgren 1949, Lü 1963, Li and Thompson 1981), monosyllabic words are disyllabified to avoid homophonous ambiguities. Lü 1963 predicts that the more monosyllabic homophones there are in a language, the more likely disyllabic words would be created. Duanmu 1999 argues against the HA approach and claims that no supporting evidence has been found in Chinese. This paper argues for the HA approach and provides supporting evidence from corpora of Mandarin Chinese and Cantonese. Our discoveries support the HA motivation for the disyllabification of Chinese. Additionally, the HA theory applies cross-linguistically. The HA theory has interesting implications about the disyllabification of Chinese from a diachronic perspective, which are supposed to accompany the simplification of syllable structures in archaic Chinese.

1 Introduction

It has been widely accepted that Chinese monosyllabic words were disyllabified to avoid ambiguities of interpretation that would otherwise have arisen because of homophony (Guo 1928, Wang 1944, Karlgren 1949, Lü 1963, Li and Thompson 1981). For example, both ‘wood’ and ‘to shampoo’ are pronounced [mu⁵¹] in Mandarin. In order to avoid ambiguities of interpretation caused by this pair of monosyllabic homophones, we use [mu⁵¹ tou] ‘wood’ to contrast with [mu⁵¹ y⁵¹] ‘to shampoo’ (Duanmu 2007:152). Similar phenomena are observed cross-linguistically (see e.g., Arcodia 2007, Kaplan 2010, Baerman 2011). For example, in Teiwa, a Papuan language spoken in the Alor Island of Indonesia, possessed nouns may be prefixed for person and number of a possessor. A CV-prefix is attached to C-initial noun stems and the vowel of the prefix is deleted when it is attached to V-initial noun stems. The prefixes of 1SG and 1PL are *na-* and *ni-*, respectively. Thus, both the 1SG and 1PL forms of the V-initial stem *-uar wa* ‘ear’ would expectedly be *n-uar wa*. However, to avoid homophony the 1PL form of *-uar wa* is *ni-uar wa*, which contrasts with the 1SG form *n-uar wa*. See Baerman 2011 for a detailed discussion.

Lü 1963 predicts that the more monosyllabic homophones there are in a language, the more likely disyllabic words would be created. For example, Mandarin is expected to

have more disyllabic words than Cantonese, which has more syllable types and therefore fewer monosyllabic homophones than Mandarin. Cross-linguistically, if a language has more syllable types, it should have more monosyllabic words than a language with fewer syllable types, because monosyllabic homophones can be reduced by an increase in syllable types.

Duanmu 2007 argues that homophony avoidance (HA) did not play a clear role in the increase of disyllabic words in Chinese. Instead, a large number of disyllabic words were introduced into Chinese either because they were polysyllabic names in the first place or because they consisted of two or more morphemes in the source language such as Japanese. He further argues that word length in Chinese is restricted by metrical constraints (Duanmu 2007: 172). Duanmu claims that no supporting evidence for the HA approach has been found in Chinese (Duanmu 2007:154).

Moreover, Feng 2000 argues that the disyllabification in Chinese arose because of the unmarkedness constraint FOOT-BINARY, which requires a foot to consist of two syllables cross-linguistically. Chinese words are mostly disyllabic because they satisfy this constraint, assuming that a word consists of only one foot. Feng 2000 therefore predicts that disyllabic words would predominate cross-linguistically.

I argue in favor of Lü 1963 and show that the HA approach plays a significant role in the disyllabification of Chinese. Additionally, I argue that the disyllabification of Chinese arose mainly because of HA, and the so-called ‘minimal word’ phenomena are mainly by-products of HA.

This paper argues for the HA approach and provides comparative evidence from the corpora of Mandarin, Cantonese, American English and Japanese. Mandarin has about 1,300 types of syllables (Lin and Wang 1992) while Cantonese has 1,795 ones (Kao 1971) given that Cantonese has more contour tones than Mandarin. The HA theory predicts that Cantonese should have more monosyllabic words than Mandarin because Cantonese has more syllable types, which makes disyllabification less necessary, assuming that the main function of disyllabification is to avoid homophones, which will cause ambiguities of interpretation.

The HA theory further predicts that any language with a smaller number of syllable types tends more to have multisyllabic words, because otherwise its syllable types would not suffice to differentiate words. This paper compares Chinese to English and Japanese. English has complex syllable structures while the syllable structure of Japanese is much simpler compared to Chinese. As a consequence, English is expected to have more monosyllabic words than Chinese, while Japanese is expected to have fewer monosyllabic words. The predictions of the HA theory are borne out by our statistics.

The format of this paper is as follows. Section 2 discusses the syllable structures of both Mandarin Chinese and Cantonese, and compares Mandarin Chinese with Cantonese in terms of syllable types. We show that Cantonese has more monosyllabic words than Mandarin as predicted by our HA approach. Section 3 discusses the syllable structures of English and Japanese, and compares them with Chinese in terms of syllable types. We

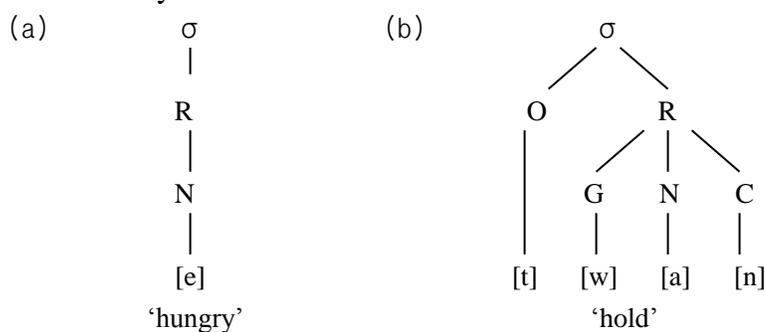
show that our HA approach makes right predictions again. Section 4 compares the HA theory with two alternative accounts such as Duanmu 1999, 2007 and Feng 2000. We also show that the HA theory has interesting implications about the disyllabification of Chinese from a diachronic perspective, which were supposed to accompany the simplification of syllable structures in archaic Chinese. Section 5 concludes and discusses some residual issues.

2 The Homophone Avoidance theory and Mandarin and Cantonese syllables

In this section, we discuss both Mandarin and Cantonese syllable types. We show that Cantonese has more syllable types than Mandarin. The HA theory predicts that Cantonese should therefore have more monosyllabic words than Mandarin because it is less necessary for Cantonese to undergo disyllabification to avoid homophones. This prediction is proved by our statistics.

A full Mandarin Chinese syllable (σ) has been traditionally considered to consist of an onset (O) and a rhyme (R). A rhyme consists of a glide (G), a nucleus (N) and a coda (C). Mandarin Chinese does not allow complex onsets and only nasal consonants such as [n] and [ŋ] can occur as a coda. The simplest Mandarin Chinese syllable consists of a nucleus only, e.g., [e⁵¹] ‘hungry’; see (1a). By contrast, a complex one consists of an onset, a glide, a nucleus and a coda, e.g., [tʰwɑn⁵⁵] ‘hold’; see (1b).¹

(1) Mandarin Syllable Structures



Mandarin Chinese has 21 different onsets and 37 different rhymes. Putting contour tones aside, there are 410 syllable types in Mandarin Chinese because some syllable types such as *[fi], *[ki] do not exist (Lin and Wang 1992). If we take into consideration the four lexical tones in Mandarin Chinese, there are about 1,300 syllable types given that lexical tones do not freely combine with a syllable. For example, only three tones can attach to [san]: [san⁵⁵] ‘three’, [san²¹⁴] ‘umbrella’, and [san⁵¹] ‘distribute’.

¹ The syllable structure of Mandarin is controversial. See Bao 1996 for a comprehensive review of possible syllable structures in Chinese.

coda. Cantonese also allows obstruent codas such as [p, t, k] and a nasal coda [m], e.g., [pik] ‘must’, [kip] ‘suitcase’ (Kao 1971: 142). Additionally, in Cantonese a nasal can independently act as a syllable, for example, [m¹³] ‘not’, [ŋ³³] ‘noon’. Kao 1971 shows that Cantonese has 1,795 syllable types; see also Duanmu 1999.

The HA theory predicts that it would be more likely for Mandarin Chinese to use other strategies such as disyllabification to avoid ambiguities of interpretation. The reason is that if both Mandarin and Cantonese used monosyllabic words only to express the same amount of meanings, there would be more monosyllabic homophones in Mandarin, which would result in ambiguities of interpretation, since Cantonese has more syllable types than Mandarin. For example, both ‘beer’ and ‘leather’ are pronounced [p^hi³⁵] in Mandarin, but in Cantonese ‘beer’ is pronounced [pe⁵⁵] while ‘leather’ is pronounced [p^hei¹³]. Native speakers of Mandarin must say [p^hi³⁵. teju²¹⁴] ‘beer’ to avoid ambiguity of interpretation due to homophony while Cantonese speakers still use the monosyllabic form [pe⁵⁵] in colloquial speech. As a consequence, the HA theory further predicts that Cantonese should have more monosyllabic words than Mandarin.

We present several types of statistical evidence to show that Cantonese has more monosyllabic words than Mandarin, which proves the predictions of the HA theory. Based on the corpora created in 1959 by Zhongguo Wenzhi Gaige Weiyuanhui Yanjiu Tuiguang Chu [Chinese Language Reform Committee Research and Popularization Office] (ZGWG hereafter), monosyllabic words amount to 29% of all the 3,624 words in the corpora. The corpora show that disyllabic words predominate in the vocabulary of modern Chinese. He and Li 1987 and ZGWG 2008 get similar results. According to Li and Bai 1987 and Yu 1993, there are few monosyllabic neologisms in modern Mandarin.

We calculated the number of monosyllabic words in Cantonese, based on a list of words drawn from various Cantonese textbooks. Our statistics shows that the ratio of monosyllabic words in Cantonese is 34.7%, see Table 1.

Table 1 monosyllabic words in Mandarin and Cantonese (%)

	Total	Monosyllabic	%	Language
ZGWG (1959)	3624	1046	29	Mandarin
He and Li (1987)	3000	809	27	Mandarin
ZGWG (2008)	3000	1000	33.3	Mandarin
Cantonese textbooks	2291	796	34.7	Cantonese

It seems that in Table 1 the ratio of monosyllabic words in Mandarin calculated by ZGWG 2008 (33.3%) is pretty close to that in Cantonese (34.7%). But a closer look will tell more difference. ZGWG 2008 gives a list of 56,008 commonly used words, which includes 3,181 monosyllabic words (5.7%), 40,351 disyllabic words (72.0%), 6,459 tri-syllabic words (11.5%), 5,855 quadri-syllabic words (10.5%), and 126 longer words (0.2%). A majority of its 3,000 most frequently used words are function words, which

tend to be short cross-linguistically. If we put aside function words and only calculate the ratios of monosyllabic lexical words (nouns, verbs, adjectives, adverbs) in both Mandarin and Cantonese, we can see that the ratio of monosyllabic words in Cantonese (31.3%) is much higher than that in Mandarin (25.5%); see Table 2.

Table 2: Monosyllabic lexical words (%)

	Total	Monosyllabic	%monosyllabic	Language
ZWGW(2008)	2479	633	25.5	Mandarin
Cantonese Textbooks	2047	642	31.4	Cantonese

Our statistics prove the prediction of the HA theory that Cantonese has more syllable types and therefore more monosyllabic words than Mandarin. If we compare the ratio of the number of Mandarin (M) syllable types divided by that of Cantonese (C) syllable types to the ratio of the number of Mandarin monosyllabic lexical words (Wds) divided by that of Cantonese monosyllabic lexical words, we can see the proximity of the two ratios ($p > 0.05$); see (2). This shows that syllable types play a clear role in determining the length of words and the necessity to resort to disyllabification.

(2) Syllable types and monosyllabic lexical words in Mandarin and Cantonese

$$1) \frac{M-\sigma \text{ types}}{C-\sigma \text{ types}} = \frac{1300}{1795} = 72.4\%$$

$$2) \frac{M\text{-monosyllabic lexical Wds \%}}{C\text{-monosyllabic lexical Wds \%}} = \frac{25.5\%}{31.4\%} = 81.2\%$$

$$3) \frac{M-\sigma \text{ types}}{C-\sigma \text{ types}} \approx \frac{M\text{-monosyllabic lexical Wds \%}}{C\text{-monosyllabic lexical Wds \%}}$$

In the vocabulary of *New Cantonese Today* (2006), if we consider lexical words only, 41.4% of the monosyllabic Cantonese words have monosyllabic Mandarin glosses and the other monosyllabic Cantonese words correspond to disyllabic Mandarin words. In (3), the Mandarin sentence uses 7 syllables while the Cantonese one uses 5 syllables. Mandarin uses disyllabic forms while Cantonese uses monosyllabic forms to express the same meanings, e.g., *zen.me* vs. *med* ‘why’, *na.me* vs. *gem* ‘so’, *huang.miu* vs. *meo* ‘ridiculous’. See also Table 3, which shows that there are more monosyllabic words in Cantonese than in Mandarin based on *New Cantonese Today* (2006).

- (3) (a) *Zen.me ni na.me huang.miu ne?* (Mandarin)
 how 2Sg so ridiculous PRT
 ‘How can you be so ridiculous!’

(b) *Med neih gem meo ga?* (Cantonese)
 how 2Sg so ridiculous PRT
 ‘How can you be so ridiculous!’

Table 3: Percentage of monosyllabic words in *Xinbian Jinri Yueyu (2006)* [*New Cantonese Today*]

	Total	Number of monosyllabic words	Monosyllabic words %
Cantonese words	613	145	23.7
Mandarin glosses	613	60	9.8

We also asked twelve bilingual speakers of Mandarin and Cantonese to translate some commonly used Cantonese words into Mandarin. We obtained the same result that Mandarin tends less to use monosyllabic words. See Table 4.

Table 4. Mandarin vs. Cantonese

	Total	Monosyllabic words	Monosyllabic words %
Mandarin	1174	297	25.4
Cantonese	1174	388	33

3 The HA theory and cross-linguistic evidence

In this section, we show that the predictions of the HA theory in terms of syllable types and word length are borne out cross-linguistically. English, for example, has more syllable types than Mandarin and therefore more of its most frequently used words are monosyllabic. Japanese, by contrast, has fewer syllable types than Mandarin and therefore its most frequently used words tend to be multi-syllabic.

3.1 English syllable types and word length

English has 15 vowels and 24 consonants, and allows complex onsets and codas (Hammond 1999). The English syllable structure is (C)(C)(C)V(C)(C)(C)(C)(C)². Compared to Mandarin and Cantonese, English syllable structures can be much more complex. Additionally, English has supra-segmental features such as vowel length and stress. English has more than 10,000 different syllable types, many more than those in Mandarin. According to HA theory, English’s most frequently used words tend much more to be monosyllabic because the amount of English syllable types makes it less necessary for English words to be longer than those in Mandarin.

Consider the 5,000 most frequently used lemmas in the Word Frequency Lists and Dictionary, which was created on the basis of Brigham Young University’s Corpus of Contemporary American English (COCA). A lemma can be either lexical or functional. A

² An example of a five-consonant coda is *angsts* /'æŋksts/.

lexeme and its inflected forms belong to one lemma. For instance, *send*, *sending*, and *sent* belong to one lemma. As a result, the effects of inflectional morphology on word length cannot be considered. For example, if both the monosyllabic forms *send* and *sends* are among the most frequently used words, only one will be counted. To make our analysis consistent, we did not consider the English morphologically complex words of this lemma-corpus, either inflectional or derivational. If we put aside functional words in addition, we can see that monosyllabic words amount to almost half of the vocabulary; see Table 5.

Table 5: Length of English lexical words based on COCA³

	Noun	Verb	Adjective	Adverb	TOTAL	percentage
monosyllabic	836	474	169	77	1556	48.34%
disyllabic	669	313	181	48	1211	37.62%
tri-syllabic	197	77	73	20	367	11.40%
others	40	19	23	3	85	2.64%
total	1742	883	446	148	3219	100.00%

English has more monosyllabic words than Chinese because the complexity of English syllable structures helps reduce potential homophones in English and ambiguities of interpretation so that there is less necessity for English words to undergo disyllabification.

3.2 Japanese syllable types and word length

In Japanese there are 136 syllables in terms of the combinations of segments and moras (Tamaoka & Makioka 1987). As predicted by the HA theory, Japanese frequently used words should be longer than those in English and Chinese.

We calculated both the number and ratio of lexical words in terms of the number of syllables in each of them, based on a list of words drawn from various textbooks of basic-level Japanese. Our statistics show that tri-syllabic and quadri-syllabic words predominate in Japanese; see Table 6.

Table 6. Number and ratio of Japanese lexical words⁴

	mono-	di-	tri-	quadri-	quintuple-	sextuple-	septuple-	TOTAL
No.	12	183	419	353	78	39	1	1085
%	0.1%	16.9%	38.6%	32.5%	7.1%	3.6%		
*loanwords are not counted.								

³ We do not count 93 words with tricky syllabic structures in this words list, such as 'theory' /θi.ə.ri/ or /θiə.ri/

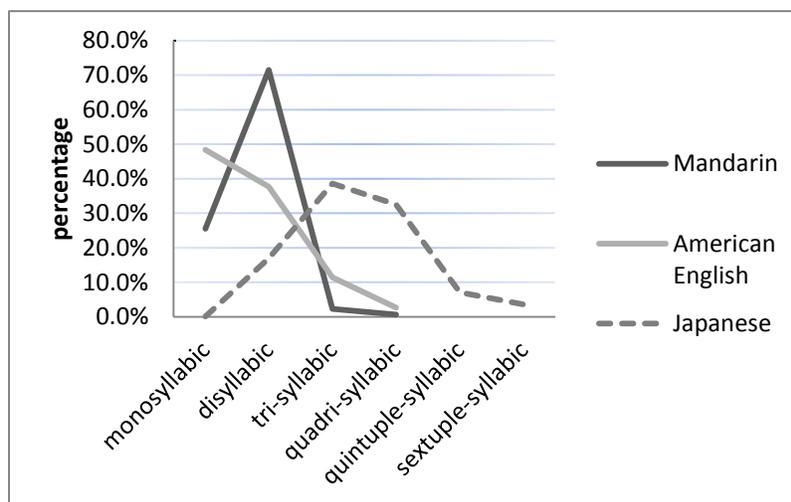
⁴ Loanwords are not counted because the number of syllable of loanwords in Japanese depend on the original word length of borrowing languages

Our results prove the predictions of the HA theory. As shown in Table 7, Japanese has much fewer mono-syllabic words than Mandarin and English because it has much fewer syllable types so that it is much more likely for Japanese to have mono-syllabic homophones. As a result, it is more likely for Japanese to refer to combinations of syllables to avoid potential ambiguities of interpretation. See also (4), which presents a diagram based on Table 7.

Table 7: Number and ratio of Mandarin, American English, and Japanese words

	mono-	di-	tri-	quadri-	quintuple-	sextuple-	septuple-	TOTAL
Mandarin	633	1772	56	16	2			2479
percentage	25.5%	71.5%	2.3%	0.6%				
English	1556	1211	367	85	1		1	3219
percentage	48.3%	37.6%	11.4%	2.6%				
Japanese	12	183	419	353	78	39	1	1085
percentage	0.1%	16.9%	38.6%	32.5%	7.1%	3.6%		

(4) A diagram of the number and ratio of Mandarin, American English, and Japanese words



Since our cross-linguistic evidence proves the predictions of the HA theory that the number of syllable types in a language will (partially) determine the length of words, we conclude that the HA theory plays an important role in the disyllabification of Chinese words.

4 Alternative accounts for the disyllabification of Chinese words

In this section, we discuss two alternative accounts for the disyllabification of Chinese, i.e., Duanmu 1999 and Feng 2000. Duanmu 1999 remarks that the HA theory does not play a clear role in the disyllabification of Chinese and implies that there is no correlation between syllable types and word length. He argues that Chinese has always had many disyllabic words, which can become monosyllabic depending on stress requirements. Feng 2000, on the other hand, refers to the constraint Foot-Binary to explain why disyllabic words predominate in Chinese. Feng's theory predicts that disyllabic words should predominate cross-linguistically. We argue that neither of them is correct.

4.1 Duanmu 1999, 2007

Duanmu 1999, 2007 remark that the increase in the disyllabic vocabulary of Chinese is not due to the simplification of its syllable structures. Duanmu 1999, 2007 argue against Lü's 1963 prediction that Cantonese should have more monosyllabic words than Mandarin, and remark that "no evidence for the prediction is offered." (Duanmu 2007: 154). Duanmu claims that "[the] HA approach does not play a clear role in the increase of disyllabic words in Chinese." (Duanmu 2007: 172) Duanmu's claim is incorrect because our results show that Cantonese does have more monosyllabic words than Mandarin. Additionally, our cross-linguistic evidence proves the correctness of the HA theory in terms of the relationship between syllable types and word length.

Duanmu 1999, 2007 argues that some monosyllabic items still remain monosyllabic though they can always cause ambiguity, and uses *ta* 'he', *ta* 'she' and *ta* 'it' as examples. However, as Mandarin is rarely regarded as having grammatical genders, these three pronouns may have the same underlying structure. Some other languages, such as Arabic, French, which are commonly regarded as having grammatical gender, each noun of these languages will belong to one of the genders. For example, the second pronoun 'you' in Arabic is distinguished by gender, an Arabic person will say *uhibbuka* 'I love you' to male, and *uhibbuki* 'I love you' to female. For the non-gender languages, such as Malay, Japanese, nouns in these languages do not belong to gender. In Malay, *Dia* can represent both 'he' and 'she'. No evidence shows that pronouns in Mandarin have been classified by genders. Thus, this example fail to argue against HA approach because there is no evidence offered against the third pronouns of Mandarin has three different underlying structures. Duanmu 1999, 2007 argues that most increase in disyllabic words comes from borrowings after the Opium War, whereby disyllabic (or longer) words are introduced either because they are polysyllabic names in the first place, or because they require two or more morphemes independent of the borrowing language (Duanmu 2007: 172). According to Duanmu 2007, most of the loanwords after the Opium War are borrowed from Japanese. We count for the loanwords from Japanese in *Chinese Loanwords Dictionary*, there are 853 loanwords from Japanese in total and 99% of them can be written in Kanji, which is similar to Chinese characters. Thus, there are two

methods of borrowing words from Japanese, borrowing phonetic features (5a), or borrowing Kanji (5b), while English and other Latin languages can only use the first method.

- (5) (a)
- | | | | | |
|------------------------------|-----------------|---|------------------------------|-----------------|
| Pronunciation in
Japanese | <i>o.ba.san</i> | | Pronunciation in
Mandarin | <i>o.ba.san</i> |
| Kanji | 老婦人 | → | Chinese
Characters | 欧巴桑 |
| jia.li.fo.ni.ya | ‘California’ | | | |
| niu.yue | ‘New York’ | | | |
- (b)
- | | | | | |
|------------------------------|------------------|---|------------------------------|---------------|
| Pronunciation in
Japanese | <i>bo.ku.shi</i> | | Pronunciation in
Mandarin | <i>bo.shi</i> |
| Kanji | 博士 | → | Chinese
Characters | 博士 |

According to our data, Kanji tend to occur in a binary form. As a result, most words loaned from Japanese are disyllabic. Chinese prefer Kanji borrowing to phonetic borrowing because disyllabic template has been formed as a byproduct of HA.

We do not deny the metrical approach (Duanmu 1999, 2007) when dealing with disyllabic words of Chinese. Metrical approach proposes that word length variation in Chinese is influenced by stress, which is determined cyclically by Nonhead Stress and the each foot must have two syllables (Duanmu 2007: 159), and word lengths are constrained by metrical structure, hence, in that some position prefer a disyllabic words and others prefer a monosyllabic word.

4.2 Feng 2000

Feng 2000 argues that Chinese is obliged to the constraint of FOOT-BINARY, which is the internal reason for the disyllabification of Chinese. From archaic Chinese to modern Mandarin, Chinese syllable has been simplified by a large scale. Feng 2000 argues that in Archaic Chinese, every syllable has two moras and forms a heavy syllable while the syllable structure becomes simpler in the modern times. This argument has been admitted by large amount of scholars (Ding 1979, Yu 1985, Yip 2002, Arcodia 2007, etc.), two of them are given in (6).

- (6)
Ding (1979) and Yu (1985):

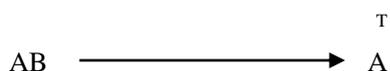
Early Archaic Chinese:	(C)C(C)(G)(G)(V)VC(C)
Middle Archaic Chinese: Zhou-Qin Dynasty	(C)(C)(G)(G)(V)V(C)
Wei and Jin Accent to Middle Chinese:	(C)(G)(G)V(C)
Modern and Contemporary Chinese	(C)(G)V(N)

Arcodia (2007)

Stage	Minimal Syllable	Maximal Syllable	Final consonants
Archaic Chinese	CVC	CCCMVCCC	At least ten different
Middle Chinese	CV	{C,S}V{C,S}	[m], [n], [ŋ], [p], [t], [k]
Modern Chinese	V	{C,S}VC	[n], [ŋ]

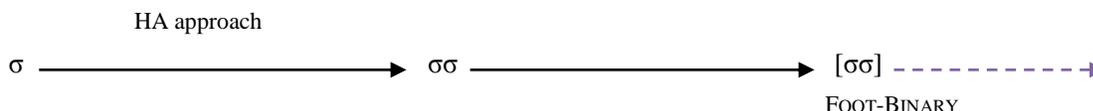
It is obvious that Chinese syllable has a complex structure during the ancient time. Feng 2000 argues disyllabification is caused by FOOT-BINARY. Feng 2000 discusses a foot obey binary branching condition. However, after the disappearance of consonant codas, the new syllable structure cannot form a foot because this structure violates the binary branching condition. As a result, disyllabification occurs. Feng 2000 fails to explain the reason why those syllables with entering tone, which have (C)VC structure, disyllabified; and no supporting evidence can be found to show that archaic Chinese syllables can form a single binary foot.

According to Feng 2000, the disappearance of entering tone should happen before the large scale of disyllabification. There is no evidence show that tone change occurred before disyllabification. In fact, contemporary Cantonese still has disyllabic words with entering tone syllables. Feng 2000 argues that the occurrence of contour tones balance duration, which can be summarized as below:



According to Feng 2000, archaic Chinese should not have any contour tones. However, no supporting evidence can be found. In fact, Zhang 2002 adds on that contour tones tend to occur on a phonemic long vowel, which has two moras, based on cross-linguistic statistical evidence. Thus, archaic Chinese has more possibility to have contour tones than modern Chinese. We hypothesize that FOOT-BINARY occurs after functional movement, that is, HA approach. In order to avoiding homophonous ambiguity, two-syllable combinations occur. A disyllabic morphological template has been modeled with the high type frequency of disyllabic items, then, FOOT-BINARY established. This process can be treated as Blocking, see (7).

(7): diachronic development of disyllabification (template)



On the other hand, Feng 2000 also fails to explain the shortening of archaic Chinese syllables. If Chinese prefer binary foot, Chinese would like to keep the original heavy syllables, but evidence has shown that Chinese syllables get shortened since ancient time. This tendency still exist today, Wang 1992 argues contemporary Chinese syllable begins to shorten to (C)(G)Vnasal structure, which has already happened in Beijing dialects.

Heavy syllables may change into light syllables because simple syllables are unmarked cross-linguistically. Although Middle Chinese has contour tones, but there is no evidence showing that Archaic Chinese do not have contour tones, in fact, according to Zhang 2002, Middle Chinese may have fewer tonal changes than Archaic Chinese. Thus, we can strongly believe that Chinese loses syllable types diachronically, which then integrated ambiguities arises.

Additionally, languages tend to use short structures to explain semantic meanings may follow the Gricean theory, which is also called Cooperative Principle. According to Gricean theory, four rules⁵ govern rational, cooperative conversational behavior in general. Among these four rules, Maxim of Quantity may contribute to the wide use of simple structures.

Maxim of Quantity: Be as informative as required.

Evidence can be found cross-linguistically. For example, in English, *be gonna*, which is in a tri-syllabic form, is shortened from a quadri-syllabic words, *be going to*. In oral conversations, speakers tend to say *be gonna* instead of *be going to*, because the rule, Maxim of Quality, contributes to the shorter one, *be gonna*, which is easier for speakers to say. Another rule of Griean theory may be applied to:

Maxim of Quality: Make your contribution true; so do not convey what you believe false or unjustified.

This rule contributes to conversation from a listener's aspect. As a listener, he would like to get enough information so that he can understand well. Thus, monosyllabic may not provide enough information because of integrated ambiguity.

Modern Chinese has about 1300 types of syllable, which can only explain 1300 different semantic meanings, while a disyllabic form can reflect 1300*1300 semantic meanings, which is more or less adequate for daily use. For those which cannot be explained in disyllabic forms, people tend to choose tri-syllabic forms or even polysyllabic forms. Word length depends on the both the Quality and the Quantity, there is no need for words can be well explained in disyllabic forms to have longer syllables.

⁵ The four rules are: Maxim of Ouality. Maxim of Ouantity. Maxim of Relation. Maxim of Manner

For instance, Zhongguo Wenzhi Gaige Weiyuanhui Yanjiu Tuiguang Chu [Chinese Language Reform Committee Research and Popularization Office] is shortened for Zi Gai Wei, not Zhong Zi Gai Wei, Gai Wei, or Zi Gai, etc. The reason is that only tri-syllabic form can explain this well, if we choose Gai Wei or Zi Gai, it may lead to ambiguity, such as ‘reform committee’ or ‘language reform’. On the other hand, we do not have to use Zhong Zi Gai Wei and everyone can know that it stands for ‘Chinese language reform committee’. Both Gai Wei or Zhong Zi Gai Wei obeys FOOT-BINARY, but the output of this construction is Zi Gai Wei. Such examples also add evidence to HA approach. Thus, we may insist that word change is motivated by HA theory and Feng 2000’s argument on FOOT-BINARY as the motivation of Chinese disyllabification may be wrong in this case. But we cannot deny FOOT-BINARY as an unmarked constraint for prosodic words cross-linguistically.

5 Concluding Remarks

This paper contributes to the debate of motivation of Chinese disyllabification and claims that disyllabic words are affected by several approaches, while HA is the most important one and can be treated as the motivation of Chinese disyllabification. But we do not deny metrical approach and FOOT-BINARY as important roles when dealing with disyllabic words in Chinese, especially prosodic words.

This paper argues against Duanmu 1999, 2007 that HA theory does play a clear role in the disyllabification of Chinese by providing supporting evidence for Lü 1963 prediction. Cantonese, the language has more types of syllables (1795), shows having more monosyllabic words (31.4%) than Mandarin (25.5%), the language with fewer types of syllables (1300) because the number of homophones can be reduced by an increase in the number of syllable types. According to our data, HA theory not only contributes to the disyllabification of Chinese, but also contributes to other language changes cross-linguistically. This can be proved by the evidence that languages with fewer syllable types have fewer monosyllabic words, such as Japanese, and languages with more syllable types have more monosyllabic words, such as American English, see (4). Duanmu 1999, 2007 fail to explain a relationship between syllable types and monosyllabic distributions.

This paper argues against Feng 2000 that FOOT-BINARY may not be the motivation of disyllabification of Chinese. Instead, it is a byproduct of HA approach. Mandarin has about 1300 syllable types, and disyllabification can create about 1300×1300 different new forms ideally, which is enough for daily use. In some languages, which have limited syllable types, like Japanese (about 136 syllable types), tri-syllabification may be applied to avoid homophonous ambiguities. In some other language, such as American English, which has more than 10,000 syllable types, there is no need to create too many disyllabic words.

This paper also argues against Feng 2000 that contour tone does not occur to

balance duration. In fact, according Zhang 2002, archaic Chinese has a higher possibility to have contour tones than middle Chinese. Thus, contour tone cannot be treated as a coordinator of duration. On the other hand, according to Baerman 2009, contour tone is a mechanism of HA. Baerman 2009 studies Chiquihuitlán Mazatec, a Mexico language. In this language, negation is marked partly by a distinct ending and partly by tonal alternations, that is, tonal alternations has been added to avoid integrated ambiguity. However, there is no significant evidence show that contour tone is another mechanism of HA approach in Chinese. This question needs further discussion.

The present discussion also raises some further questions. We discuss the relationship between syllable types and % monosyllabic words in section 2 and section 3, according to (4), syllabic distribution shows different curves in different languages, we would like to know to what extent HA can be explained as the motivation of polysyllabification cross-linguistically. Additionally, a lack of monosyllabic words in modern Chinese can be seen obviously, and Japanese may play an important role in this process. As we mentioned above, Japanese is a language with fewest monosyllabic forms because it has simplest syllable types and the Chinese loanwords from Japanese is a combination of both writing system and phonological system. Thus, to what extent Japanese influence the disyllabic words in Mandarin also needs further discussion.

Compare to Mandarin and Cantonese, which get similar ratio of the number of Mandarin syllable types divided by that of Cantonese syllable types to the ratio of the number of Mandarin monosyllabic lexical words divided by that of Cantonese monosyllabic lexical words. American English does not apply to this rule. American English has more than 10,000 syllable types, which are seven times of Mandarin, but the % monosyllabic words in American (48.34%) only twice of that of Mandarin (25.5%). However, cross-linguistically, a complex syllable is marked and languages prefer a simple syllable. We calculated the number of syllable types (only simple syllables) in English and put aside those with complex onset or complex coda. There are about 3000 different simple syllable types in total. Interestingly, we can see the proximity of the two ratios ($p > 0.05$); see (8).

(8) Syllable types and monosyllabic lexical words in Mandarin and Cantonese

$$a) \frac{M-\sigma \text{ types}}{E-\sigma \text{ types}} = \frac{1300}{3000} = 43.3\%$$

$$\frac{M-\% \text{ monosyllabic lexical Wds}}{E-\% \text{ monosyllabic lexical Wds}} = \frac{25.5\%}{49.3\%} = 51.7\%$$

$$\frac{M-\sigma \text{ types}}{E-\sigma \text{ types}} \approx \frac{M-\% \text{ monosyllabic lexical Wds}}{E-\% \text{ monosyllabic lexical Wds}}$$

$$b) \frac{C-\sigma \text{ types}}{E-\sigma \text{ types}} = \frac{1795}{3000} = 59.8\%$$

$$\frac{C\text{-\%monosyllabic lexical Wds}}{E\text{-\%monosyllabic lexical Wds}} = \frac{31.4\%}{49.3\%} = 63.7\%$$

$$\frac{C-\sigma \text{ types}}{E-\sigma \text{ types}} \approx \frac{C\text{-\%monosyllabic lexical Wds}}{E\text{-\%monosyllabic lexical Wds}}$$

This shows that marked constraints of complex onset and complex coda also contribute to word length. We would like to know to what extent HA contribute to those language with complex segmental features and how HA and such constraints work out together. Such questions will be left for future research.

REFERENCES

- Abby Kaplan, 2010, *How Much Homophony Is Normal?*. CUNY Conference on the Word in Phonology.
- Chen, Huiying & Rao, Hui. 2002. *Cantonese--Putonghua in the Chinese language*. Beijing: Beijing Language and Culture University.
- Ding, Bangxin, 1979. *Shanggu Hanyu de Yinjie Jiegou [The syllable structure of Archaic Chinese]*, Taiwan: Taipei.
- Duanmu, San. 1999. "Stress and the development of disyllabic words in Chinese". *Diachronica* 16: 1-35.
- Duanmu, San. 2007. *The Phonology of Standard Chinese*. UK: The Oxford University Press
- Feng Shengli, 2000, *The Origin of Disyllabicity in Chinese 漢語雙音化的歷史來源, Contemporary Research in Modern Chinese (Zhongguoyu yanjiu 《現代中國語研究》)* 2000, 1:123-138. Japan. Also included in the book *From Semantic Information to Typological Study*, ed. by Youwei Shi. Beijing Language and Cultural University Press. 2001: 22-47. Beijing.
- Giorgio Francesco Arcodia, 2007. *Chinese: A Language of Compound Words? In Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, ed. Fabio Montermini, Gilles Boyé, and Nabil Hathout, 79-90.
- George Kingsley Zipf, 1949, *Human Behavior and The Principle of Least Effort*. New York: Hafner Publishing Company.
- Grice, H. Paul. 1989, *Studies in the way of words*, Cambridge, Mass: Harvard University

- Press.
- Helen S. Eaton. 1961. *An English-French-German-Spanish Word Frequency Dictionary*. New York.
- He, Kekang & Li, dakui. 1987. *Xiandai Hanyu Sanqian Changyongci Biao [Three thousands commonly used words in Mandarin]*. Beijing: Beijing Normal University.
- John Xiang-Ling Dai, 1990. *Historical Morphologization of syntactic words. Evidence from Chinese derived verbs. Diachronica VII: 1.9-46 (1990)*
- Kao, Diana L. 1971. *Structure of the syllable in Cantonese*. the Hague.
- Katsuo Tamaoka, Shogo Makioka, 2009. *Japanese Mental Syllabary and Effects of Mora, Syllable, Bi-mora and Word Frequencies on Japanese Speech Production, Language and Speech: 52(1), pp. 79-112, 2009*
- Lin, Tao & Wang, Lijia. 1992. *Yuyinxue Jiaocheng [The course of phonetics]*. Beijing: Peking University
- Lü, Shuxiang. 1963. “Xiandai Hanyu dan shuang yinjie wenti chu tan” [A preliminary study of the problem of monosyllabism and disyllabism in modern Chinese]. *Zhongguo Yuwen 1963.1: 11-23*.
- Martin Haspelmath, 2005. *Against markedness (and what to replace it with)*, appear in *Journal of Linguistics 41 (2005) or 42 (2006)*, Cambridge University Press.
- Matthew Baerman, 2011. *Defectiveness and homophony avoidance. J. Linguistics 47 (2011), 1-29*. Cambridge University Press.
- Rao, Bingcai, Ouyang, Jueya & Zhou, wuji. 1991. *Guangzhouhua Fangyan Cidian [Dictionary of Guangzhou dialect]*, Hong Kong.
- Wang, Jenny. 1992, *The Geometry of Segmental Features in Beijing Mandarin*. Ph.D. Dissertation, University of Delaware.
- Wayne A. Davis, 1998. *Implicature: Intention, convention, and principle in the failure of Gricean theory*. UK: Cambridge University Press.
- Yu, Naiyong, 1985. *Shanggu Yinyi Yanjiu [A study of Archaic Chinese phonology]*, CUHK
- Zhan, Bohui. 2002. *An outline of Yue dialects in Guangdong*. Guangzhou: Jinan University.
- Zhang, Jie (2002). *The effects of duration and sonority on contour tone distribution--A typological survey and formal analysis*. Routledge, New York.
- Zheng, Ding'ou & Pan, Xiaoluo. 2006. *Xinbian Jinri Yueyu [New Cantonese today]*. Beijing: Beijing Language and Culture University.
- The Corpus of Contemporary American English (COCA), created by Mark Davies of Brigham Young University in 2008
- Zhongguo Wenzhi Gaige Weiyuanhui Yanjiu Tuiguang Chu [Chinese Language Reform Committee Research and Popularization Office]. 2008. *Xiandai Hanyu Changyongci (Cao'an) [A Draft of commonly used words in Standard Chinese]*. Preliminary edition. Beijing: Wenzhi Gaige Chubanshe.

On the Origins of Sinitic

Scott DeLancey
University of Oregon

A striking fact about the Sinitic branch of Sino-Tibetan is that, while its lexicon, phonological structure and some reconstructible morphology clearly link it genetically to Tibeto-Burman, its basic morphosyntactic profile is the isolating SVO type characteristic of mainland Southeast Asia rather than the agglutinating SOV structure characteristic of Tibeto-Burman. This results from the history of Sinitic as a more typical Tibeto-Burman SOV language which was came to be used as a lingua franca by the Tai-Kadai, Hmong-Mien, and Austroasiatic “Bai Yue” populations of the south even before their incorporation into the Chinese Empire. The resulting creoloid syntax (Ansaldo and Matthews 2001) has remained robust in China and mainland Southeast Asia.

Evidence for early contact with Tai-Kadai languages is shared vocabulary (representing loans in both directions) in Proto-Tai-Kadai and Sinitic, matching in tone class. The oldest stratum is of Old Chinese date or older (Li 1945; Manomaivibool 1975), which implies that it dates from first contact between Sinitic and TK. Assuming Ostipirat’s (2000) reconstruction of PTK to 4,000 years ago, the date correlates with Xia. Presumably Sinitic was still Tibeto-Burman-like, and TK had yet to diverge dramatically from Austronesian (Sagart 2005). The dramatic creolization which produced the grammatical structure of Sinitic, Tai, Kadai, and other mainland Southeast Asian languages originated in this original contact.

1 The problem

The essential problem in the formation of Chinese is that it has strong lexical, phonological, and grammatical connections both with the Tibeto-Burman languages to the west and with the Southeast Asian languages to the south, especially Tai-Kadai, but also Hmong-Mien and Mon-Khmer. While basic vocabulary and some reconstructible morphology clearly link Sinitic genetically to Tibeto-Burman, its basic morphosyntactic profile is the isolating SVO type characteristic of mainland Southeast Asia rather than the agglutinating SOV structure characteristic of Tibeto-Burman. This occasionally even leads to doubts about the genetic relationship of Sinitic and Tibeto-Burman (Beckwith

2002, 2006), but few linguists doubt that the history of Sinitic is of a Tibeto-Burman language, with the SOV syntax characteristic of that family, which was adopted by a population speaking Kadai, Austronesian and/or Austroasiatic, and quite possibly other, languages¹.

The evidence which requires explanation falls into four broad categories: lexical correspondences among Chinese and one or more other languages or families, morphological correspondences between Chinese and Tibeto-Burman, and the striking similarities in both syntactic and phonological structure between Chinese and the mainland Southeast Asian families. The difficulty is that there is significant evidence linking Chinese with several different language groups, including Tai-Kadai, Hmong-Mien, Austroasiatic, and Austronesian, but it cannot be genetically related to all or even several of them. Most of what Chinese shares with most of these languages must thus have resulted from language contact. The fundamental problem of Sinitic linguistics is to unravel the various linguistic threads which make up Old Chinese and its predecessors and understand how they came to be woven together into the language which we know.

The idea that many of the features of Old Chinese came about as a result of contact between the language of Shang (and/or possibly Zhou) conquerors and an indigenous population speaking a language or languages related to those currently spoken in mainland Southeast Asia is old (Terrien de la Couperie 1887). My purpose in this paper is to propose a more explicit model of this contact, and to argue that it offers a path toward solutions to several long-standing problems. Little of the content of this paper is novel. In sections 2 and 3 I will briefly summarize the kinds of evidence which a hypothesis of Sinitic origins needs to account for. In Section 4 I will suggest a hypothesis which I hope will account for at least some of it, invoking a model of language contact and interaction which has not really been applied previously to the problem of Chinese.

2 Lexical and phonological correlations

Chinese shows striking points of correspondence with Tai-Kadai, Hmong-Mien, and Vietnamese in both lexicon and phonological structure. In this section I will very briefly review some of the reasons why these congruences do not argue for a genetic relationship among any of these languages.

¹ We do not need to deal here with the status of Tibeto-Burman as a genetic unit, i.e. the question of whether Sinitic represents one of two primary branches of the family, is a branch coordinate with higher-level branches like Bodic, or, as van Driem (1997, 2008) suggests, is a actually subordinate unit. However, the conventional view of the family as Sinitic + Tibeto-Burman is based on traits which Sinitic shares with the Bai Yue rather than the Tibeto-Burman languages. Therefore the explanation offered here for those correlations can be seen as undermining the basis of that view.

2.1 Issues in Lexical Comparison

Since the earliest days of serious linguistic study of Chinese, scholars have noted the huge amount of vocabulary shared between Chinese and neighboring languages. A great deal of this was obviously borrowed from Chinese, which throughout historic times has been the major cultural force in East Asia. But there is also a very substantial body of vocabulary shared with Kadai, Hmong-Mien, and/or Vietnamese, which is much older than that, and it is not easy to determine whether such shared forms are common inheritance or borrowing, and in the latter (more likely) case, borrowing in what direction.

Thus, aside from the Tibeto-Burman languages (Matisoff 2003), Chinese has been linked with Tai-Kadai (Wulff 1934, Nishida 1975, FK Li 1945, 1976, Manomaivibool 1975, 1976a, b, *inter alia*), Austroasiatic (Norman and Mei 1976), Hmong-Mien (Downer 1963, 1971, Wang 1986, Haudricourt and Strecker 1991), and Austronesian (Sagart 1994, 1995, 1999).² On the one hand, all of these proposals are supported by serious lexical comparisons, and some sort of historical connection with Tai-Kadai, Hmong-Mien and Vietic within Austroasiatic is established both by shared lexicon and by the astonishing correspondence in phonological typology (see below). But these three groups are not evidently related, so Sinitic can hardly be genetically related to all of them, much less to all of them and Tibeto-Burman as well, except at some very hypothetical, very high, unrecoverable level. Thus it has long been clear that some of the evidence which has been adduced to argue for genetic relations among these languages in fact reflects sustained intense contact among unrelated languages (Terrien de la Couperie 1887, Matisoff 1973, LaPolla 2001, *inter alia*).

The interpretation of the lexical evidence has sometimes been confused by unrealistic notions of when and how borrowing can take place, in particular simplistic ideas that lexical borrowing only occurs from a more dominant into a smaller and less “advanced” population, or under some kind of necessity. For example, Manomaivibool (1975:364), discussion shared Tai-Chinese vocabulary, says “It seems implausible that Tai *had to* borrow that many items of such common vocabulary from Chinese” (emphasis added). But without a more thorough understanding of the social, economic, and political context of Shang and Zhou era China, it is simply impossible to distinguish borrowings from cognates purely on how easily one can imagine a motivation for borrowing a word with a particular meaning.³ I will suggest a model of language contact which makes room for exactly the sort of unsystematic lexical mixture which we find in Chinese.

² I ignore more far-fetched suggestions of connections to Uralic, Indo-European, North Caucasian, and Na-Dene.

³ Consider the following French borrowings into English: *family, dinner, supper, soup, easy, difficult, quiet, silent, noisy, lake, river, mountain, valley, forest, marsh, flower, village, city, language, story, color, attack, defend, protect, accept, argue, agree, beautiful, stupid, count, real, false, very, front*. It would be hard to argue that any of these, or thousands of similar items, “had” to be borrowed for cultural reasons.

It has been suggested (FK Li 1976, Manomaivibool 1975, 1976a, b, Nishida 1975, 1976) that if a Sino-Tai form can be reconstructed for Proto-Tai-Kadai, this is evidence for genetic relationship between Sinitic and T-K, presumably on the grounds that PTK is too old to have been contemporary with any stage of Chinese, so that there would be no time at which borrowing could have taken place. But there is no logic to this argument – whether we imagine the common vocabulary to reflect a common proto-language or to represent borrowings, in either case PT-K or something ancestral to it, and Old Chinese or something ancestral to it, must have been contemporaneous. Noting this fact does not constitute an argument for one hypothesis or the other. What is important is that Li and other scholars consider the oldest layer of shared Tai-Chinese vocabulary (which certainly represents loans in both directions, not only from Chinese to Tai) to be of at least Old Chinese date, so that this common lexicon probably dates from the earliest contact.

2.2 The Southeast Asian phonological profile

The most impressive correspondence between Sinitic and the Southeast Asian Tai-Kadai, Hmong-Mien, and Vietic languages is in their phonological structure. All share the stereotypical monosyllabic morpheme structure and elaborate tone systems. The most striking, and puzzling, fact about this congruence is the perfect correspondence of the tone systems (Wulff 1934, Haudricourt 1954a, b, FK Li 1945, 1976, Matisoff 1973, Ostapirat 2000, Ratliff 2010). Sinitic, Tai-Kadai, Hmong-Mien, and Vietnamese all have a four-tone system, with a three-way distinction on “smooth”, i.e. open or sonorant-final syllables, and all “checked”, i.e. obstruent-final, syllables manifesting a distinct fourth tone. Each of the other three shares with Sinitic (and to some extent with each other) a substantial body of shared vocabulary which shows regular correspondence in tone class. In all of the languages tones originated out of final laryngeal features, so that the original correspondence is in the type of rime: obstruent coda, coda *-h (sometimes < *-s), final *-ʔ, and “smooth” syllables with none of these (Haudricourt 1954a, b, 1961/1972, Mei 1970, 1980). The shared vocabulary which shows these correspondences must have been borrowed⁴ at a stage when both the donor and recipient languages still retained these final laryngeal distinctions, and had not yet developed phonemic tone; if we imagine that these items were borrowed with phonemic tone, it becomes impossible to explain the regularity of the correspondences. (For a very clear exposition of this argument see Ratliff 2010:187-93). The languages must have still been in close contact when they underwent a shared tonogenetic episode in which these laryngeal distinctions were reinterpreted as tonal, as they were still centuries later when they all shared in the “Great Tone Split” conditioned by mergers of initial consonant series.

⁴ Even on the hypothesis that some or all of these languages might be genetically related at some level, there is no question that a great deal of the most obvious shared lexicon represents borrowing in one direction or the other.

The monosyllabic pattern is not really characteristic of Austroasiatic, or even of Kra-dai, and the Sinitic developments do have parallels in the phonological development of other Sino-Tibetan groups. So Sagart is probably right in attributing the original locus of monosyllabic structure to Chinese:

From a typological point of view, Old Chinese was more similar to modern East Asian languages like Gyarong, Khmer or Atayal than to its daughter language Middle Chinese: its morphemes were nontonal and not strictly monosyllabic; its morphology was essentially derivational, and largely prefixing; but it also made use of infixes and suffixes. At some point between Old Chinese and Middle Chinese, and for unknown reasons, a cascade of changes caused the language to move away from this model. Its affixing morphology began to freeze; its loosely attached prefixes were lost, while other affixes clustered with root segments and were reinterpreted as root material. A new morphemic canon tending toward strict monosyllabism, with a great variety of initial and final consonant clusters, emerged. Further shifts saw the reduction of initial clusters, this resulting in a more complex inventory of initial consonants, and in new vowel contrasts. Final clusters were also reduced and the inventory of final consonants restricted to resonants and stops, this leading to the emergence of tones. Thus the classical ‘Indochinese’ typology common in its major features to Middle Chinese, Vietnamese, Miao-Yao, Tai, Burmese etc., was born. (Sagart 1999:13)

3 Grammatical evidence for Sino-Tibetan

The strongest evidence for the genetic affiliation of Sinitic with the Tibeto-Burman languages is grammatical, specifically correspondences in personal pronouns and in some reconstructable verb morphology.

3.1 The pronouns

An important argument for the Sino-Tibetan affiliation of Chinese has always been the correspondence of the 1st and 2nd person pronominal roots. We will look here at the 1st person forms 吾 (*nga* in Baxter’s reconstruction) and 我 (*ngaX* < **ngaj?*), which neatly match the reconstructed PTB root **ŋa* (Matisoff 2003).⁵ Sagart has argued on the basis of its late appearance in Shang and Zhou inscriptions that the 1st person **ŋa* root is a secondary development in Chinese (1999:142-4), and a late borrowing from there into the rest of Tibeto-Burman (145-6). Instead of the well-established **ŋa*, Sagart proposes that the PTB 1st person pronoun was the stop-initial *ka* which occurs as the primary 1st person

⁵ In fact the OC forms very specifically resemble the Tibetan 1st person pronominals *nga* and *nged*.

root in three geographically marginal branches of the family, northern Qiangic, Kuki-Chin, and “a few languages of eastern Nepal and neighboring areas”, i.e. Kiranti in Nepal and Western Himalayan in northwest India. He suggests that this distribution reflects a spread of **ŋa*, ultimately from Chinese, through the contiguous central TB area, leaving only the few branches on the edges of the TB area untouched. These three (actually four, plus a few strays in Nepal and Arunachal Pradesh) then retain what Sagart takes to be the original Sino-Tibetan 1st person root **ka*.

This proposal cannot be correct, since the nasal root is found in the 1st person agreement suffix which is reconstructable for Proto-Tibeto-Burman (Bauman 1975, Sun 1983, van Driem 1993, DeLancey 1989, 2010, *inter alia*), and thus long predates its first appearance in the Chinese inscriptions.⁶ Since a late borrowing from a Tibeto-Burman source into Chinese does not seem likely here, we have to recognize this root as dating back to their common ancestor. But it is likely that the **ka* root may also be ancient; it now appears that it was a possessive or oblique form contrasting with the nominative **ŋa* (Jacques 2007, DeLancey 2011b). What we see in the languages where this form has replaced the original nominative **ŋa*, is the replacement of the original finite construction with an innovative finite form based on a nominalization, which thus takes a genitive rather than a nominative “subject” (DeLancey 2011c).

3.2 Ancient morphology

While we find no inflectional morphology recorded in any form of Chinese, the fossils of pre-Chinese prefixes and suffixes can be found in the phonological alternations of semantically and graphically related words. Two morphological constructions which are securely reconstructable for both Tibeto-Burman and Sinitic are a causative prefix **s-* (Conrady 1896, Mei 1980, 2008, Dai 2001) and a nominalizing **-s* suffix (Downer 1959, Forrest 1960, Mei 1980, Mazo 2002).

The **s-* causative is retained in Written Tibetan and a handful of other languages, though in many it is no longer productive:

Tibetan	<i>log</i> ‘return (intransitive)’, <i>slog</i> ‘turn (transitive)’
Boro	<i>gab</i> ‘cry’, <i>səgab</i> ‘make s.o. cry’
Tarong	<i>ip</i> ⁵⁵ ‘sleep’, <i>səip</i> ‘cause to sleep’

In most modern TB languages, we find the prefix reflected in devoicing of the initial consonant:

Tibetan	<i>nub</i> ‘sink’, <i>snuv</i> ‘destroy, abolish’
Zaiwa	<i>nop</i> ‘sink in mud’, <i>ɲop</i> ‘make s.t. get bogged in mud’

⁶ The **ka* 1st person root which Sagart suggests is the original Tibeto-Burman form is in fact old, but as an oblique alternant to the nasal root (Jacques 2007, DeLancey 2011).

Boro	<i>gi</i> ‘afraid’, <i>si-gi</i> ‘frighten’
Newar	<i>gya-</i> ‘afraid’, <i>khya-</i> ‘frighten’

And we find the same in Old Chinese (Mei 2008):

见 xiàn ‘be visible’	< *gians
见 jiàn ‘see’	< *kians < *s-k < *s-g
别 píe ‘leave, separate (intr.)’	< *bjät < *brjat
别 piè ‘discriminate, distinguish’	< *pjät < *prjat < *s-p < *s-b
墨 mò ‘ink’	< *mæk
黑 hēi ‘black’	< *s-mæk

Dai (2001) demonstrates that this construction is ancient in Tibeto-Burman; based on this and the abundant evidence for it in Chinese, Mei (2008) suggests that it is a defining feature of Sino-Tibetan languages; note that it is preserved, at least in fossil form, in some branches which have lost almost all other inherited morphology, e.g. Bodo-Garo.

The nominalizing *-s suffix is reflected in Old Chinese tonal alternations (Downer 1959, Forrest 1960); which correspond neatly to the suffix which is preserved in Written Tibetan (Mei 1980):

Chinese		Tibetan
量 liáng ‘to measure’	< *lianj	‘grang ‘to count’
量 liàng ‘a measure’	< *liangs	grangs ‘a number’
织 zhī ‘to weave’	< *tjæk	‘thag ‘to weave’
织 zhì ‘woven goods’	< *tjæks	thags ‘web, woven stuff’

Morphological comparisons like this are the sine qua non of comparative linguistics, and without some strong argument discrediting comparisons like these, we can take this evidence as conclusively establishing the genetic relationship of Sinitic with the rest of Sino-Tibetan.

3.3 The South East Asian syntactic profile

Mainland Southeast Asia is well-known for its striking areal linguistic typology, characterized both by the elaborate and congruent tone systems discussed above (this is not shared by most Mon-Khmer languages) and by radically isolating SVO morphosyntax. Indeed the examples put forward to illustrate isolating typology are

always languages from this area; aside from modern European-based creole languages, few if any other languages in the world are as resolutely free of any sort of inflectional morphology. In this respect Chinese clearly sorts with the Southeast Asian rather than the Tibeto-Burman languages, which are characteristically agglutinative, SOV, and often morphological very complex.

There is no serious question that the Southeast Asian syntactic profile in Chinese is a secondary development:

From the fact that we can clearly see changes in the word order of these three languages [Sinitic, Karen, and Bai] over time, and cannot see such changes in the Tibeto-Burman languages other than Bai and Karen, we assume that it was Bai, Karen and Chinese that changed rather than all the other Tibeto-Burman languages. (LaPolla 2003:28)

A few scholars see this development as internal to Sinitic:

The new linguistic standard of the Han dynasty ... typologically characterized by its incipient isolating morphology, and its emergent tonal and monosyllabic phonology, gradually spread to all parts of the empire, north and south, and this same typology further spread to all non-Chinese languages spoken in territories under Chinese rule after the Han: all of Miao-Yao, Viet-Muong (but not the rest of Mon-Khmer), all of Kam-Tai, some south-eastern Tibeto-Burman languages including Lolo-Burmese (but not Tibetan, Qiang, Gyarong, etc.). (Sagart 1999:8)

But most scholars, from Terrien de la Couperie on, see the shift in Sinitic as due to influence from neighboring languages to the south; Egerod (1976:59) points out that since SVO order is inherited in Thai, “Chinese was largely a recipient rather than a donor in the early times ... it is Chinese which borrows a new word order”⁷ (see also Benedict 1976). Indeed, all of the Southeast Asian groups have SVO syntax as far back as we can trace. And there are ample traces of earlier SOV patterning in Old Chinese (Cheng 1983) and even in modern Mandarin, as detailed in Wu Fuxiang’s contribution to this conference (2011).

4. The Formation of Chinese

There is no question that the formation of Chinese involved contact with neighboring languages, definitely including Austroasiatic, Tai-Kadai, and Hmong-Mien, and very possibly others which have completely disappeared. There has been a certain tension on

⁷ Egerod links this claim to a putative shift in transitivity type, but this part of his argument is unnecessary; it is sufficient that a Bai Yue-type substratum was present to contribute the basic grammatical pattern.

the question of what sorts of contact might be involved. Traditionally there seem to be two basic possibilities: contact between adjacent languages, i.e. imagining Proto-Sinitic, Proto-Hmong-Mien, etc. as spoken in adjacent states, or super-substratum influence, i.e. an “elite dominance” model in which Proto-Sinitic formed in a state consisting of immigrant Tibeto-Burman conquerors interacting with indigenous Proto-Tai-Kadai or Proto-Hmong-Mien subjects. Both of these have important deficiencies; in this section I will develop an alternative model which is better suited to explain the kinds of data we have been considering.

The territory where Sinitic languages are spoken was an area of substantial linguistic diversity from prehistoric times (Terrien de la Couperie 1887, Pulleyblank 1983, 1995, Ballard 1984, LaPolla 2001, inter alia). We are particularly concerned with the “Bai Yue” 百越 languages, which seem to have been of mixed provenance, including both Austroasiatic and pre-Kadai and pre-Hmong-Mien languages (JZ Li 1994, Meacham 1996, LaPolla 2001). The Yue people and kingdom to the south are a long-term presence in Chinese history, but the first explicit reference to the Hundred Yue is in the Qin era *Annals of Lü Buwei*:

For the most part, there are no rulers to the south of the Yang and Han rivers, in the confederation of the Hundred Yue tribes [lit. 百越之际 ‘among the Hundred Yue’], in the territories of Bikaizhu, Fufeng, and Yumi, and in the states of Fulou, Yangyu, and Huandou. (Lü et. al. 2000:112 / Book 20/1.3)

This term is important because it makes clear that the reference of Yue is multiethnic (Luo 1990:268):

Leaving aside the Austronesian question, it seems highly likely that the peoples called Yue at various times by the Han Chinese spoke Austroasiatic languages, early forms of Hmong/Mien, Tai-Kadai languages and perhaps languages in other families now extinct. (Meacham 1996:98)

The question is, how does Chinese come to share large bodies of vocabulary, and characteristic phonological and morphosyntactic typological profiles, with these languages?

The similarity of the isolating Chinese-Southeast Asian morphosyntactic profile to creole languages has been noted for some time⁸. But history gives us no reason to suppose that Chinese at any stage was ever a true creole, in the traditional sense of a language which develops from a grammarless pidgin. Ansaldo and Matthews (2001)

⁸ I don’t know who first made this observation; I heard it in the 1970’s from David Strecker.

suggest the term “creoloid” to refer to such a language, which shows the constellation of typological features traditionally associated with creole languages but does not have the history of a creole:

Perhaps a more constructive way to see the “prototypical creole” traits is that languages which have been subject to intensive contact involving several typologically distant varieties will tend to show some combinations (or subset) of these features. (Ansaldo & Matthews 2001:317)

Such languages arise in conditions of intense contact, when for whatever reason some significant portion of the language community are second-language rather than native speakers (McWhorter 2007). This kind of development has occurred, and continues to occur, repeatedly in Tibeto-Burman (DeLancey 2010a, to appear a, b), and it is clear that Sinitic has the same kind of history.

The broad account which I suggest here is the familiar picture of a contact situation between western invaders speaking a TB tongue and locals speaking languages affiliated with one or more of the attested mainland Southeast Asian stocks. But it is not enough to simply say “contact” and pretend that we have explained anything. In this view of Sinitic we have a very specific outcome, with Sino-Tibetan lexical and grammatical core, heavy Bai Yue, especially Kadai, lexical influence, creoloid syntax based more on Bai Yue than on Sino-Tibetan patterns, and innovative phonological structure. This did not come about through people overhearing each other’s languages on market day, or learning a few phrases for doing business; we have to imagine a situation of widespread bi- or multilingualism. This would be the case in a scenario in which Chinese or pre-Chinese speakers conquered a Bai Yue population, as happened as the kingdoms of Chu and then Yue were incorporated into Qin China. But this does not automatically explain the extent of the influence which we find on the whole language. Ballard’s (1984) “Mother Soup” metaphor captures the problem but doesn’t solve it. More importantly, the most important contact evidence predates the assimilation of the southern kingdoms into imperial China.

Instead, I propose that the features which so dramatically distinguish Sinitic from other Tibeto-Burman branches reflect the use of Proto-Sinitic as a lingua franca, used widely by non-Chinese (by whatever definition) outside of the actual administrative control of the Chinese state. As we have noted, the term Bai Yue refers to the multiethnic and multilingual situation in the south. One can imagine the utility of a vehicular lingua franca even without reference to the Chinese state and its influence; by the time the Chinese state is present on the historical stage, some version of its language would be a likely candidate. Thus, with the increasing power and prestige of Zhou, perhaps even Shang, China, a pidginized version of its Tibeto-Burman language became a lingua franca throughout the region. Cheng (1983) speaks of “two sublanguages

coexisting in early archaic Chinese”, an earlier SOV stratum and an innovative SVO syntax. This would, essentially, be “pure” Sino-Tibetan Chinese with SOV syntax, and innovative “foreigner” Chinese, spoken with the SVO pattern of the Bai Yue languages. Ultimately the widespread lingua franca version of Proto-Sinitic replaced the original everywhere.

REFERENCES

- CAAAL = Computational Analysis of African and Asian Languages
 ICSTLL = International Conference on Sino-Tibetan Languages and Linguistics
 LTBA = Linguistics of the Tibeto-Burman Area
- Ansaldo, Umberto, and Stephen Matthews. 2001. Typical creoles and simple languages: The case of Sinitic. *Linguistic Typology* 5.2–3:311–326.
- Ballard, W.L. 1984. The Mother Soup: A South China recipe for tonometamorphogenesis. *CAAAL* 22:43-70.
- Beckwith, Christopher. 2002. The Sino-Tibetan problem. *Medieval Tibeto-Burman Languages*, ed. by C. Beckwith, 113-57. Leiden: Brill.
- _____. 2006. Introduction: Toward a Tibeto-Burman theory. *Medieval Tibeto-Burman Languages II*, ed. by C. Beckwith, 1-38. Leiden: Brill.
- Benedict, Paul. 1976. Sino-Tibetan: Another look. *Journal of the American Oriental Society* 96.2:167-197.
- Cheng, Tsai Fa. 1983. SOV sentence patterns in Early Archaic Chinese. 16th ICSTLL, Seattle.
- Conrady, August. 1896. Eine indochinesische causative-denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten. Leipzig: Harrasowitz.
- Dai, Qingxia. 2001. Zang-Mian yu zu yuyan shidong fanchou de lishi bianhua (The historical development of the causative category in Tibeto-Burmese languages). *Journal of Chinese Linguistics* 29.1: 1-10.
- DeLancey, Scott. 2010a. Language replacement and the spread of Tibeto-Burman. *Journal of the Southeast Asian Linguistics Society* 3.1: 40-55.
- _____. 2010b. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics Journal* 9.1: 1-38.
- _____. 2011a. Nocte and Jinghpaw: Morphological correspondences. *North East Indian Linguistics* 3, ed. by Gwendolyn Hyslop, S. Morey, and M. Post, 61-75. New Delhi: Foundation/Cambridge University Press India.
- _____. 2011b. Agreement prefixes in Tibeto-Burman. *Himalayan Linguistics Journal* 10.1:1-35.
- _____. 2011c. Finite structures from clausal nominalization in Tibeto-Burman languages. *Nominalization in Asian Languages: Diachronic and Typological Perspectives*,

- ed. by Foong Ha Yap, K. Grunow-Hårsta & J. Wrona, 343-62 [Typological Studies in Language 96]. Philadelphia: Benjamins.
- _____. to appear a. On the origins of Bodo-Garo. *North East Indian Linguistics* 4, ed. by Gwendolyn Hyslop, S. Morey, and M. Post. New Delhi: Foundation/Cambridge University Press India.
- _____. to appear b. Creolization in the divergence of Tibeto-Burman. *Papers from the 16th Himalayan Languages Symposium*, ed. by Nathan Hill.
- Downer, Gordon. 1959. Derivation by Tone-Change in Classical Chinese. *Bulletin of the School of Oriental and African Studies* 22.1/3:258-290.
- _____. 1963. Chinese, Thai, and Miao-Yao. *Linguistic Comparison in South East Asia and the Pacific*, ed. by H.L. Shorto, 133-139.
- _____. 1971. The further relationships of the Miao-Yao languages. 4th ICSTLL, Bloomington.
- Driem, George van. 1993. The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* 56.2: 292-334.
- _____. 1997. Sino-Bodic. *BSOAS* 60.3: 455-488.
- _____. 2008. To which language family does Chinese belong, or what's in a name. *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*, ed. by A. Sanchez-Mazas et.al., 219-53.
- Egerod, Søren. 1976. Benedict's Austro-Tai hypothesis and the traditional view on Sino-Thai relationship. *CAAAL* 6:51-59.
- Forrest, R. A. D. 1960. Les occlusives finales en chinois archaïque. *Bulletin de la Société Linguistique de Paris* 55:228-239.
- Hashimoto, Mantaro. 1976c. Genetic Relationship, Diffusion and Typological Similarities of East & Southeast Asian Languages: Papers for the 1st Japan-US Joint Seminar on East & Southeast Asian Linguistics. Tokyo: Japan Society for the Promotion of Science.
- Haudricourt, André. 1954a. De l'origine des tons en vietnamien. *Journal Asiatique* 242:69-82.
- _____. 1954b. Comment reconstruire le chinois archaïque. *Word* 10.2-3:351-64.
- _____. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient. *Bulletin de la Société de Linguistique de Paris* 56: 163-80.
- _____. 1972. Two-way and three-way splitting of tonal systems in some Far Eastern languages. *Tai phonetics and phonology*, ed. by Jimmy Harris and Richard Noss, 58-86. Bangkok: Central Institute of English Language. (English translation of Haudricourt 1961).
- _____, and David Strecker. 1991. Hmong-Mien (Miao-Yao) loans in Chinese. *T'oung Pao* n.s. 77.4-5: 335-41.
- Henderson, Eugenie. 1965. The topography of certain phonetic and morphological characteristics of Southeast Asian languages. *Lingua* 15. 400-434.

- Jacques, Guillaume. 2007. A shared suppletive pattern in the pronominal systems of Chang Naga and Southern Qiang. *Cahiers de linguistique Asie Orientale* 36.1:61-78.
- LaPolla, Randy. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. *Areal Diffusion and Genetic Inheritance*, ed. by Alexandra Aikhenvald and R.M.W. Dixon. Oxford: Oxford University Press.
- _____. 2003. Overview of Sino-Tibetan morphosyntax. Thurgood and LaPolla (eds.), 22-42.
- Li, Fang-kuei. 1945. Some old Chinese loan words in the Tai languages. *Harvard Journal of Asiatic Studies* 8: 333-42.
- _____. 1971. Shang gu yin yanjiu. *Tsing Hua Journal of Chinese Studies* 9.1-2: 1-61.
- _____. 1976. Sino-Tai. Hashimoto (ed.), 230-39.
- _____. 1983. Archaic Chinese. Keightley (ed.), 393-408.
- Li, Jingzhong. 1994. *Yuyan yanbian lun*. Guangzhou: Guangzhou Chubanshe.
- Lü, Buwei, John Knoblock and Jeffrey Riegel. 2000. *The Annals of Lü Buwei: A Complete Translation and Study*. Stanford: Stanford University Press.
- Luo, Xianglin. 1990. *Zhongguo Minzu Shi*. Taibei : Zhonghua Wenhua Chuban Shi Ye She.
- Manomaivibool, Prapin. 1975. *A Study of Sino-Thai Lexical Correspondences*. Ph.D. dissertation, University of Washington.
- _____. 1976a. Thai and Chinese – are they genetically related? *CAAAL* 6:11-31.
- _____. 1976b. Layers of Chinese loanwords in Thai. *Tai linguistics in honor of Fang-kuei Li*, ed. by Thomas Gething, J. Harris and P. Kullavanijaya, 179-184. Bangkok: Chulalongkorn University Press.
- Matisoff, James. 1973. Tonogenesis in Southeast Asia. *Consonant Types and Tone*, ed. by Larry Hyman, 73-95. (Southern California Occasional Papers in Linguistics 1). Los Angeles: Linguistics Program, University of California Los Angeles.
- _____. 2003. *Handbook of Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. [University of California Publications in Linguistics 135]. Berkeley; Los Angeles: University of California Press.
- Mazo, Olga. 2002. Reflexes of the suffix *-s in Classical Tibetan and Old Chinese. *Tibet, Past and Present*, ed. by Henk Blezer, 455-472. Leiden: Brill.
- McWhorter, John. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press.
- Meacham, William. 1996. Defining the Hundred Yue. *Bulletin of the Indo-Pacific Prehistory Association* 15: 93–100.
- Mei, Tsu-lin. 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. *Harvard Journal of Asiatic Studies* 30: 86-110.
- _____. 1980. Sisheng bieyi zhong de shijian cengci. *Zhongguo Yuwen* 1980.6: 427-43.

- _____. 2008. Shanggu Hanyu dongci zhuoqing bieyi de lai yuan (The original differences between verbs with voiced and voiceless initials in Ancient Chinese). *Minzu Yuwen* 2008.3:3-19.
- Nishida, Tatsuo. 1975. Common Tai and Archaic Chinese. *Studia Phonologica* 9:1-12.
- _____. 1976. Some problems of the morpheme stock in Sino-Tibetan: A preliminary observation. Hashimoto, (ed.), 30-38.
- Norman, Jerry, and Mei Tsulin. 1976. The Austroasiatics in ancient south China: Some lexical evidence. *Monumenta Serica* 32:274-301.
- Ostapirat, Weera. 2000. Proto-Kra. *LTBA* 23.1:1-251.
- Pulleyblank, E. G. 1983. The Chinese and their neighbors in prehistoric and early historic times. Keightley (ed.), 411-466.
- _____. 1995. Historical and prehistorical relationships of Chinese. *The Ancestry of the Chinese Language*, ed. by William S-Y. Wang, 145-94. [Journal of Chinese Linguistics Monograph Series no. 8]. Berkeley: Project on Linguistic Analysis, University of California.
- Ratliff, Martha. 2010. *Hmong-Mien Language History*. (Studies in Language Change 8). Canberra: Pacific Linguistics.
- Sagart, Laurent. 1994a. Proto-Austronesian and Old Chinese evidence for Sino-Austronesian. *Oceanic Linguistics* 33.2:271-308.
- _____. 1994b. Questions of method in Chinese-Tibeto-Burman comparison. 27th ICSTLL, Paris.
- _____. 1995. Some remarks on the ancestry of Chinese. Wang (ed.) 195-223.
- _____. 1999. *The Roots of Old Chinese*. Philadelphia: Benjamins.
- Terrien de la Couperie. 1887. *The Languages of China Before the Chinese*. London: David Nutt.
- Wang, Fushi. 1986. A preliminary investigation of the genetic affiliation of the Miao-Yao languages. *International Symposium on the Minority Nationalities of China*, Santa Barbara.
- Wu, Anqi. 2008. Shiqian huanan diqu de yuyan anchu. *Minzu Yuwen* 2008.3:21-35.
- Wu, Fuxiang. 2011. Hanyu zhong jige you tese de jyushi (Language-specific syntactic constructions in Chinese). presented at the 23rd North American Conference on Chinese Linguistics, Eugene, OR, June 18, 2011.
- Wulff, Kurt. 1934. *Chinesisch und Tai: Sprachvergleichende Untersuchungen*. (Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser 20.3). København: Levin & Munksgaard.

Analogy and Reanalysis in the Postposing of Durative and Iterative Adverbials in the History of Chinese

Jeeyoung Peck

University of the Pacific

This study attempts to examine both speakers' motivations for and grammatical mechanisms involved in the postposing of temporal adverbials denoting duration (e.g. *santian* 'for three days') and iteration (e.g. *sanci* 'three times') over the course of the history of the Chinese language. It will point out that durative and iterative adverbials delimit event type, similar to quantized NPs in direct object position (e.g. *san ping jiu* 'three bottles of wine') which have consistently been postverbal. It analyzes how the event delimiting function shared by these temporal adverbials and quantized NPs motivated speakers to align all these structures in the postverbal position. It also discusses that as a mechanism of word order change, situation-delimiting adverbials such as *sanci/santian* were reanalyzed as pseudo-objects (e.g. *jie zhebenshu jie le sancisantian* 'borrowed this book three times/for three days') in Early Mandarin, and distinguished syntactically from other purely temporal adverbials denoting frequency (e.g. *changchang*, 'often') or temporal frame (e.g. *zai santian zhinei*, 'in three days').

1. Introduction

Chinese word order has been SVO continuously from the time of Old Chinese through Modern Chinese (Sun & Givón, 1985).¹ While the basic word order has been stable, the sentential position for adverbial phrases, i.e. those non-subcategorized adjunct phrases, has gone through noticeable changes. For example, in Old Chinese, iterative adverbials typically appeared in the preverbal position, and durative adverbials were allowed in this position as well.

- (1) a. 今一日飲酒，而三日寢之 (晏子春秋)
b. 三讀九思，方服淵致 (全劉宋文)

In contrast, in Modern Chinese, these adverbial phrases generally follow verbs and cannot appear in the preverbal position (Klippel, 1991).

¹ Periodization in this work follows Sun (2006), with each historical period subdivided into Early and Late periods: Early Old Chinese (771 BCE to 207 BCE), Late Old Chinese (206 BCE to 220 CE), Early Middle Chinese (220 to 589 CE), Late Middle Chinese (589 to 960 CE), and Early Mandarin (960 to 1900 CE).

- (2) a. 全仗着喝了一天酒啊。(海边的雪, PKU²)
 b. (你的文章)我至少看过三遍。(盡管我們不相識, PKU)

There have been a number of studies that discuss usage and frequency of iterative and durative adverbials in different stages of history such as Xu (1997), Yin (2002), Wang (1997), Liu (1992), Tang (1990), Zhuang (1980), Liu (1965), and Oota (1958) among others.³ While these previous studies mainly focus on specific historical stages, Zhang (2010) is the first attempt to analyze exhaustive quantitative data through the course of history, explain the usage of different patterns of these elements and to provide statistics.

However, none of these previous studies provide qualitative analysis regarding why and how the sentential distribution of durative and iterative adverbials has changed and how they ended up appearing in the postverbal position. This work is the first attempt to analyze speakers' motivations and grammatical mechanisms in the postponing of durative and iterative adverbials. For quantitative data, I will rely on the findings from Zhang (2010) whenever required.

2. Historical facts

2.1 Gradual shift of iterative adverbials to the postverbal position

In Old Chinese, iterative adverbial phrases were expressed only by a bare number and occurred preverbally most of the time (Oota, 1958; Wang, 1958; Fan, 1982; Zhang, 1987).

² PKU refers to the online corpus developed by Center for Chinese Linguistics at PKU. For details, visit http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp?dir=xiandai

³ Usage and semantic differences in preverbal and postverbal iteratives (i.e. Num V and V Num) in Old Chinese are introduced in Xu (1997). Frequency of each pattern, i.e. Num-ECL V and V Num-ECL in the Wei and Jin period is dealt with in Liu (1965), Zhuang (1980) and Liu (1992). Syntactic function of both preverbal and postverbal iterative adverbials and their frequency in Tang dynasty documents, *Bianwen*, is analyzed in Wang (1997).

A diachronic analysis of the rise and fall of each variation of iterative adverbials between the pre-Qin and the Wei and Jin period is provided in Tang (1990). He argues that iterative adverbials consisting of bare number first shifted to the postverbal position (Num V > V Num), and then the postverbal bare-number iterative adverbials were replaced by iterative adverbials consisting of number and ECL (V Num > V Num-ECL).

Oota (1958) provides general observations on the usage of iterative and durative adverbials in the history of Chinese, and argues that preverbal iteratives with bare number (Num V) were affected by analogy with postverbal durative adverbials and became V Num-ECL in the Wei and Jin period.

(3) 三責我 (晏子春秋)

While most of the iterative adverbials appeared only in the preverbal position in Old Chinese, there were some exceptions. Tang (1990) observed postverbal iterative adverbials from periods between the Western Zhou and Qin dynasties⁴.

(4) 鞭女五百 (贖匱, Tang, 1990)

However, iterative adverbials were found far more frequently in the preverbal position than the postverbal position. The followings are more examples of iterative adverbials in the preverbal position between Old Chinese and Early Middle Chinese.

- (5) a. 齊與魯三戰而魯三勝 (戰國策)
 b. 昔齊魯三戰，魯人再克而亡不旋踵 (三國志)

As early as the Han dynasty (observed from wooden slips of the Han Dynasty from the Juyan) or the Wei and Jin period (reflected in transmitted documents), a few more iterative adverbials began to appear in the postverbal position. These postverbal iterative adverbials were no longer bare number expressions but nominal phrases that consisted of number expressions and event classifiers (hereafter, ECL) (Oota, 1958).

- (6) a. 農呼妻相出於庭，叩頭三下 (搜神記)
 b. 凡種小麥地，以五月內耕一遍 (齊民要術)

Over the course of history, while there was always variation, the ratio of preverbal iterative adverbials continuously decreased and more and more postverbal iterative adverbials surfaced. Between the Tang dynasty and the Yuan and Ming dynasties, iterative adverbials began to occur primarily in the postverbal position, as illustrated in the following examples.

- (7) a. 師以手敲拄杖三下 (鎮州臨濟慧照禪師語錄)
 b. 問言誦咒幾千遍 (李白)
 c. 誦《金剛經》三遍 (太平廣記)
 d. 牧敲柱三下 (五燈會元)

⁴ Tang studied iterative adverbials from eleven pre-Qin classic books, and observed that 6% of iterative adverbials were in the postverbal position. The fact that iterative adverbials were allowed in either the preverbal or postverbal position shows that there was not yet a linguistic constraint preferring these adverbials in certain positions.

- e. 隻聽得番將馬上敲了三下 (三寶太監西洋記)

2.2. Durative adverbials gradually constrained to the postverbal position

Oota (1958) observed that in Old Chinese both preverbal and postverbal positions have been available for noun phrases denoting temporal duration, though such phrases have occurred primarily postverbally.

- (8) a. 三日哭 (春秋)
 b. 居三日, 魏王乃聽起賈 (呂氏春秋)

Between late Han Dynasty and late Tang Dynasty, the postverbal position continued to be the primary position for durative adverbials. Even though some preverbal occurrences of durative adverbials were still found, as shown in (9b), in the later historical stages the postverbal position was the most common for durative adverbials, as shown in (9c-f).

- (9) a. 臥三日 (戰國策)
 b. 而孔子三日樂 (淮南子)
 c. 武迎置獄三日 (前漢紀)
 d. 醉臥三日 (支謙譯經)
 e. 宋州僧尼, 行道三日 (唐代墓志匯編續集)
 f. 當服小功者只服七日 (朱子語類)

2.3. Frame and frequency adverbials stayed in the preverbal position

In Modern Chinese, adverbials expressing time-frame such as *in an hour*, *within a day*, etc. and adverbials denoting frequency such as *everyday*, *often*, etc. are found in the preverbal or the pre-sentential position.⁵

- (10) a. 一只猴子一天可摘 1 千多个椰子。 (中国儿童百科全书, PKU corpus)
 b. 每天练歌唱歌。 (中國北漂藝人生存實錄, PKU corpus)

Throughout history, frame adverbials ('in *x* time') have appeared only in the preverbal position. Between Old Chinese and Early Middle Chinese, frame adverbials

⁵ As in (10a), frame adverbials in Modern Chinese is often expressed by the preverbal temporal NP, but sometimes a rather classical style *zai...zhinei* ('within') is also used together, also denoting a temporal frame.

(i) 一位厂长在一天之内竟签下了 1 5 7 次自己的名字。 (1994 年报刊精选, PKU.)

were often expressed by bare NP temporals such as *yiri* ('one day') or by PPs such as *yi yiri* ('in one day')

- (11) a. 一日浸百畦 (庄子)
 b. 常以一日行千裡 (前漢紀)
 c. 驚牛一日行百裡 (世說新語)

From the time of the Tang dynasty, frame adverbials have been also expressed by the extended pattern *yiri zhinei* ('one day in').

- (12) a. 一日之內，無留訟牘 (唐文拾遺)
 b. 一日之內遇大雨扳土 (天工開物)

Frequency adverbials have also only been found in the preverbal position throughout history.

- (13) a. 常仗趙而畔楚 (戰國策)
 b. 故貧民常衣牛馬之衣 (漢書食貨志)
 c. 公常攜兄子邁及外生周翼二小兒往食 (世說新語)
 d. 終日常行乞食而活生命 (敦煌變文集新書)
 e. 時常到村店中吃酒 (碾玉觀音)
 f. 學生常在他家看病 (金瓶梅)

So far, we have shown that among temporal adverbials, frame and frequency adverbials remained in their original preverbal position. On the other hand, iterative and durative adverbials, which could appear in the preverbal position without any constraint, gradually came to be constrained to the postverbal position.

3. Iterative and durative adverbials are event delimiters

Some adverbials are distinguished from purely modifying adverbials due to their function to temporally quantify the event or state specified by a verb phrase. Examples of adverbials which are event delimiters⁶ include durative and iterative adverbials

⁶ According to Tenny (1994), studies of the properties of event delimitation have a long tradition, even going back as early as Aristotle, and event delimiters have been an important concept in verb aspectuality, relied upon by many linguistic and philosophical works such as Kenny (1963), Ryle (1949), Vendler (1967), Dowty (1979), Bach (1981, 1983, 1986), Mourelatos (1981), Moens and Steedman (1988) and Jackendoff (1990).

(Wechsler and Lee 1996). On the other hand, purely modifying adverbials such as temporal point adverbials ('at two o'clock'), frequency adverbials ('often') and frame adverbials ('in two hours') simply modify an event without affecting the temporal course of events.

In our work, we follow Pustejovsky (1991) and others in using the concept of aspectuality (also called as eventuality) and classify event types (also called situation types in other work) as State, Process and Transition. Transition is the term which collapsed the traditional notions of achievement and accomplishment.

3.1. Event delimitation by quantized NPs

In order to introduce the notion of event delimitation, I will start by showing how a nominal of definite quantity (e.g. *a glass of wine*) delimits the temporal course of an event by "cutting out an entity of a certain size from a continuum" (Wechsler and Lee, 1996).

(14) 他喝了一杯红酒。

The nominal *yibei hongjiu* ('a glass of wine') in the example above denotes a definite quantity. A definite quantity of wine delimits the temporal course of the drinking event. As part of the wine gets consumed, part of the wine-drinking event progresses, and when the last sip of wine is consumed, the drinking event ends. The quantification process is transferred from the domain of object ('wine') to the domain of events ('drinking').

Whether or not the addition of quantized NPs delimited the primitive event type can be tested by using a diagnostic with the frame adverbials (*in*-adverbials) and durative adverbials (*for*-adverbials). The idea is that a delimited event, a Transition, will be grammatically modified by a frame adverbial, because a Transition involves the natural culmination of an event. Meanwhile, a non-delimited event, a Process or a State, will co-occur with a durative adverbial grammatically.

- (15) a. *他喝了两个小时的一杯红酒。
 b. 他两个小时喝了一杯红酒。
 c. 他喝了两个小时的红酒。
 d. *他两个小时喝了红酒。

In (15a-b), the spatially-quantized quality of 'a glass of wine' can be transferred to the temporally-quantized event of drinking. Thus, the frame adverbial naturally co-occurs with this VP-denoting Transition as in (15b), rather than with the durative adverbial as in (15a). (15c-d) shows Process, the non-quantized event 'drinking wine' instead, and this Process co-occurs harmoniously with the durative adverbial (15c), rather than with the frame adverbial as in (15d). The Process of drinking can occur for an indefinite period of

time, because the non-quantized quality of wine can be transferred to the temporally non-delimited event⁷.

3.2. Event delimitation by durative adverbials

Usually States or Processes expressed in the progressive aspect are also inferred in the perfective aspect (based on de Swart, 1998), as shown in (16).

(16) a. 张三在游泳。 → b. 张三游泳了。

When listeners hear a statement like (16a), they can assume (16b).

However, the same inference is not found in a Transition in which a quantized object delimits the given event.

(17) a. 张三在画一张画。

≠> b. 张三画了一张画。

However, when a State or Process co-occurs with a durative adverbial, it does not allow the same inference from the progressive aspect to the perfective anymore, similar to the case of the Transition in the example above.

(18) a. 张三在游泳。他今天要游三个小时，现在还在游。

≠> b. 张三游了三个小时的泳。

When listeners hear the statement (18a), they cannot assume the statement (18b) is true. In short, a predicate co-occurring with a durative adverbial in (18) behaves more similarly to a predicate with a quantized object, i.e. a Transition in (17), than to a simple Process or State in (16). De Swart (1998) proposes that a durative adverbial combines with a State or Process and turns the given event into a quantized event, i.e. Transition (also see Krifka, 1989; Kamp and Reyle, 1993; Naumann, 1995; Moens, 1987; Vet, 1980)

⁷ Such delimitation of events by quantized object shows correlation between the domain of events and the domain of objects. For further details, see Verkuyl (1972), Hinrichs (1985), Krifka (1989), Dowty (1991), Tenny (1994) among others.

⁸ The phenomenon in which a primitive aspectual class (i.e. event type) of a verb is changed to a derived aspectual class of a VP by adding inherent or added arguments is called type-shifting (Bach, 1986). For example, a direct argument, e.g. an object NP, can type-shift to a primitive aspectual class: a verb like 'eat' could describe delimited (Transition) or non-delimited (Process)

- (19) Durative adverbials (*for*-adverbials)

FOR *x time* & State U Process → Transition (based on de Swart 1998:357)

On the other hand, a frame adverbial ('in two years') maps a set of quantified events onto another set of quantified events without affecting the event type of the modified predicate.

- (20) Frame adverbials (*in*-adverbials)

IN *x time* & Transition → Transition (based on de Swart, 1998)

Frame adverbials do not participate in event composition, whereas durative adverbials do. Such a difference can be understood easily by the following contrast.

- (21) 那本书张三写了两年。

In (21), due to the durative adverbial, i.e. the postverbal temporal NP *liangnian* 'two years', the process of writing is expressed as lasting for a definite (minimum) period of time (based on de Swart, 1998).

Meanwhile, in (22) the amount of time expressed by the frame adverbial is also two years. However, the frame adverbial phrase, i.e. the preverbal temporal NP *liangniani* 'in two years', specifies the temporal scope within which the event of writing was completed.

- (22) 张三两年写了一本书。

The length of time that the completion of writing actually took might be shorter than two years, let's say 21 months. In other words, a frame adverbial might not express the actual temporal course of the modified event. A frame adverbial merely modifies the given event, without altering the event type of the modified event (de Swart, 1998).

3.3. Event delimitation by iterative adverbials

Iterative adverbials quantify events and present the events as countable, which are comparable to countable objects. Even though frequency adverbials also semantically quantify events, they quantify events in an atelic way, comparable to mass objects. This difference between iterative and frequency adverbials is reflected in their tendency to select tense or aspect (de Swart, 1991).

events depending on whether a direct object is a quantized NP ('an apple') or an unquantized NP ('apples') (Verkuyl, 1993; Tenny, 1994; Krifka, 1995).

In Modern Chinese, iterative adverbials denoting simple counts of events can occur in the perfective aspect. On the other hand, they do not sound natural in the imperfective aspect, as shown in (23).

(23) *张三在弹两次钢琴。(Imperfective)

On the other hand, frequency adverbials are not sensitive to specific aspect in Chinese. They sound perfectly natural in the imperfective aspect⁹.

(24) 张三常常在弹钢琴。(Imperfective)

In French also, according to de Swart (1991), iterative adverbials do not select for the imperfective or present contexts because events quantified by iterative adverbials describe an absolute quantity. Iterative phrases directly participate in forming the aspectuality of a given event by taking a sub-event, which is delimited ('playing the piano once') and turning it into a series of sub-events, which are also delimited ('playing the piano twice') (Klippel, 1991). Delimited events are incompatible with the imperfective aspect, which is 'mass-like.'

On the other hand, events quantified by frequency adverbials denote homogeneous events that have an atelic property. Such events combine harmoniously with either a 'mass-like' aspect or tense (e.g. the imperfective and the present) or a 'count-like' aspect or tense (e.g. the perfective and the past)¹⁰ (de Swart, 1991).

4. Motivations and mechanisms

In this section, we will show that iterative and durative adverbials were aligned in the postverbal position in Chinese on analogy with quantized NPs in object position. These different syntactic structures had in common their semantic functions as event delimiters, which resulted in their sharing a common sentential position, i.e. following the verb.

⁹ For the same phenomenon found in French and further analysis on this issue, see de Swart (1991).

¹⁰ From a semantic point of view, both iterative and frequency adverbials quantify events with which they co-occur, and these quantified events refer to plural events. Events quantified by iterative adverbials can be compared to countable plural NPs. Events quantified by frequency adverbials are similar to mass NPs. In the domain of determiners, count/mass determiners produce delimited/undelimited properties of NPs respectively. In the respect that both durative and iterative adverbials produce plurality of events, they are distinguished from modifying adverbs (e.g. run *fast*), which add more precise characterization to the events that they combine with, and do not affect the singularity or plurality of modified events. For a semantic analysis on quantifying and modifying adverbs, see de Swart (1991).

4.1. Reanalysis in VP with postverbal durative adverbials

Previous studies such as Zhang (2010) and Oota (1958) argued that iterative adverbials were analogized with durative adverbials, on the grounds that these two adverbials both quantify events, and as a result, shifted to the postverbal position. However, more should be said about this process in terms of motivations and mechanisms. For example, it is still not clear why the postverbal position eventually became where these two situation-delimiting types of adverbial appeared.

Generally speaking, temporal adverbials were typically found in the preverbal position: in the Han Dynasty, the majority (90%) of iterative adverbials could be found in the preverbal position, whereas the majority (91%) of durative adverbials were in the postverbal position (based on Zhang, 2010); Other purely temporal adverbials such as frequency and frame adverbials have continuously appeared only in the preverbal position throughout history. Such a distribution enables us to imagine a reverse scenario in which durative adverbials could have shifted to the preverbal position in analogy with iterative and other temporal adverbials.

So why would the postverbal position rather than the preverbal position have become the locus of durative and iterative adverbials? Considering the fact that the sentential position for objects has continuously been the postverbal position (Sun and Givón, 1985), having quantized NPs in object position is therefore a better candidate for the model of analogy that attracted both durative and adverbial phrases rather than durative adverbials which experienced slight positional changes themselves.

In what follows, we will explain that the position of durative adverbials came to be constrained to the postverbal position through the reanalysis, which modeled the position of quantized NP objects, and then the position of iterative adverbials changed also, modeling after durative adverbials.

When a hearer gives an analysis of structure and meaning to a form that is different from the intention of speaker, reanalysis can take place (Hopper and Traugott 2003: 50). Let us postulate a situation where a hearer was told a term *hamburger*. Even though a speaker intended to mean “item (of food) from Hamburg” and the original structure of this term was “[Hamburg]+[er]”, the hearer might analyze it as “[ham]+[burger]” and understand its meaning as a bread that contained ham. At this moment, reanalysis can take place. Reanalysis is a “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation” (Langacker 1977:58).

While this example of reanalysis took place at morphological level (word formation), our case of durative adverbials took place at a syntactic level (sentence formation). Let us imagine that someone heard a sentence that included a durative adverbial, whose original structure was [V+Adjunct] and the intended meaning was doing activity for the amount of time expressed by the adverbial phrase.

(25) 伏读一周 (六朝, 三国志)

For example, as in (25), the length of time expressed in the durative adverbial *yizhou* ('one week') functions to delimit the temporal course of the reading event. The reading event gets terminated when one week has passed. Such a semantic function is similar to that of quantized objects. If someone reads a book, the time required to finish reading depends on the volume of the book expressed in a quantized NP. Objects always appeared in the postverbal position by default in Chinese as in (26).

(26) 如使读一卷书 (东汉, 太平经)

Reanalysis might take place when the hearer applies the analysis of [V+NP complement] to example (25), whose original structure is [V+NP adjunct], on the ground that both duration of time and quantized objects provide a specific endpoint to an event.

Because it was a hearer who applied a different syntactic analysis to the VP with durative adverbials, speakers and hearers can not notice any change after reanalysis. The following summarizes reanalysis in VPs with durative adverbials.

(27) [reanalysis of delimiting adjuncts as complements, covert change]

Early Old Chinese		Late Old Chinese
the model: 读 [一卷书] _{argument}		
Read one CL book		
=		
读 [一周] _{adjunct}	>	读 [一周] _{complement}
read one week		read one week

4.2. Analogy of iterative adverbials with durative adverbials

When speakers apply the same new rule to other material, the new grammar of constraining delimiting adverbial phrases to the postverbal position will become noticeable. Such rule expansion or rule generalization is a primary characteristic of analogy.

In Meillet's view, analogy is a "process whereby irregularities in grammar, particularly at the morphological level, were regularized. The mechanism was seen as one of proportion or equation...It is overt." (Hopper and Traugott 2003:64). Let us imagine that other speakers of English, in the context of '-burger', substituted *ham* with other words such as *cheese* or *beef* or even *veggie*. Analogy must have a pre-existing example as a model (Hopper and Traugott 2003:64). In this example, the model for the new form of *cheese-burger* was *ham-burger*. Now, once unapparent reanalysis from *hamburg-er* to *ham-burger* became apparent through *cheese-burger*. Such attraction to other words makes reanalysis which was once invisible visible (Hopper and Traugott 2003).

While this example shows an analogy within a word level, our case with iterative adverbials serves as an example of analogy within a sentence level. Recall that iterative

PECK: ANALOGY AND REANALYSIS

adverbials used to appear in the preverbal position.

(28) 三读九思，方服渊致 (东汉末年-南朝梁时，弘明集) (=1b)

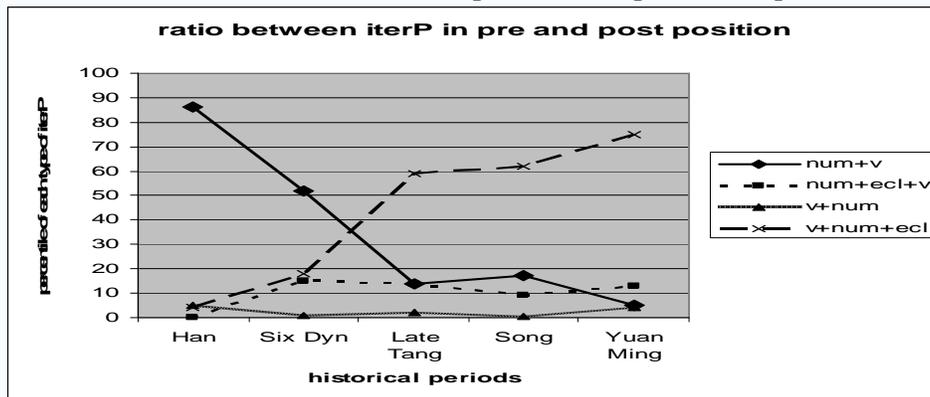
(29) 主人三遍讀文 (北魏末年，齊民要術)

Iterative adverbials also provide a definite endpoint to events that they describe, similar to durative adverbials. Based on this shared semantic function to delimit events, let us imagine a speaker expanded the new rule of constraining delimiting adverbials to the postverbal position to iterative adverbials. As a result, we get a postverbal iterative adverbial as follows.

(30) 必当先读百遍 (六朝，三国志)

It is plausible to say that speakers might have expanded this new sentential position to an increasing number of iterative adverbials because such rule regularization can save speakers from memorizing a number of various grammars. As a result, the preverbal position for iterative adverbials lost its competition with the postverbal position.

The chart below is drawn from Zhang (2010). It illustrates the overall picture of gradual shift of iterative adverbials from the preverbal to postverbal position.¹¹



¹¹ Token numbers for Chart 1 (Zhang, 2010).

	Han	Six Dynasties	Late Tang	Song	Yuan and Ming
Num+V	154	124	22	183	10
Num+ECL+V	0	35	22	94	29
V+Num	9	2	3	5	8
V+Num+ECL	8	45	89	655	162

Chart 1. Sentential distribution of iterative adverbials throughout history (Zhang, 2010)

The preverbal bare-number iterative adverbials occupied more than 85% of total iterative expressions in the Han dynasty. As time passed by, a new expression for iteratives, that is, the string of number (Num) and event classifier (ECL), emerged in the Han dynasty and increased in both preverbal and postverbal positions in the Six dynasties period. From the Six dynasties period, this new expression began to increase with a high rate in the postverbal position, but stopped increasing in the preverbal position. From the late Tang Dynasty, the postverbal Num-ECL became the major expression of iteration until Yuan and Ming dynasties (Zhang, 2010) as shown in the following.

(31) 对大师读一遍 (唐, 六祖坛经)

4.3. Triggering effect for analogy

We argue that what triggered the analogical process between iterative adverbials and durative adverbials in the postverbal position was the emergence of event classifiers in iterative adverbials. Due to the emergence and spreading of event classifiers, iterative adverbials such as *yi hui* ('one time') came to have a similar syntactic form to duration of time such as *yi zhou* ('one week'), as well as quantized nominal expressions such as *yi juan shu* ('one CL book')¹². This new pattern with ECL enabled speakers and listeners to analogize the syntax of event delimiting adjunct phrases (adverbial phrases) with that of event delimiting arguments (quantized NPs as objects) through their shared form, i.e. classifiers.

(32) [analogy and reanalysis in the position of iterative adverbials]

Late Old Chinese		Early Middle Chinese
the model: 读 [一 周] _{complement} read one week		
= [一] _{adjunct} 读 one read	>	读 [一 遍] _{complement} drink one ECL
[一 遍] _{adjunct} 读 one ECL drink		

In addition, as reflected in Chart 1, while the frequency of both the Num+ECL+V pattern and the V+Num+ECL pattern increased between the Han Dynasty and the Six

¹² The occurrence of classifiers in NPs (e.g. 一尺布, 一斗粟 from 史记) emerged in the Qin dynasty and became widespread in the Han dynasty (Wang, 1958). We abbreviate nominal classifiers as CL here.

Dynasties period, the frequency of the Num+V pattern decreased concurrently. Then, starting in the Six Dynasties period, the preverbal Num+ECL+V pattern stopped increasing in frequency, losing in competition to the postverbal V+Num+ECL pattern. Finally V+Num+ECL pattern became the majority in the Late Tang period. This correlation between the increase in frequency of V+Num+ECL and the decrease in frequency of Num+V and Num+ECL+V strongly suggests that the emergence and the spreading of ECL was a trigger for the appearance and increasing occurrence of the postverbal V+Num+ECL pattern.

To summarize, the previously possible constituent order of iterative and durative adverbials was reanalyzed into a new constituent order, in analogy with the V-quantized NP order.

(33) a. change in constituent order between durative adverbials and verbs

[Num Temp_{unit}]_{+del} V, V [Num Temp_{unit}]_{+del} > V [Num Temp_{unit}]_{+del}

b. change in constituent order between iterative adverbials and verbs

[Num]_{+del} V > [Num]_{+del} V, [Num ECL]_{+del} V, V[Num ECL]_{+del} > V [Num ECL]_{+del}

What enabled duratives to stay in the postverbal position and iteratives to shift to the postverbal position was speakers' motivation to equate the sentential position of delimiting adjuncts with that of delimiting arguments.

(34) Old Chinese

Middle Chinese

the model: [V+Argument_{+del}]_{VP}

=AdjunctP_{±del} [V AdjunctP_{±del}]_{VP} > AdjunctP_{-del} [V AdjunctP_{+del}]_{VP}

Due to this analogy and reanalysis process in history, delimiting adjuncts, durative and iterative adverbials must appear as complements in the postverbal position in Modern Chinese, patterning with direct objects (i.e. arguments) in terms of sentential position. Gradually, situation-delimiting adverbials (i.e. internal adjuncts) came to be distinguished from non-delimiting adverbials (i.e. pure adjuncts) in terms of sentential position. Non-delimiting temporal adverbials such as *meitian* ('everyday', frequency adverbial) or *zai san tian zhi nei* ('in three days', frame adverbial) cannot appear in the postverbal position.

5. Syntactic correlates between delimiting temporal adjuncts and delimiting objects

Once the postponing of durative and iterative adverbials from Late Old Chinese through Middle Chinese took place through reanalysis and analogy, these adverbials were again reanalyzed as pseudo-objects. In Old to Middle Chinese, durative and iterative adverbials could follow objects in the postverbal position as illustrated below.

(35) 讀書百遍 (三國志, Fang 2006)

Since the time of Early Mandarin, a verb-copying construction has emerged and these situation-delimiting adverbials have come to be required to immediately follow the verb, at the cost of duplicating the verb (Fang, 2006). See the following examples.

- (36) a. 請人請到四五次 (盧太學詩酒傲公侯, Early Mandarin)
 b. 我借这本书借了三次 (Modern Chinese)

When there is no object in the postverbal position, durative and iterative adverbials must follow the main verb immediately in Modern Chinese.

- (37) a. 我看了两个小时。
 b. *我两个小时看了。
 c. 我看了两次。
 d. *我两次看了。

In this way, durative and iterative adverbials, which are complements, pattern with objects, which are subcategorized arguments in Modern Chinese. Arguments of verbs do not take overt case markers in Chinese. Instead, word order contains essential information for understanding a verb's argument structure. Noun phrases that act as objects follow verbs.

(38) 我看书了。

The so-called NP complement is a middle-status element postulated by Klipple (1991). Klipple (1991) distinguishes some adjuncts which are syntactically internal from pure adjuncts. These syntactically internal adjuncts are also semantically internal. In other words, they directly affect the aspectual structure of events. Klipple (1991) classifies these adjuncts as complements and proposes that durative and iterative adverbials in Modern Chinese are one of examples. Complements and arguments have in common that they are syntactically and semantically internal.

We accept Klipple (1991)'s proposal regarding Modern Chinese and argue that over the course of history durative adverbials continuously appeared in the postverbal position as a result of the reanalysis process, rather than getting analogized to the preverbal position where most of the pure temporal adverbials appeared. Our point is that in Old Chinese internal temporal adjuncts (delimiting adverbials) and pure temporal adjuncts (non-delimiting adverbials) were not differentiated in terms of syntax, but through reanalysis and analogy over the course of history, internal adjuncts came to appear in the sentential position where arguments usually appear.

6. Conclusion

This paper examined historical changes that affected the sentential position of durative and iterative adverbials and analyzed speakers' motivations and grammatical mechanisms reflected in the postposing of these adverbials in the history of Chinese.

The position of iterative adverbials was typically preverbal but gradually changed to be postverbal. Durative adverbials were allowed to appear in the preverbal position, but came to be constrained to the postverbal position. In contrast, the position of general temporal adverbials such as frequency and frame adverbials has not changed over the course of history.

Durative and iterative adverbials show a semantic correlate with quantized NPs in that these elements are event delimiters. These adverbials directly participate in event composition and present events as quantized ones, similar to countable or quantized NPs, which make events delimited. In contrast, frequency adverbials quantize situations in a non-delimited way, similar to mass NPs, and frame adverbials do not change the event composition of predicates but merely modify them.

We also provided a detailed analysis for why and how delimiting temporal adverbials came to be constrained to the postverbal position: durative adverbials were reanalyzed as complements, modeling after objects (i.e. arguments) whose default position have always been postverbal since Old Chinese; the iterative adverbials were reanalyzed as complements on analogy with durative adverbials. We showed that this analogy was triggered by the emergence and the spreading of event classifiers in iterative adverbials: while all iteratives, duratives and quantized NPs shared a similar semantic function to delimit events, once iterative adverbials had event classifiers, these structures came to share a similar form as well.

We also explained that delimiting temporal adverbials came to be constrained in the postverbal position, patterning with delimiting NPs first, and they were reanalyzed to pseudo-objects that had to follow V immediately in Early Mandarin.

The fact that event delimiting adjuncts exhibit a syntactic correlation with event delimiting objects in some other languages such as Korean and Finnish also supports that the sentential positions of durative and iterative adverbials were changed, patterning with quantized NPs in object position in the history of Chinese.

References

- Bach, E. 1981. On time, tense, and aspect: an essay in English metaphysics. In *Radical pragmatics*. Ed. by Cole, 62-81. New York: Academic Press.
- _____. 1983. On the relationship between word grammar and phrase grammar. *Natural Language & Linguistic Theory* 1: 65-90.
- _____. 1986. The algebra of events. *Linguistics and Philosophy* 9 (1):5-16.
- de Swart, H. 1991. Adverbs of quantification: a generalized quantifier approach. Doctoral dissertation, University of Groningen.

- _____ 1998. Aspect shift and coercion. *Natural Language & Linguistic Theory* 16 (2): 347-385.
- Dowty, D. R. 1979. *Word meaning and Montague grammar*. Dordrecht: D. Reidel Publishing Company.
- Fan, Y. 1982. *Hanyu yufashi gaiyao* [Essentials on the historical syntax of Chinese]. Zhongzhou Shuhuaaju.
- Fang, J. 2006. The verb copy construction and the post-verbal constraint in Chinese. Doctoral dissertation, Stanford University.
- Heinämäki, O. 1984. Aspect in Finnish. In *Aspect bound - a voyage into the realm of Germanic, Slavonic, and Finno-Ugrian aspectology*. Ed. by de Groot and Tommola, 153-177. Foris: Dordrecht.
- Hopper, P & E. Traugott. 2003. *Grammaticalization* (Second Edition). Cambridge: Cambridge University Press.
- Jackendoff, R. 1990. On Larson's treatment of the double object construction. *Linguistic Inquiry* 21:427-455.
- Kamp, H. and U. Reyle. 1993. *From discourse to logic*. Kluwer: Dordrecht
- Kenny, A. 1963. *Action, emotion, and will*. London: Routledge & Kegan Paul Limited.
- Klippel, E. 1991. The aspectual nature of the thematic relations: locative and temporal phases in English and Chinese. Doctoral dissertation, MIT.
- Krifka, M. 1989. Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and contextual expressions*. Ed. by Bartsch, van Benthem, and van Emde Boas, 75-115. Foris: Dordrecht.
- Langacker, R. 1977. Syntactic reanalysis. In *Mechanisms of Syntactic Change*, Ed. by C. Li, 57-139. Austin, Texas: University of Texas Press.
- Levin, B. and M. Rappaport Hovav. 2005. *Argument Realization*. Cambridge, UK: Cambridge University Press.
- Liu, Shiru. 1965. *Weijin Nanbeichao Liangci Yanjiu* [Study on classifiers in the Wei and Jin Dynasties]. Beijing: zhonghua shuju.
- Liu, Shizhen. 1992. *Weijin Nanbeichao Lishi Yufa* [Historical grammar of the Wei and Jin Dynasties]. Nanjing: Nanjing University Press.
- Mourelatos, A. 1981. Events, processes and states. In *Syntax and semantics*. Ed. by Tedeschi and Zaenen, 15:191-211. New York: Academic Press.
- Naumann, R. 1995. Aspectual Composition and Dynamic Logic. Doctoral dissertation, Dusseldorf.
- Oota, T. 1958. *Chuugokugo rekishi bunpo* [A historical grammar of Chinese]. Kokyo: Konan shoin.
- Pustejovsky, J. 1991. The Syntax of Event Structure. *Cognition* 41:47-81.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson's University Library.
- Sun, Chaofen. 2006.
- Sun, C. and T. Givón. 1985. On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implication. *Language* 61:329-351.

PECK: ANALOGY AND REANALYSIS

- Tang, Y. 1990. Guhanyu dongliang biaoshifa tanyuan [Study on Old Chinese grammar of event iteration]. *Guhanyu yanjiu* 1: 71-73.
- Tenny, C. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.
- Vendler, Z. 1967. Verbs and times. In *Linguistics in philosophy*. Ed. by Vendler, 97-121. Ithaca and London: Cornell University Press.
- Verkuyl, H.J. 1972. *On the Compositional Nature of the Aspects*. Dordrecht: Reidel.
- Vet, C. 1980. *Temps, aspect et adverbes de temps en Franqais contemporain*. Droz: Geneve.
- Wang, L. 1958. *Hanyu sihgao* [The history of Chinese]. Beijing: Zhonghua shuju.
- Wang, Wenyi. 1997. Guanyu dunhuang bianwen liangci yufa gongneng de jige wenti [Some problems about the grammatical function of classifiers from Dunhuang Bianwen], *Guizhou minzu xueyuan xuebao* 4: 44-48.
- Wechsler, S. and Y.-S. Lee 1996. The domain of direct case assignment. *Natural Language and Linguistic Theory* 15:629-664
- Xu, Lili. 1997. Mawangdui hanmu boshu suo jian chengshufa kaocha [Study on the grammar of measuring and counting reflected on silken books from Mawangdui tomb], *Guhanyu yanjiu* 1:19-23, 41.
- Yin, Z. 2000. Dongliangci qianzhi tedian lunlüe [Research on characteristics of preposed event classifiers], *Yufa yanjiu he Tansuo* 9:299-307. Beijing: Shangwu yinshuguan
- Zhang, C. 2010. *Hanyu yuxu de lishi yanbian* [Historical changes of Chinese word order]. Beijing: Beijing Language and Culture University Press
- Zhang, R. 1987. Hanyu yufa fazhan jianshi gangyao [Essentials on the historical syntax of Chinese]. In *Er Wu Shi yufa luncong* [Er Wu Shi's collected essays on Chinese syntax]. Guiyang: Guizhou People's Press.
- Zhuang, Zhengrong. 1980. Shishuoxinyu zhong de chengshufa [grammar of measuring and counting reflected on Shishuoxinyu], *Zhongguo yuwen* 3: 188-193.

Distinction between Mandarin Tones 2 and 3 for L1 and L2 Listeners

Yung-hsiang Shawn Chang

University of Illinois at Urbana-Champaign

Among the four Mandarin lexical tones, Tone 2 and Tone 3 have been found to be the most confusable perceptually for both native speakers and non-native speakers. This study investigates whether the overall tone duration serves as a perceptual cue and whether L1 and L2 listeners weigh the cue in the same manner in distinguishing the two tones. In doing so, two perceptual experiments using duration-normalized stimuli and non-normalized stimuli respectively were conducted on L1 and L2 Mandarin listeners. The results show that native listeners do not heavily rely on duration for the Tone 2- Tone 3 distinction, whereas L2 listeners give more perceptual weight to duration. The overall tone duration is concluded to be an important parameter for L2 tone training and acquisition.

1. Introduction

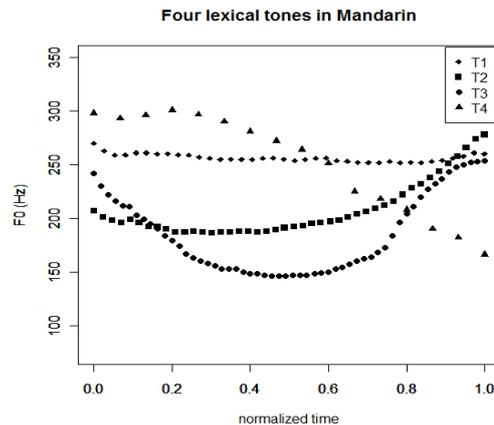
Mandarin is a tonal language and uses pitch to distinguish lexical meanings. Perception studies on Mandarin tones have found the distinction between Tone 2 (T2) and Tone 3 (T3) to be the most difficult for native speakers (Chuang et al. 1972; Shen & Lin 1991; Zue 1976) and non-native speakers (Kiriloff 1969; Wang et al. 1999; Shih et al. 2010) alike. While the F0 turning point and the F0 difference between the tonal onset and the turning point are often manipulated to study the perception of T2 and T3 (e.g., Blicher et al. 1990; Moore & Jongman 1997; Shen 1993), one temporal cue—the overall tone duration—has not received much attention in tone perception studies. On the one hand, T3 is found to be consistently longer than T2 such that duration is suggested to be a perceptually relevant acoustic cue for tone distinction (Blicher et al. 1990). On the other hand, Shih (2007) argued that in conversational speech, T3 is most likely to become the shortest tone because its rising tail is not realized. The conflicting views of overall duration serving as a cue for perception motivates the current study to investigate whether duration indeed plays a role in T2-T3 distinction and whether the durational cue is weighed in the same manner for native (L1) and second language (L2) speakers. To this end, three research questions are asked: 1. Does duration normalization, which keeps all other spectral cues intact, affect the accuracy of T2-T3 distinction? 2. If the performance difference of distinction of duration-normalized vs. non-normalized T2-T3 pairs is reflected in accuracy, is it also reflected in reaction time measures? 3. Do native

speakers and non-native speakers exhibit the same perceptual patterns in T2-T3 distinction?

1.1. Acoustic characteristics of Mandarin T2 and T3

Mandarin T2 and T3 have been traditionally characterized as having a high-rising pitch contour and a low falling-rising pitch contour respectively (Chao 1948). When produced in citation form, T2 and T3 appear distinguishable acoustically, as seen in Figure 1, where all tones are time normalized. In the literature, high confusability for the T2-T3 pair is suggested be due to similar F0 contours between the two tones (Kirilloff 1969; Chuang *et al.* 1972; Gandour 1978; Li and Thompson 1978) or a similar rise in F0 during the final portion of the two tones (Blicher *et al.* 1990).

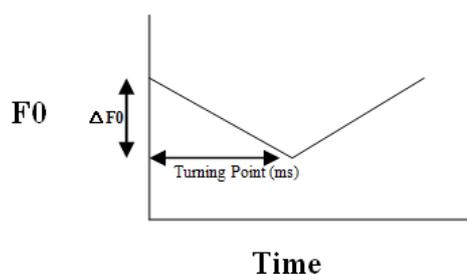
Figure 1. Four lexical tones in Mandarin



To further characterize the two tones, besides both having a concave shape, T2 and T3 was argued by Shen (1990) to differ to a great extent in the timing of the turning point in the pitch contour (schematized in Figure 2). The turning point is claimed to occur close to the onset in Tone 2 and toward the middle of Tone 3. Shen & Lin (1991) tested this claim by having subjects listen to stimuli that varied in the timing of the turning point. They concluded timing dose constitute a perceptual cue for differentiating T2 and T3. Shen & Lin also found that the degree of the initial fall (i.e., F0 difference between the tonal onset and the turning point, shorthand as $\Delta F0$) to be correlated with the turning point, as they reported perception of stimuli with mismatched turning point and $\Delta F0$ to incur more errors. Moore & Jongman (1997) further investigated this potential covariance of the two parameters by using synthetic stimuli, whose turning point and $\Delta F0$ were systematically manipulated, in a perception experiment. They found that both acoustic

dimensions trigger categorical identification and “operate in tandem as perceptual cues to tones” (p. 1871).

Figure 2. Schematized turning point and the degree of the initial fall ($\Delta F0$) for a contour tone (figure reproduced from Moore & Jongman 1997)



While most production and perception studies suggest that turning point and $\Delta F0$ are the main acoustic cues for T2-T3 distinction, Blicher et al. (1990) found syllable duration to also have an effect on perception. In fact, an intrinsic durational difference among four Mandarin tones has been noted as early as in Lin (1965), with T3 being the longest, and T4 being the shortest. An acoustic study by Shen in 1990 confirmed T3 to be consistently longer than T2 when produced in isolation. Assuming that listeners may use this durational difference to enhance the perceptual contrast between T2 and T3, Blicher et al. (1990) varied T2-T3 along a duration continuum and studied whether lengthening would shift the perceptual boundary. Their findings show that listeners tend to produce T3 responses when a stimulus is longer in duration. That is, lengthening may constitute a cue in favor of T3 perception. However, Shih (2007) challenged the view that duration can be a reliable cue for tone identification by pointing out the divergent reports in the literature on Mandarin tone duration. She argued that T3 is longer than T2 only when produced in isolation; in conversational speech, T3 most likely becomes the shortest among all four tones because its rising tail is not realized.

A difference in intrinsic amplitude was found for the four Mandarin tones in Chuang et al. (1997), with Tone 3 having the lowest amplitude, and Tone 4 the highest. In this regard, the contribution of the amplitude contour to Mandarin tone identification has also been studied in the literature. The findings diverge with respect to the effect of amplitude contours on tone perception. Lin (1988) manipulated the amplitude contour but found no effect in the presence of F0. In the absence of F0, however, Whalen & Xu (1992) found amplitude contours to be a useful cue for their subjects in identifying Mandarin tones, although their results do not particularly generalize the effect of amplitude to T2-T3 distinction. Fu et al. (2003) also studied how the amplitude cue contributes to tone perception. They concluded that the amplitude cue only contributes to T3 and T4

discrimination and that use of amplitude in tone recognition has a large variability across listeners.

Besides F0 and temporal envelop cues, Yang (1989) reported that vowels have an effect on T2-T3 distinction, although such an effect was not observed in discrimination of other tone pairs. Yuan (2003) followed up on this study using selective adaption tests to see if tone identification is influenced by repeated presentation of a stimulus that shares similar acoustic features. He found that only the *da2* and *da3* adaptors, but not *bi2* and *bi3*, had a selective adaptation effect on the identification test on the *da2-da3* continuum. Based on this discrepancy, Yuan suggested that tone perception may not be independent from segmental perception.

One acoustic attribute oftentimes associated with T3 that is considered extralinguistic in Mandarin tones is creakiness (or laryngealization). Yu (2010) reported in her corpus study that laryngealization occurred most frequently for T3 and 68% of the time across speakers. While Davison (1991) and Yu suggest laryngealization as a feature to aid tonal perception, Yang (2011)'s study provides evidence for categorical perception in discriminating all tonal pairs over continua, varying voice qualities but holding F0 constant. That is, voice quality can contain functions beyond auditory enhancement in Mandarin tone perception.

1.2. Native vs. non-native speakers' perception of lexical tones

Many studies found that only native listeners exhibit categorical perception of lexical tones, whereas non-native listeners do not (e.g., Hallé *et al.* 2004; Xu *et al.* 2006). Therefore, for tones to be used in speech communication, they must be perceived by the hearer as linguistic objects instead of musical pitches. Before learners acquire phonemic boundaries between tones, any within category phonetic variation may be misperceived as linguistically relevant. In addition to a psychoacoustic approach of studying lexical tone perception, a flux of recent research pursue the same line of research using neurophysiological methods and has reported neural response differences between native and non-native speakers of tone languages. For example, in Zatorre & Gandour's (2008) fMRI study, they found that English speakers who have no prior experience with a tone language exhibit a lack of left-hemisphere dominance in the processing of lexical tones. In contrast, there is strong left-hemisphere involvement in native speakers' processing of lexical tones. However, the processing difference between native speakers and non-native speakers is not strictly dichotomous but has scalar dimensions. Many studies (e.g., Bent 2005; Huang 2004) that investigate native vs. non-native speakers' perception of lexical tones found that L1 and L2 listeners weigh various acoustic cues differently, which results in different perceptual patterns. In distinguishing Mandarin tones, Gandour (1978) pointed out that tone height and direction are the most important dimensions. He found that his native speaker subjects attended more to the direction dimension whereas the non-native counterparts directed more attention to the pitch height of the stimuli. However, the non-native like perceptual system is not unmodifiable. Many auditory training studies show that given intensive laboratory training, L2 learners improved in

perception and production of non-native sounds as their cue weighting approximates that of native speakers', both at segmental and suprasegmental levels. Especially with regards to Mandarin tone perception, Wang et al. (1999) reported significant improvement for their American English subjects over a short training period. And the suprasegmental perceptual modification was even retained six months after training.

1.3. Operationalization of the study

The current study investigated whether duration is used as a perceptual cue for Mandarin T2-T3 distinction in addition to F0 information. Although Blicher et al. (1990) has reported that syllable lengthening would enhance the percept of T3 for native listeners, we are not certain as to whether when the durational difference between T2 and T3 is removed, T2-T3 discriminability would be similarly affected for native and non-native listeners. To answer this question, two auditory discriminations tasks were conducted on native listeners and L2 learners of Mandarin. Our hypothesis was that listeners would have a lower accuracy score or a longer reaction time in distinguishing the stimuli with the duration cue removed. To avoid the practice effect being a confounding factor, we gave the task with unaltered stimuli (where T3 was naturally longer than T2) first. If the practice effect is strong, we might see both groups perform better in the second task or we might see no statistically significant performance difference between the two tasks. If we observe significantly lower accuracy score and/or longer reaction time in performing the second task, we could conclude that the effect of duration as an auditory enhancement is robust enough to override any practice effect.

2. Methodology

2.1. Auditory stimuli

Six native speakers of Mandarin (3 male and 3 female) were recruited to record the stimuli for this study. They were graduate students at the University of Illinois at Urbana-Champaign, aged between 25 and 30. All recordings were conducted in a sound-treated booth in the Phonetics Lab at UIUC. An AKG C520 head-worn condenser microphone was used to record acoustic signals onto a Marantz PMD570 recorder. The stimuli were two syllables *ma* and *da* carrying four Mandarin tones. Each syllable-tone combination was repeated five times in a randomized order. The stimuli were visually presented to the speakers in E-prime.

All recordings were later screened and only one female speaker's data was used because 1. she had no creaky voicing in her T3 production, which could be a potential cue for T2-T3 distinction, and 2. her T3 production was fully realized. We did not use more than one speaker's data as perception of tone has been suggested to be a talker-contingent process (Moore & Jongman 1997). At the end, only two sweeps of each syllable-tone combination was chosen out of five repetitions to limit the number of stimuli pairing.

Duration normalization was performed using the PSOLA function in Praat. This method preserves the spectral structure of the syllables regardless of modified duration values. The duration changes were made consistently in that the longer syllable was shortened to match the duration of the shorter syllable within a tone minimal pair. More specifically, *ma2* and *ma3* were all duration-normalized to 470 ms while *da2* and *da3* to 370 ms. The stimuli were normalized for RMS amplitude at 60 dB. At the end of post-processing, the stimuli were screened by two native speakers for accuracy and all of them were correctly identified.

2.2. Subjects for the discrimination tasks

Eight Mandarin native speakers and 8 non-native speakers participated in the discrimination tasks. The native speakers included 3 males and 5 females, aged between 24 and 30. The non-native speakers were intermediate-level Mandarin learners. They included 2 males and 6 females, aged between 20 and 24.

2.3. Experiment procedure

The experiment consisted of 2 forced-choice discrimination tasks. Task 1 used the stimuli where T3 was naturally longer than T2; task 2 used duration-normalized stimuli. In each task, 30 items (16 T2-T3 pairs and 14 non-T2-T3 pairs as distractors) were randomized and presented in E-Prime. All our stimuli can be represented by various lexical items. Since we were not particularly looking into the lexical frequency effect on the T2-T3 distinction, all stimuli were presented auditorily and no orthographic information was provided. The inter-stimulus interval was set at 1.3 seconds, in which there was a 1 second beep. Both accuracy and reaction time were logged automatically in E-Prime. Subjects were asked to press the button Y when the tones in the word pairs were the same and N for different tones. The subjects were encouraged to respond as quickly and accurately as possible.

3. Results

For the purpose of this paper, only the T2-T3 discrimination data were analyzed. Table 1 shows the percentage accuracy scores in the two tasks and for both native and non-native listener groups. We see that the native group's performance almost reached the ceiling. Non-native group had a rather high accuracy as well, although with a greater variance. The result is generally in line with Shih & Lu's (2010) tone training study in that the only few tone discrimination mistakes native listeners made were all on T2-T3 pairs and that T2-T3 distinction yielded most of the errors for L2 learners of Mandarin. In task 2, both native and non-native groups scored lower than in Task 1, where the durational cue was present. For statistical analysis, a mixed-design ANOVA was run with group as a between-subject variable, and task type as a within-subject variable. The results revealed significant main effects of group ($F(1,14)= 8.34, P=.012$), task ($F(1, 14)=33.51, p <0.001$), and task*group interaction ($F(1, 14)=12.064, P<.01$). Pairwise comparisons with a Bonferroni correction exploring the task*group interaction effect

showed a significant difference between two groups in Task 2 ($p < .05$), but not in Task 1. For the native listener group, there was no significant difference between the accuracy scores in Task 1 and Task 2 ($p=0.0796$), while for the non-native group, a significant difference was observed ($p < .001$). The results suggest that our intermediate-level learners performed no differently from their native counterparts in Task 1, where the T3 tokens were all naturally longer than T2. In contrast, in Task 2, where duration normalization was performed on the stimuli, L2 learners did significantly worse than L1 speakers.

Table 1. Accuracy score (%) for native and non-native listeners in 2 tasks

	Native	Non-native
Task 1	0.99 (SD=0.02)	0.95 (SD=0.06)
Task 2	0.97 (SD=0.047)	0.86 (SD=0.07)

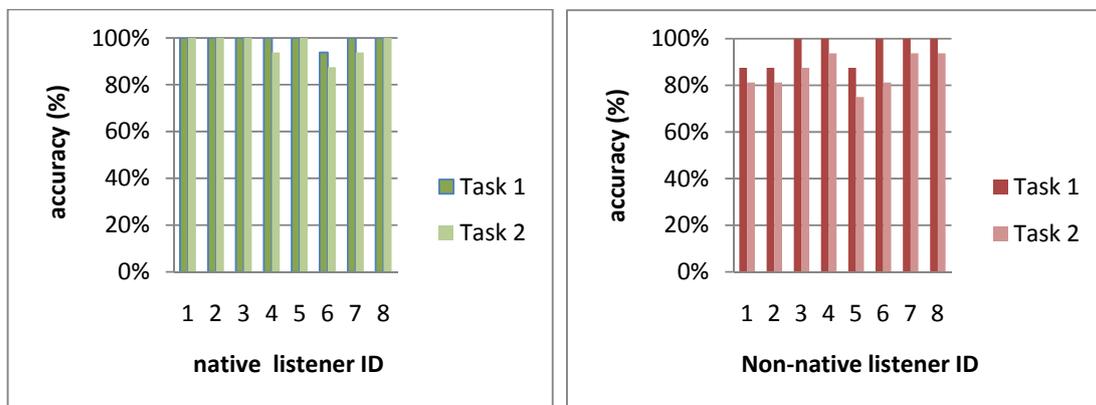
In terms of reaction time, the descriptive statistics are presented in Table 2. The reaction time data exhibits the same pattern as the accuracy data in that the native listener group has a lower mean reaction time than the non-native group in both tasks. And more reaction time was needed for both native and non-native listeners when the durational cue for T2-T3 distinction was absent. Note that the standard deviation is fairly large for both groups in both tasks. We reason that some listeners had decided on the answer as soon as the second word finished playing, while other listeners could have taken their time to retrieve the first word they heard after hearing the second word. Once again, a mixed-design ANOVA was conducted on the reaction data. The results showed a significant main effect of group ($F(1,14)= 4.695$, $P < .05$) and task ($F(1, 14)= 12.099$, $p < .01$), but no significant task*group interaction. That is, native listeners had a shorter reaction time than their non-native counterparts in both tasks. Absence of the durational cue in distinguishing T2 and T3 would induce longer reaction time for both groups.

Table 2. Reaction time (ms) for native and non-native listeners in 2 tasks

	Native	Non-native
Task 1	514.6 (SD=110.69)	720.23 (SD=217.12)
Task 2	573.17 (SD=128.46)	754.85 (SD=234.05)

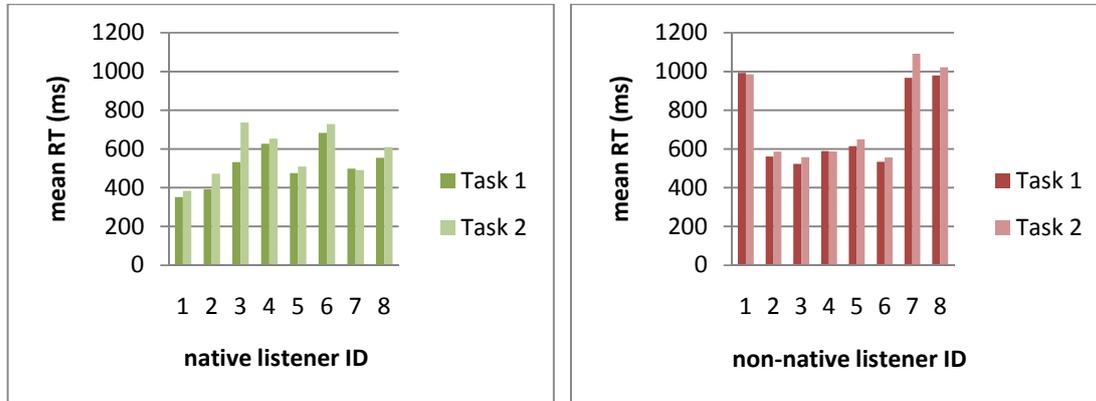
Besides the group mean performance, we were also interested in within-group variability to see if there is any perceptual pattern within or across individuals in the two tasks. In figure 3, all listeners' scores in Task1/Task2 were plotted back-to-back. We see that almost all of our native listeners reached the ceiling in the first task. In the second task where the stimuli were duration normalized, three listeners scored lower. As for our non-native listeners, five out of eight scored a hundred percent in the first task. No one got all the items correct in the second task; in fact, all non-native listeners' scores dropped in Task 2.

Figure 3. Individual accuracy scores (%) in both tasks



The individual reaction time data was plotted in Figure 4. For both native and non-native listeners, their reaction time mostly ranged between 400-600 ms, except for non-native listeners 1, 7, 8. Note that non-native listeners 7 and 8 had the highest accuracy scores compared to other non-native listeners. Therefore, there could be some accuracy-speed trade-offs for these two listeners. From Figures 3 and 4, we generally observe a longer reaction time across subjects in the second task. Only native listener 7 and non-native listeners 1 and 4 show the opposite pattern. However, the reaction time difference between the 2 tasks for these speakers was noticeably smaller than for all other speakers such that this divergence may be negligible.

Figure 4. Individual reaction time (ms) in both tasks



Although orthographic information was purposely not provided for the auditory stimuli in the experiment, we seem to observe a lexical frequency effect on the T2-T3 distinction. In analyzing the T2-T3 discrimination data, we pooled the data of *ma* and *da*. Upon a closer look at the *ma* and *da* data respectively, we see more errors on *da2-da3* pairs for both native listeners and non-native listeners. We have two speculations regarding this pattern. First, the lexical frequency difference between the *ma* and *da* syllables may play a role. Also, although the stimuli were correctly identified by two native speakers, whether all tone productions were the equally same/different has not been acoustically quantified. That is, it may well be the case that the pitch contours of *da2-da3* were more similar, and hence more confusable, than *ma2-ma3*.

4. Discussion and Conclusion

Previous literature on Mandarin tone perception generally agrees upon F0 being the most important cue in the T2-T3 distinction. Given that T3 is consistently longer than tone 2 in citation form, Blicher et al. (1990) tested whether duration difference can be used to enhance the perceptual contrast. While Blicher et al. did find syllable lengthening to auditorily enhance the percept of T3, the current study took a step further to investigate the role of duration in T2-T3 distinction. Our results show that native speakers' accuracy did not suffer in discriminating two duration-normalized tones. However, the reaction data indicates that native listeners would take significantly more time to respond when the durational cue was not available. In contrast, non-native listeners' performance exhibited significantly lower accuracy and significantly longer reaction time where the durational cue was absent. Our findings suggest that native and non-native listeners weigh the durational cue differently. For native listeners, syllable duration may be an enhancing cue that facilitates T2-T3 distinction. For non-native listeners, however, duration serves more as a primary cue in addition to F0. This psychoacoustic account for the performance difference between native and non-native listeners as well as natural stimuli and duration-normalized stimuli is compatible with previous L2 acquisition

literature in that native speakers and non-native speakers give different weight to various acoustic properties in perception (e.g., Bohn 1995; Flege et al. 1997).

In terms of L2 acquisition of lexical tones, some pedagogical implication can be drawn from the different perceptual patterns observed in our L1 and L2 listeners. It is not uncommon that in teaching Mandarin T3, language instructors generally exaggerate the syllable duration in order to fully realize its F0 contour. As a result, length differences can be taken as one of the primary cues for distinction between T2 and T3. Indeed, our study found that duration only appeared to be an enhancing cue for our native listeners, whereas L2 listeners' discrimination accuracy dropped significantly in the absence of the durational cue. Therefore, giving duration-normalized T2-T3 pairs may allow L2 learners to improve in perception as their cue weighting is shifted from duration to the F0 dimensions.

At last, we acknowledge that some native speakers do consistently produce T3 shorter than T2 even in isolation, as their T3 is not fully realized. It is unclear to us whether they also perceive T3 to be shorter than T2. In future studies, it will be of our interest to include another experiment where T3 tokens are shorter than T2 and see whether this durational warping will hinder or facilitate the discrimination of T2 and T3.

REFERENCES

- Alexander, J., Wong, P. C. M., & Bradlow, A. (2005). Lexical tone perception by musicians and non-musicians. *Proceedings of Interspeech 2005*.
- Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, 18(1), 37-49.
- Bent, T. (2005). *Perception and Production of Non-Native Prosodic Categories*. Unpublished Ph.D. dissertation, Northwestern University, Chicago.
- Bohn, O.-S. (1995). Cross-language speech perception in adults: first language transfer doesn't tell it all. In Strange, W. (ed.). *Speech perception and linguistic experience: theoretical and methodological issues*. Timonium, MD: York Press.
- Chao, Y. R. (1948). *Mandarin Primer*. Cambridge: Harvard University Press.
- Chuang, C.-K., & Hiki, S. (1972). Acoustical Features and Perceptual Cues of the Four Tones of Standard Colloquial Chinese. *The Journal of the Acoustical Society of America*, 52 (1A), 146.

- Davison, D.S. (1991). An acoustic study of so-called creaky voice in Tianjin Mandarin. *Working Papers in Phonetics, UCLA*, 78. 50–77.
- Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In *Speech perception and linguistic experience: issues in cross-language research* (W. Strange, ed.), pp. 233–277. Baltimore: York Press.
- Flege, J. E., Bohn, O-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437–470.
- Fu, Q.J., Zeng, F.G. 2000. Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech Language and Hearing* 5, 45-58.
- Gandour, J. (1978). The Perception of Tone. In V. A. Fromkin (Ed.), *Tone: A Linguistics Survey* (pp. 41-76). New York: Academic Press.
- Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12(2).
- Hallé, P. A., Chang, Y., & Best, C. T. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32 (3). pp. 395-421.
- Howie, J. M (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge University Press, Cambridge.
- Huang, T. (2004). Language-specificity in auditory perception of Chinese tones. Unpublished Doctoral Dissertation, Ohio State University, Columbus.
- Kiriloff, C. (1969). On the Auditory Perception of Tones in Mandarin. *Phonetica*, 20(2-4), 63-67.
- Li, C. N., & Thompson, S. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4(2), 185-199.
- Lin, M.-C. (1965). Yingao xianshiqi yu Putonghu shengdiao yingao texing [The pitch indicator and the pitch characteristics of tones in Standard Chinese]. *Shengxue Xuebao* [Chinese Journal of Acoustics] 2 (1), 8–15.
- Lin. M.-C. (1988). Pulonghua silcngdiao de shengxue texing he zhijue zhengzhao [The acoustic characteristics perceptual cues of tones in Standard Chinese]. *Zhongguo*

- Yuwen [Chinese Language], 204. 182- 193.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of Acoustical Society of America*, 102, 1864-1877.
- Shen, X-N. S. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics*, 18, 281-295.
- Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech*, 34, 145-156.
- Shen, X. S., Lin, M., & Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America*, 93, 2241.
- Shih, C. (2007). *Prosody learning and generation*. Berlin: Springer.
- Shih, C., Lu, H.-Y., Sun, L., Huang, J.-T., & Packard, J. (2010). An Adaptive Training Program for Tone Acquisition. In *Proceedings of Speech Prosody 2010*.
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, 106, 3649-3658.
- Wang, Y., Sereno, J. A., Jongman, A. & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *J. Cogn. Neurosci.* 15, 1019–1027.
- Whalen, D.H., Xu, Y. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49, 25-47
- Xu, Y., Gandour, J., & Francis, A. (2006). Effects of language experience and stimulus complexity on categorical perception of pitch direction. *Journal of the Acoustical Society of America*, 120 (2). pp. 1063-1074.
- Yang, R. (2011). The Phonation Factor in the Categorical Perception of Mandarin Tones. In *Proceedings of ICPHS XVII*.
- Yang, Y. (1989). Yuanyin he Hanyu Shengdiao de Ganzhi Wenti [Vowels and the Perception of Mandarin Tones]. *XinLi XueBao [Acta Psychologica Sinica]*, 34

(1).

- Yu, K. M. (2010). Laryngealization and features for Chinese tonal recognition. In *Proceedings of Interspeech 2010*.
- Yuan, J. (2003). Selective Adaptation Tests on Mandarin Tone2 and Tone3. *Proceedings of ICPhS XV*, 1719-1722.
- Zatorre, R. J., & Gandour, J. T. (2008) Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society Biological Sciences*, 363 (1493). 1087-1104.
- Zue, V. W. (1976). Some perceptual experiments on the Mandarin tones. *Journal of Acoustical Society of America*, 60 (1), 45.

Representation of Mandarin intonations: boundary tone revisited

Ping Jiang Aishu Chen

The Chinese University of Hong Kong

In order to find where exactly the Mandarin interrogative cues are located and whether Mandarin question intonation could be represented using a boundary high tone (H%), we conducted a perceptual experiment, asking 10 native listeners to identify manipulated Mandarin utterances with question or statement intonations. Listeners judged four types of stimuli, which are full utterance, utterance without the last syllable, utterance without the last prosodic word, utterance remaining the first prosodic word. The results show that the correct question identification percentage drops significantly when the last prosodic word is missing, but the majority of listeners are still able to identify the questions correctly, even when the stimulus only contains the first prosodic word. This result supports the claim that Mandarin interrogative cues distribute over an entire utterance. High pitch in both edges of the intonation phrase should be marked out. *

1. Introduction

Boundary tone, proposed by Pierrehumbert (1980) as a phonological unit for representing the internal makeup of an intonation phrase, has been disputed in the study of Mandarin intonation. Most scholars argue that the boundary tone described as the element of intonation does not exist in Mandarin as the contrasts of Mandarin sentence moods are not differentiated by the “pitch height of the sentence-final syllable”. Instead, it is the “pitch register contrast over whole utterance domain” that signals Mandarin intonations (Ho 1976, 1977, Shi 1980, Shen 1985, Shen X.N. 1990, Yuan, Chih and Greg 2002, Lee 2005). As early as 1930s, Chao has proposed his “simultaneous addition” hypothesis to describe intonation’s register adjusting effect on tones. He says that there is a general raised or lowered levels of pitch when imposing sentence intonations on connected lexical tones, and the pitch range over whole utterance domain will be widened or narrowed (Chao 1932, 1933, 1968). Therefore, the concept of boundary tone, if it exists

in Chinese, only represents boundaries of prosodic units (something like a break between prosodic units), but does not represent intonations which convey pragmatic functions of sentences (Shi F. 2010). There are, however, exceptions which support that the boundary tone could be used in Mandarin to represent intonations. Lin (2004, 2006) argues that “information about question or statement is carried by an overwhelming majority of the last one or two syllables without neutral tones in the final prosodic word.” He proposes using a boundary high tone to represent question intonation in Mandarin. Lin is not the only one who insists sentence-final syllables’ decisive roles in carrying intonational meanings. Hu (1987), without instrumental help, summarizes the last syllables’ distinctive features in different intonation environments. He thinks the features of sentence-final syllables’ pitch, duration and intensity are the key factors in describing different types of intonations. Following Hu’s claim, He Y. and Jing S. (1992) did acoustic measurements of the sentence-final stressed syllables in statement, question and imperative intonations’ environments, and they reported the acoustic data of the final syllable’s pitch, duration and intensity in different sentence moods. Neither Hu nor He Y. has proved the last syllable’s independent function in distinguishing different sentence types before they proceed to investigate the final syllable’s phonetic features. Lin, Hu and He.Y all adopt some ideas from Chao (1933)’s “successive addition hypothesis” which also describes pitch interactions between lexical tones and sentence intonations. Different from his “simultaneous addition”, this hypothesis proposes that sentence intonation is like a rising or falling pitch tail attached to the original sentence-final lexical tones’ contours, thus the original tones’ shapes are changed. Lin, Hu and He.Y acknowledge that intonation’s major modification domain is in the last syllable, but they do not agree that lexical tone shapes were changed by sentence intonations. In their analysis, intonations only raise or lower the pitch registers of the last syllables. Boundary tone is also adopted in recent ToBI conventions for Mandarin intonation transcription. Peng et al (2005) use a boundary high tone (H%) to transcribe Mandarin echo question and yes-no question with sentence-final particle like “ma”. However, they mark out the overall pitch level rising at the same time, by a tag “%q-raise”, which means the boundary high tone at the right edge of the question utterance is not the only high pitch feature that carries the interrogative information, it combines with the overall pitch raising to signal questions.

As discussed above, it seems that three disputed views still exist regarding boundary tone’s status in Mandarin intonation representation. The first view thinks it is not necessary to include the boundary tone as a unit in the phonological representation of Mandarin intonation. In great contrast, the second view considers boundary tone as an indispensable element, almost all the necessary phonetic cues for signaling sentence moods are included in the last syllable’s domain, or, the weak version of this view thinks

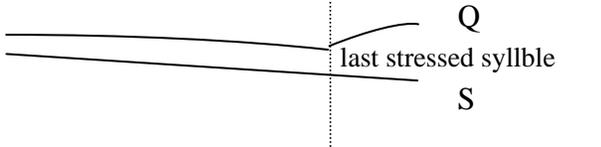
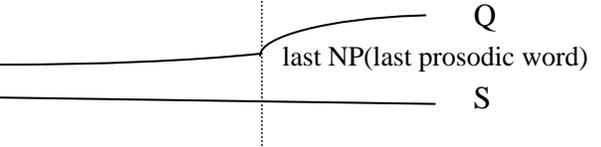
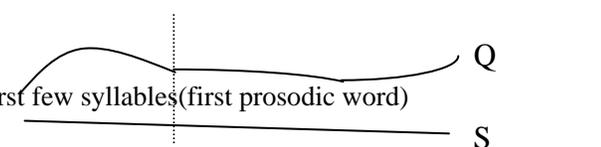
that terminal high pitch combines with the overall high pitch to signal questions. The third view acknowledges boundary tone's existence. However, boundary tone has nothing to do with sentence moods in this proposal, it only indicates pause or break between prosodic units. In order to make this controversial issue more clear, the current paper is going to discuss whether the "boundary high tone" is a proper phonological component of Mandarin interrogative intonation by looking into some experimental evidences.

2. Purpose and methods

2.1 The goal and reasoning of the experiment

The goal of conducting the perceptual experiment is to test listeners' judgments on manipulated incomplete question utterances, so as to find out in which part of the intonational phrase the overwhelming interrogative cues locate. Here, based on previous studies, we assume the major acoustic correlate of intonation is pitch (Howie 1976, Tseng 1981, Shen 1990). High or raising pitch will be interpreted as having interrogative meanings (Ohala 1983, 1984), so the major interrogative cues are located in the part of the question which has most significant higher pitch contrast to its statement counterpart. However, previous production studies reported different results about this most significant pitch rising domain. Schemas listed in figure1 compare those results.

figure1: the domain of the most significant pitch expansion and rising in questions.

different findings from production data	schematized representation of that findings
Lin (2004, 2006): last stressed syllable	
Lee (2005): last NP	
Shen X.N.(1990); Shi(1980): first few syllables	

Lin (2004, 2006) reported that the localized most significant pitch raising occurred in the sentence-final stressed syllable. The F0 values at the starting point and the ending point of sentence-final lexical tones are higher in questions than in statements. The F0 slopes' values of the four lexical tones increase in questions. Lee (2005)'s acoustic study on un-marked yes-no questions found that the most significant pitch rising and expansion occurred in the last noun phrase (last prosodic word) domain. Shen X.N. (1990) and Shi (1980) both mentioned the starting few syllables in questions had gone through notable pitch rising, so the first prosodic unit is also a suspect domain for major interrogative cues location. Based on their findings on production data, we hypothesized three possible locations of the interrogative cues in question utterances: 1) the last stressed syllable, 2) the last prosodic word, 3) the first prosodic word. If we take off the hypothetic portion of the question utterance which may carry most of the interrogative cues, native listeners' correct identification rate will drop significantly. By this way, we could locate the interrogative cues by observing listeners' performances.

2.2 Methods

2.2.1 What could be improved in previous perceptual experiments

Some perceptual studies on question intonations have been done. Experimenters used synthesized, manipulated and natural tokens for different research purposes. However, there are some insufficiencies in the previous stimuli designs.

First, the semantic meanings of utterances become incomplete after manipulation. For example, Lin (2006)'s cutting manipulation break the meaning of the last disyllable word; Wu, Tao and Lu (2006) cut off question markers in the middle part of the utterance, leaving some of the stimuli like fragments. We may question that the low ratios of questions identifications are because of listeners' confusions about the utterance meanings rather than lack of interrogative cues in the stimuli.

Second, little attention has been paid to the tones of the ending syllables in previous manipulated utterances. However, some perception studies have found the influence of sentence-final tones in intonation identification (Yuan, 2004 2006). If there is no variation of sentence-final tones, the perception results are questionable. In Lin(2006)'s perception study, when the final object is cut off, the utterances for identification are always ended in falling lexical tone of the verb 去 “qu4”. The low correct rates of questions identifications may be influenced by the pitch interactions of ending lexical tones rather than influenced by loss of interrogative cues.

Third, studies using synthesized materials for identification (e.g., Liang and Van, 2007) generated different pitch patterns by adjusting both the boundary point F0 value and the overall F0 level value. The problem with this synthesizing method is that it limits

the localized significant pitch rising effect in the last syllable domain without proving the real localized effect domain in advance.

Although there are some weaknesses in the previous studies, these investigators laid down solid foundations in perception studies looking for interrogative cues. Based on previous studies, the present study tries to design a corpus taking into considerations the above mentioned problems.

2.2.2 Our corpus design

A corpus of 64 utterances was built up. The original recording materials are four pairs minimally contrasted simple declaratives and their corresponding unmarked yes-no questions without narrow focus. The recording materials are listed below (with pinyin transcriptions and numbers to indicate four citation tones in Mandarin):

- 1) ma3xiao3ming2chi1han4bao3bao1. “Ma Xiaoming eats hamburger.”
- 2) wang2jiao4shou4jiang3xin1wen2xue2. “Prof. Wang gives lectures on journalism.”
- 3) lin2xiao3jie3lai2yun4dong4chang3. “Miss Lin comes to the playground.”
- 4) zhang1xian1sheng0zai4ban4gong1shi4. “Mr. Zhang is in the Office.”

Four citation tones and the neutral tones were mapped into four specific positions in the utterances, they are: the last syllable of the first prosodic word, the syllable before last prosodic word, the syllable before last syllable and the utterance-final syllable.

2.2.3 Recording

All these sentences were recorded from a native female speaker of standard Mandarin Chinese. The sentences were randomized and presented one by one to the speaker in a computer screen. The speaker was asked to say the sentence naturally without any emphasis. She was reminded to pay attention to the punctuations and say the sentences in appropriate intonation. She was required to repeat each utterance for three times. We choose the one which is more natural and clear from the second or third time repetition. The recording was carried out in a quiet room by means of a high quality microphone connected to the computer. The speech was recorded with a sampling rate of 16000 Hz. by a computer software called “Cool Edit Pro 2.0”. The recorded materials were stored in the computer for further manipulation.

2.2.4 Manipulation

We manipulated the original sentences using Praat5.1.23. The semantic meaning of each token is kept intact after manipulation and the four lexical tones were mapped onto the

end of the manipulated utterance. Table1 shows the meaning and tone mapping of the tokens after manipulation (with pinyin transcriptions and numbers to indicate tones):

table1: experiment design

Remain the first prosodic word(I)	Cut off the last prosodic word(II)	Cut off the last stressed syllable(III)	Utterance in full-length (IV)	Last syllable tone mapping			
				I	II	III	IV
马小明 ma3xiao3ming2 Ma Xiaoming (name of a boy)	马小明吃 ma3xiao3ming2chi1	马小明吃汉堡 ma3xiao3ming2chihai4bao3	马小明吃汉堡包 ma3xiao3ming2chihai4bao3	R	H	L(R)	H
林小姐 lin2xiao3jie3 Miss Lin	林小姐来 lin2xiao3jie3lai2	林小姐来运动 lin2xiao3jie3lai2yun4dong4	林小姐来运动场 lin2xiao3jie3lai2yun4dong4chang3	L(R)	R	F	L(R)
王教授 wang2jiao4shou4 Prof. Wang	王教授讲 wang2jiao4shou4jiang3	王教授讲新闻 wang2jiao4shou4jiang3xinwen2	王教授讲新闻学 wang2jiao4shou4jiang3xinwen2xue2	F	L(R)	R	R
张先生 zhang1xiansheng0 Mr. Zhang	张先生在 zhang1xiansheng0zai4	张先生在办公 zhang1xiansheng0zai4ban4gong1	张先生在办公室 zhang1xiansheng0zai4ban4gong1shi4	N	F	H	F

After manipulation, we got 32 utterances. Each token was repeated one time and stored, we obtained 64 tokens in total (4 utterances * 2 intonation types * 4 ways of manipulation * 2 repetitions = 64 stimuli).

2.2.5 Further processing of the stimuli

The 64 tokens were grouped into 4 separated sound files by syllable numbers. In each sound file, the utterance orders were randomized. Each stimulus was repeated once again and put together with 1s intervals. The inter-stimulus interval (pause between offset of the second repetition of the previous stimulus and the beginning of the next stimulus) was fixed at 3s. For example, in the sound file containing three syllable stimuli, listeners will hear:

“wang2jiao4shou4 (1s interval) wang2jiao4shou4 (3s interval) lin2xiao3jie3 (1s interval) lin2xiao3jie3 (3s interval).....”

2.2.6 Listening test

The test was carried out in a quiet office in The Chinese University of Hong Kong. 10 native listeners took the perception test one by one. The experimenter played 4 sound files to the listeners through a lap-top and a headphone. The playing order was from 3

syllables file to 4 syllables file, then 6 syllables file and finally 7 syllables file. By this order, listeners could not guess the intonations of short stimuli from longer stimuli which contain more phonetic information. Before the real experimental section, there is a practice section including three stimuli. There are 16 trials in each experimental section. Subjects were required to choose between statement and question by marking their choices in four separated answer sheets provided to them. The test was a forced-choice task. The whole process took about 20 minutes. Only questions identification results were analyzed statistically.

3. Results and discussions

3.1 Results

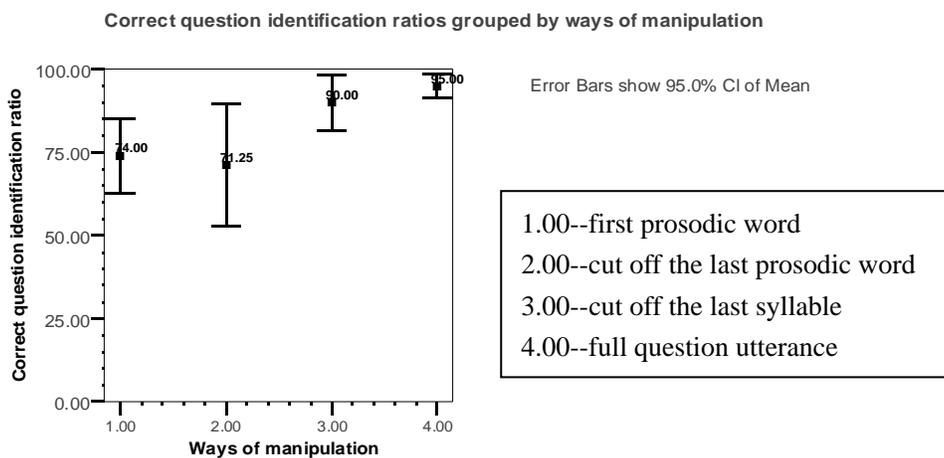
Listeners' questions identification results are listed in table2 and figure2. Table 2 shows the mean correct questions identification results on four kinds of stimuli, while figure2 shows the mean correct identification rates plus one positive and one negative standard deviations. The deviation values let us know the data distribution patterns.

table2: listeners' mean correct questions identification results

<u>Utterance in full length</u>	<u>Cut off the last syllable</u>	<u>Cut off the last prosodic word</u>	<u>Utterance remaining the first prosodic word</u>
马小明吃汉堡包?	马小明吃汉堡?	马小明吃?	马小明?
林小姐来运动场?	林小姐来运动?	林小姐来?	林小姐?
王教授讲新闻学?	王教授讲新闻?	王教授讲?	王教授?
张先生在办公室?	张先生在办公?	张先生在?	张先生?
Mean correct rates: 95.00	Mean correct rates: 90.00	Mean correct rates: 71.25	Mean correct rates: 74.00

We found from table2 that mean correct questions identification rates on “utterance in full length” and “utterance without the last stressed syllable” are close to each other, both are around 90%. Even for full utterances, listeners' identifications are not 100% correct. On the other hand, mean correct questions identification rates on “utterance without the last prosodic word” and “utterance remaining only the first prosodic word” are close to each other, both are around 70%. When the last prosodic word is missing, the mean correct identification rate drops about 20%. All the correct identification rates are above chance level (50%).

figure2: mean correct questions identification results, with two standard deviations



The numbers 1~4 represent four ways of manipulation. Figure2 tells us that:

1) 10 listeners' correct questions identification rates for the "full utterance" has little variations, most of the data distribute in the 90~100% interval. It means that most of listeners have no problem in figuring out the question intonation.

2) 10 listeners' correct question identification rates for the "utterance without the last syllable" has the second smallest variations, most of the data distribute in the 80~100% interval. It means that listeners have a little difficulty in figuring out the question intonation when the last syllable is missing. However, most of them could catch the interrogative meaning from the "utterances without the last syllable".

3) 10 listeners' correct question identification rates for the "the utterance without the last prosodic word" has the largest variations. It means that individual differences exist among listeners. Some of them could get the question interpretation; others are confused in judging "the stimuli missing the last prosodic word".

4) 10 listeners' identification rates for the "the first prosodic word" also has relatively large variation. But the variation is smaller than group 2. It means that listeners have individual differences in identifying "utterance remaining only the first prosodic word", but the differences are smaller than group2's.

5) There is few overlap of group2 and group4's error bars. That means data in group2 contrasts statistically significantly to data in group4. Listeners' correct question identification rate drops significantly when the last prosodic word is missing.

6) The bar of group 4 overlaps with the bar of group3. That means data in these two groups have no significant difference, taking off the last syllable did not cause many misjudgments of the question intonation.

3.2 Discussions

3.2.1 Marking out boundary high tones according to interrogative cues' locations

The results in 3.1 suggest two major findings: first, it is the final prosodic word rather than the sentence-final stressed syllable which carries most of the interrogative information. Second, both the global pitch level rising and the localized significant pitch rising at the final prosodic word contribute to the question detection.

Our result supports Lee (2005)'s acoustic measurement result, i.e., pitch rising begins from the first syllable of the question and goes to the highest place in the sentence ending place. Significant pitch swing occurs in the last prosodic word, thus the localized pitch rising and expansion carry most of the interrogative information. The perceptual data shows that listeners' correct identification rates amount to 74% even when the utterance remains only the first prosodic unit. Therefore, the upward pitch moving in front portion of the sentence also contributes a lot to signaling questions.

The result did not support the opinion that sentence-final syllable is the only carrier of question intonation. Cutting off the last stressed syllable does not influence the correct question identification a lot. Mandarin question intonation is different from intonation language like English which signals questions by boundary syllables' pitch shapes.

In syntactically un-marked yes-no questions, interrogative cues are distributed over the whole sentence rather than limited to the boundary stressed syllable at the end point of the utterance. Therefore, if we equals a "boundary high tone (H%)" to a phonological representation of "high or rising pitch which conveys the question meaning", an adequate representation of Mandarin question intonation should mark out the beginning high tone of question sentence in addition to the ending high tone. A syntactically un-marked question utterance which contains only one intonation phrase should be marked as:

"H% [intonational phrase] H%".

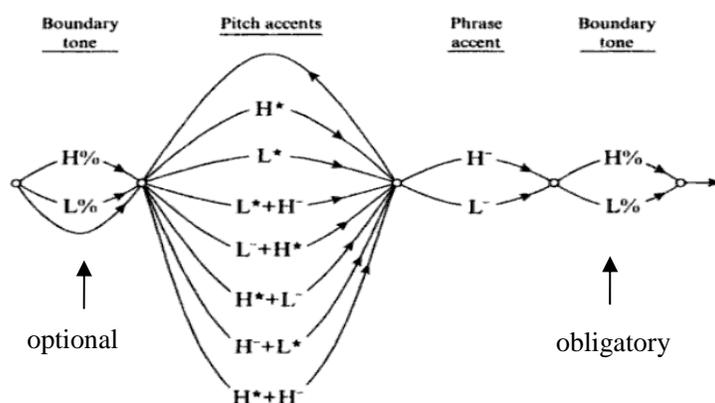
The phonetic realizations of these two boundary high tones lead to the pitch rising in the entire question utterance, and the pitch rising for the prosodic units at the beginning part and the ending part of the questions are more significant because they are near the high tone target.

3.2.2 Whether boundary tone is suitable for representing Mandarin intonation

The above analysis may be invalid if the notion of "boundary tone" has nothing to do with the sentence moods. If the boundary tone does not convey sentences' pragmatic meanings, we should not use this notion to categorized question or statement intonations. Since we adopt the term "boundary tone" from Pierrehumbert (1980)'s work on phonological aspects of English intonation, we need to make it clear whether the notion of "boundary tone" should be used to represent intonation's pragmatic meanings at all.

Figure3 shows Pierrehumbert (1980)'s finite state grammar for generating intonational tunes in English.

figure3: finite grammar for generating intonation (Pierrehumbert, 1980: p29)



It displays the internal makeup of the melody for an intonational phrase. In Pierrehumbert's definition, "boundary tone" is a high tone or low tone target associated to either beginning or ending edges of an intonational phrase. In her original work, the boundary tone unit has two functions. Firstly, it signals the pause position or boundary of an intonational phrase. The intonation phrase boundary (marked by "%") occurs "where there is a non hesitation pause", or the last syllable of a phrase is lengthened. In most of cases, [intonational phrases] corresponds to syntactic structures. The high tone or low tone target appeared in this position are marked by H% or L%. Secondly, boundary tone helps discourse interpretation. It "contributes information about the [intonational phrase] as a whole, it conveys information about whether the current phrase is to be interpreted with particular respect to a succeeding phrase or not". A high boundary tone means "forward-looking", "connects to following discourse" or "needs further interpretation with respect to a succeeding phrase". A low boundary tone means "all the relevant information to an utterance is already in the previous discourse" (Pierrehumbert & Hirschberg 1990). All polar questions are ended in a high boundary tone because they need further interpretation with respect to the succeeding answers.

Although the "boundary tone" notion has been adopted in Mandarin intonation discussions, its' definition seems to be different from Pierrehumbert's original descriptions. Boundary tones' functions are changed in Mandarin intonation analysis. In Lin (2006)'s work, "boundary tone" has the exclusive function of distinguishing question from statement, so a boundary high tone is the signal for question intonation. However, original "boundary tone" only signals phrasing positions between intonational phrases.

Polar question is just one kind of intonational phrases ending with high boundary tone whose discourse interpretation requires forward reference to the following context. Shi F. (2010)'s "boundary tone" is more similar to Pierrehumbert's original usage. Following Pierrehumbert's definition, Shi F. thinks it is better to study the phonetic features of the prosodic boundary and the features of the semantic focus in a separate way when describing Mandarin intonation. By his definition, boundary tones do not indicate sentence moods. Boundary tones tell only boundaries between prosodic words. In his example "zhang1zhong1bin1/xing1qi1tian1/xiu1shou1yin1ji1/", boundary tones realized in the final syllables of each prosodic words, which are bin1, tian1 and ji1. However, Shi F.'s usage of boundary tone also has a little difference from Pierrehumbert's. He gave up boundary tone's function in discourse interpretation, viewing it as a pure break without conveying any semantic meanings. Boundary tones are not limited to the high (H%) or low(L%) tone targets in Shi. F.'s system.

Based on our experiment's results and phonological point of view, we suggest retaining boundary tones in Mandarin intonation representation. Our interpretation of the boundary tone unit strictly follows Pierrehumbert (1980)'s definition. Boundary tone is not a purely sentence mood's indicator or a simply break in our interpretations. Therefore, when we transcribe a syntactically un-marked polar question as "H% [intonational phrase] H%", the boundary high tones (H%) represent the structural boundaries of this question phrase. They also indicate that this question phrase needs further information from the following discourse.

4. Conclusions

The experiment shows that in syntactically un-marked polar questions, both global pitch register rising and significant localized pitch rising at both edges of the intonational phrase contribute to cue questions.

Un-marked yes-no questions which contains only one intonational phrase is represented as "H% [intonational phrase] H%". The global pitch rising and localized extra pitch swing are viewed as the phonetic realizations of these two boundary high tones.

Our interpretation of the boundary high tone (H%) follows Pierrehumbert(1980)'s original definition. Structurally, the H% signals intonational phrase boundary. Functionally, the high tone means the question needs further information from the following discourse, which is, the answer.

REFERENCE

- CHAO, Y. R. 1929. BeiPing YuDiao YanJiu (Study of Beijing Mandarin intonation). In Wu, Zongji and Zhao, Xinna (Eds.), ZhaoYuanRen YuYanXue LunWen Ji(Proceedings of Chao Yuen Ren's linguistic papers) (2002 ed.). Beijing: ShangWu YinShu Guan (The Commerical Press).
- CHAO, Y. R. 1933. ZhongGuo ZiDiao Gen YuDiao (Tone and intonation in Mandarin Chinese). GuoLi ZhongYang YanJiuYuan LiShi YuYan YanJiuSuo JiKan (Journal of the institute of history and philosophy, Academia Sinica) 4.2:121-135.
- CHAO, Y. R. 1968. A grammar of Spoken Chinese. Berkeley: University of California Press.
- HE, YANG, and JING, SONG. 1992. Beijing YuDiao de ShiYan TanSuo (Experimental exploration on Beijing Mandarin intonation). YuYan JiaoXue Yu YanJiu (Language Teaching and Research)2: 71-96.
- HU, MINGYANG. 1987. GuanYu BeijingHua de YuDiao Wenti (On the intonation of Beijing Mandarin). In BeijingHua ChuTan (Exploring Beijing Mandarin). Beijing: Commercial Press.
- HO, AICHEN T. 1976. The acoustic variation of Mandarin tones. *Phonetica* 33:353-367.
- HO, AICHEN T. 1977. Intonation variation in a Mandarin sentence for three expressions: interrogative, exclamatory and declarative. *Phonetica* 34:446-457.
- HOWIE, JOHN, M.. 1976. Acoustical studies of Mandarin vowels and tones. New York: Cambridge University Press.
- LADD, D.R.. 1981. On Intonational Universals. In Myers, T. et al. (eds). *The Cognitive Representation of Speech*. 389-397. Amsterdam: North Holland Publishing.
- LEE, OK JOO. 2005. The prosody of questions in Mandarin. Doctoral dissertation. Ohio State University, Columbus, OH.
- LIANG, J and VAN HEUVEN. 2007. Chinese tone and intonation perceived by L1 and L2 listeners. In Gussenhoven, C. and Riad, T. (eds.), *Tones and Tunes*. 27-61. Berlin, New York: Mouton de Gruyter.
- LIN, MAOCAN. 2004. On production and perception of boundary tone in Chinese intonation. Proceedings of the international symposium on tonal aspects of languages: with emphasis on tone languages, pp. 125–130. Beijing.
- LIN, MAOCAN. 2006. YiWen he ChenShu YuQi de BianJieDiao (Interrogative mood and boundary tone in Chinese). Report of phonetic research. Beijing: Institute of Linguistics.
- OHALA, J. J.. 1983. Cross-language use of pitch: An ethological view. *Phonetica*40: 1-18.
- OHALA, J. J.. 1984. An ethological perspective on common cross-language utilisation of

- F0 of Voice. *Phonetica*41:1-16.
- PENG, S., MARJORIE K.M.CHAN, CHIU-YU TSENG, TSAN HUANG, OK JOO LEE, and MARY BECKMAN. 2005. Toward a Pan-Mandarin system for prosodic transcription. In *Prosodic typology: The phonology of intonation and phrasing*, ed. Sun-Ah Jun, 230-270. Oxford: Oxford University Press.
- PIERREHUMBERT, J. 1980. The phonology and phonetics of English intonation. Doctoral dissertation. MIT, Cambridge, MA.
- PIERREHUMBERT, JANET, and MARY BECKMAN. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, ed. Jerry Morgan Philip Cohen and Martha Pollack, 271-311. Cambridge: Massachusetts Institute of Technology.
- SHEN, J. 1985. Beijing yuyin shiyan lu[Pitch range of tone and intonation in Beijing dialect]. In *Working papers in experimental phonetics*, ed. Lin,T. and Wang,L., 73-130. Beijing: Beijing University Press.
- SHEN,XIAO-NAN SUSAN. 1990. The prosody of Mandarin Chinese. Berkeley: University of California Press.
- SHI, PEIWEN. 1980. SiZhong JuZi De YuDiao BianHua (Intonation variations in four types of Mandarin sentences). *Yu Yan Jiao Xue Yu Yan Jiu (Language Teaching and Research)*2: 71-81.
- SHI,FENG. 2010. Boundary Tone and Focus Tone--on the hierarchy of Chinese intonation system. Public talk in the Chinese University of Hong Kong.
- YUAN JIAHONG, CHIH-LIN SHIH, and GREG P. KOCHANSKI. 2002. Comparison of declarative and interrogative intonation in Chinese. *Proceedings of the Speech prosody 2002 conference*, ed. Bel, B. and Marlien I. , pp.711-714. Aix-en-Provence: Laboratoire Parole et Language.
- YUAN JIAHONG. 2004. Intonation in Mandarin Chinese: Acoustic, Perception, and Computational Modeling. Doctoral dissertation. Cornell University, Ithaca.
- YUAN, JIAHONG. 2006. Mechanisms of Question Intonation in Mandarin. In *Chinese Spoken Language Processing*, ed. Huo., Q., et al., 19-30. Berlin:Springer.
- TSENG, CHIU-YU, 1981. An acoustic phonetic study on tones in Mandarin Chinese. Doctoral dissertation. Brown University, Providence, RI.
- Wu, Yanhong, Tao, Jianhua and Lu, Jilun. 2006. HanYu YiWen YuDiao De YunLv FenXi (Analysis of Mandarin interrogative intonations). In *Proceedings of 7th Chinese phonetic conference*. Beijing.

Perception of Mandarin fricatives by native speakers of Taiwan Mandarin and Taiwanese

Ya-ting Shih¹ Eunjong Kong²
The Ohio State University¹, University of Wisconsin-Madison²

This perception study examines bilingual Taiwan Mandarin (Guoyu) and Taiwanese listeners' perception of Mandarin fricatives. Two female talkers and two groups of listeners, one from Taiwan and the other from China, were recruited. The participants listened to fricative pairs such as /sa/, /fa/ and then rated the similarity. The results show that bilingual listeners have difficulty in distinguishing the contrast between Mandarin /s/ and /ʂ/ no matter which talker they listened to. They also perceive /x/ and /f/ to be similar when listening to the stimuli produced by a bilingual talker. The results show that the lack of the retroflex fricative /ʂ/ and labiodental /f/ in Taiwanese influence how these bilingual listeners perceive Mandarin fricatives.¹

1. Introduction

Fricatives sound like a stream of noise because fricatives are produced by air being forced through a narrow gap in the oral cavity (Ladefoged 2005) and this is where the noise comes from. According to Li (2008:25), the place of articulation of fricatives means "the narrowest constriction made by the tongue toward the ceiling of the vocal tract in the mid-sagittal plane. Take English for example; the /s/ in soup is an alveolar fricative which means that the narrowest constriction for this fricative is made in the alveolar ridge. Though, fricatives sound like a stream of noise, depending on the place of articulation, the frequency of the noise can help listeners distinguish which fricative they hear: the shorter the front cavity in front of the constriction, the higher the frequency of the noise. To distinguish fricatives with different places of articulation, frequency of fricative noise and formant movement of the adjacent vowels are often used.

Several previous studies examine cross linguistics perception show that listeners have more difficulty in discriminating fricatives that are not phonemic in their native languages. For example, Johnson and Babel (2010) explored whether listeners' fricative

¹ The authors would like to thank Dr. Mary Beckman, Dr. Marjorie Chan, Dr. Cynthia Clopper, Dr. Pusan Wong and Seth Wiener for their helpful suggestions and comments.

inventory will influence their perception of fricatives of non-native language. They recruited English and Dutch speakers and tested their perception of English fricatives. English has more voiceless fricatives than Dutch as shown in the table 1. The results show that Dutch speakers rated English /s/-/θ/, /s/-/ʃ/, /θ/-/ʃ/ to be more similar than English speakers. The lack of a dental and post-alveolar fricative in Dutch made Dutch listeners perceive the English fricatives /θ/ and /ʃ/, to be similar to adjacent fricatives such as /s/.

	Labiodental	Dental	Alveolar	Post-alveolar	Velar	Glottal
Dutch	f		s		x	h
English	f	θ	s	ʃ		h

Table 1. Fricatives of Dutch and English in IPA

Similarly, Aoyama et al. (2008) examined Japanese speakers' perception of English fricatives and the author focused only on English /θ/ which is not phonemic in Japanese. The result also shows that Japanese speakers have difficulty in discriminating /s/ and /θ/. The authors concluded that listeners are influenced by their native language inventory when perceiving fricatives of other languages. When there is a lack of fricatives in certain places of articulation in a listeners' native language, they perceive the fricative they do not have to be similar to the neighboring fricative in the target language.

	Labiodental		interdental		Alveolar		Post-alveolar		Palatal	Glottal
English	f	v	θ	ð	s	z	ʃ	ʒ		h
Japanese					s	z			ç	h

Table 2. Fricatives of Japanese and English in IPA

Lastly, Tsao et al. (2006) explored Mandarin and English speakers' perception of Mandarin fricative-affricates [ç]-[tç], [ç]-[tç^h] contrast. The authors indicated though

English speakers have contrast in the same place of articulation, aspiration is not used in contrasting fricatives. Therefore, they reported that English speakers performed poorly in discriminating the Mandarin fricative-affricate contrast.

Several production studies which focus on Taiwanese show that Taiwanese speakers' production of Mandarin fricatives are influenced by Taiwanese. The first study that examined Taiwanese speakers' production of Mandarin fricatives is Peng (1993). Peng recruited 10 Taiwanese people who learned Mandarin as a second language (L2) and differed in their Mandarin proficiency. She found that her participants tend to use Taiwanese /h/ for Mandarin /x/. Furthermore, due to the lack of a labial fricative in Taiwanese, these participants produced the labialized /h/ for the labio-dental fricative /f/. Similarly, Lin (2008) and Yang (2008) also researched Taiwanese speakers' production of Mandarin /f/. Though these two studies examined the phenomenon from a sociolinguistic perspective, they also found that Taiwanese speakers tend to produce Mandarin /f/ as [h^w]. Although /f/ and /h/ are not adjacent fricatives in terms of place of articulation, they are both non-sibilants. It might be the reason why Taiwanese speakers use labialized /h/ to mimic the labial feature of /f/.

	Labiodental	Dental	Retroflex	Palatal	Velar	Glottal
Mandarin	f	s	ʂ	ç	x	
Taiwanese		s				h

Table 3. Fricatives of Mandarin and Taiwanese in IPA

2. Research questions and predictions

The above cross-linguistic studies show that the lack of certain fricatives in one's native language will influence listeners' perception. However, there is relatively less research examining bilingual speakers' perception of one of their native languages. The study aims to answer the following research questions:

- (1) Which Mandarin fricative contrasts are more difficult for bilingual listeners of Taiwan Mandarin/Guoyu and Taiwanese to discriminate?
- (2) Do talkers' dialectal differences influence bilingual listeners' perception of Mandarin fricatives?

The predictions based on previous literature are (1) Bilinguals will perceive [f] to be similar to [x] because there is no labiodental fricative in Taiwanese and these two non-

sibilants will be perceived to be more similar than other sibilants. (2) Bilinguals will perceive [s] to be similar to [ʂ] because of the lack of a retroflex fricative in Taiwanese.

3. Methodology

3.1 Participants

Two talkers and two groups of listeners participated in the study. One talker was a bilingual Taiwanese and Taiwan Mandarin (Guoyu) speaker from southern Taiwan and the other talker was a monolingual Mandarin (Putonghua) speaker from northern China. The bilingual listeners were all from southern Taiwan ($N=22$, 7 male and 15 female, mean age=35). The other group of listeners ($N=20$, 5 male and 15 female, mean age=30) were all from China but recruited in the States. They are graduate students in a major middle west university. The monolingual listeners serve as a comparison group in order to compare their result with bilingual listeners.

3.2. Stimuli

The stimuli are five Mandarin fricatives combined with the low front vowel /a/ in tone one. They are /fa/, /sa/, /ʂa/, /ɕia/ and /xa/ (The stimuli are presented in IPA). For historic reasons, /ɕ/ cannot be combined with other vowels without /i/. Therefore, this stimuli is different from other four. Each talker produced a fricative plus /a/ 10 times. Then the stimuli were paired by randomly selecting from the 10 repetitions. The stimuli pairs are pairs such as /fa/-/fa/ or /sa/-/xa/. There are a total of 15 fricative pairs: 5 of them as “same” pairs and 10 of them as “different” pairs. The stimuli were blocked by a talker and each block contains 125 stimuli pairs.

3.3 Procedures

The experiment was run on Praat (a phonetic analysis software). All listeners listened to the block produced by the Putonghua talker first and then the Guoyu talker. The listeners used a five-point Likert scale : 1(the same), 2(very similar), 3(similar), 4(different), 5(very different) to rate the stimuli pair. Every participant listened to the stimuli pair only once and their reaction time was not recorded.

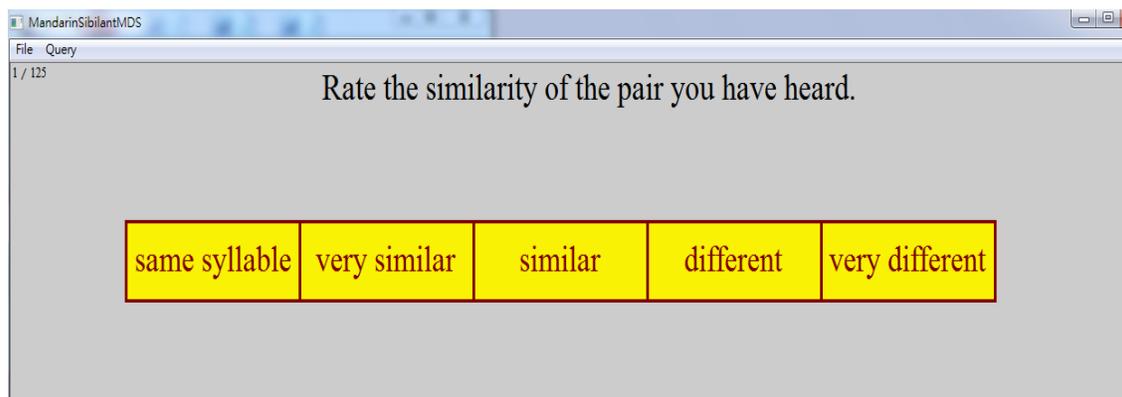


Figure 1. Snapshot of the program

3.4 Data analysis

The data was analyzed by first calculating all listeners' mean rating of the stimuli pairs. Then the data was analyzed by repeated measure of ANOVA. The three independent variables were (1) talker (bilingual of Guoyu and Taiwanese/monolingual of Putonghua), (2) stimuli pair (15 pairs) and (3) language background of the listeners (bilingual/monolingual). One-dimensional plots were also plotted to show the perceptual similarity of the listeners.

4. Results

Figure 2 shows the general patterns of all the results calculated by the mean rating of all participants. There are four lines in Figure 2. Each line represents the result of the two groups listened to the two talkers. The x-axis is the stimuli pairs. The first five pairs are identical pairs and the rest of them are different pairs. The y-axis is the mean rating of the stimuli pair ranging from 1(the same) to 5 (very different). The blue solid line represents the data of bilingual subjects listening to the Putonghua talker's production. The red solid line shows monolingual subjects listening to the Putonghua talker. The light blue dashed line represents the results of bilingual subjects listened to bilingual talker and lastly, the light red dashed line indicates the data of monolingual listeners' response to bilingual talker. The general patterns are pretty similar for both listener groups.

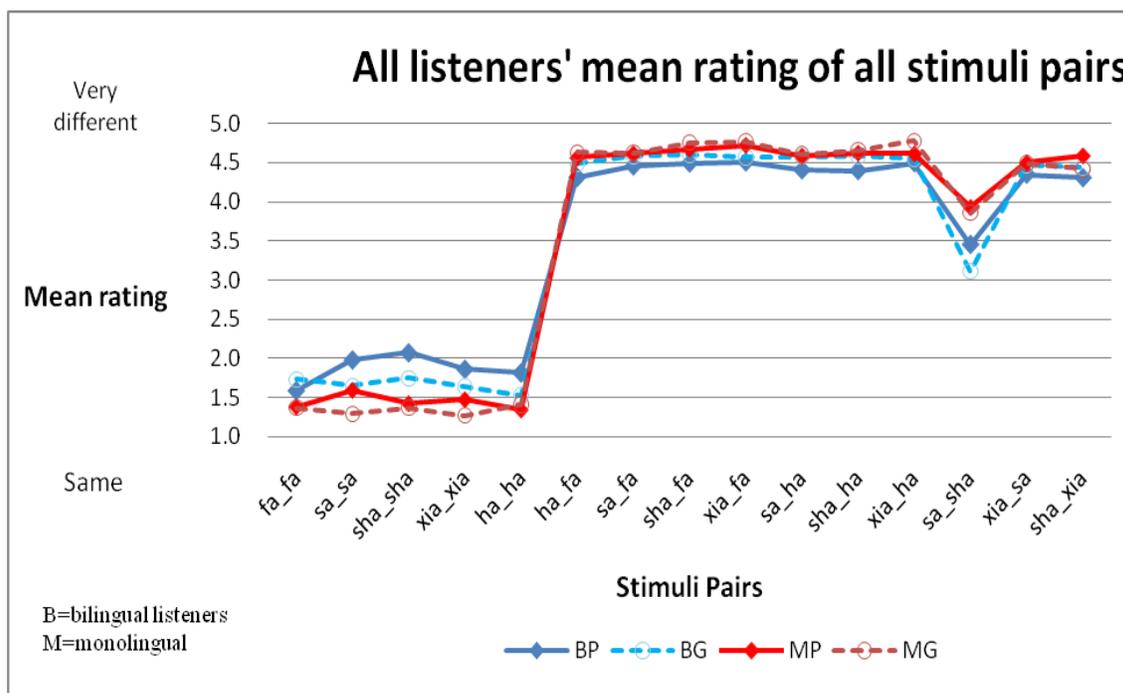


Figure 2. The mean rating of all listeners' response

The results were analyzed in a repeated measures analysis of variance. The between-listeners variable is language background (bilingual of Guoyu and Taiwanese and monolingual of Putonghua) and the within-listeners variables are (1) talker (bilingual of Guoyu and Taiwanese, monolingual of Putonghua) and (2) stimuli pairs (15 pairs). The main effect of stimuli pair was significant ($F [14,560]=507.71, p=.00$). The stimuli pair by language background interaction was also significant ($F[14,560]=5.80, p=.00$). The talker by stimuli pair interaction was also significant ($F[14 ,560]=4.68, p=.00$). Lastly, the three way interaction was also significant ($F[14 , 560]=1.92, p=.02$). In order to know whether the variation in the first five "same" pairs and the dipping in the sa-sha (the stimuli is represented by Pinyin) pair is significantly different, the researchers separated the pairs into two parts - five same pairs and ten different pairs - and ran a paired T test. The paired T test for the five same pairs showed that although there is variation among the five pairs, they are not statistically different from each other ($p>.05$). On the other hand, for the ten different pairs, sa-sha pair was the one which was significantly different from all other pairs ($p<.003$).

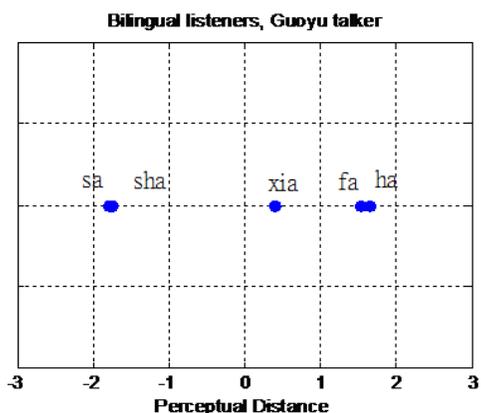


Figure 3. Bilingual listeners, Guoyu talker. (The stimuli pairs are presented in Pinyin).

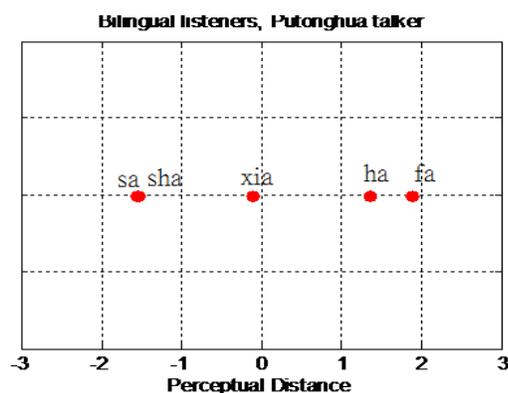


Figure 4. Bilingual listeners, Putonghua talker.

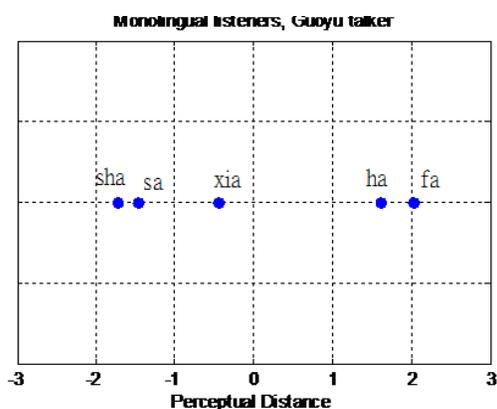
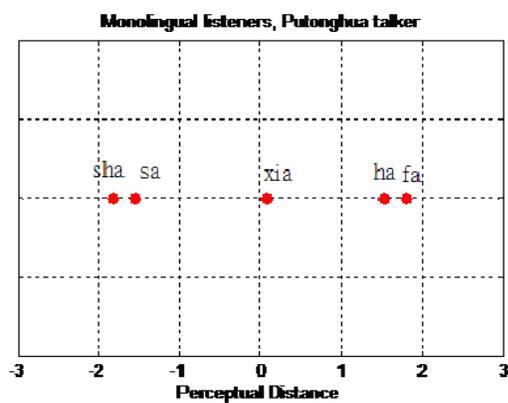


Figure 5. Monolingual listener, Guoyu talker. Figure 6. Monolingual listeners, Putonghua talker. (The stimuli pairs are presented in Pinyin).



The 4 one-dimensional plots show the perceptual space of the two listener groups. Although there are numbers on the x-axis, this number does not mean the exact distance of the sounds. It only shows the relative distance among the target fricatives. The plots with the blue dots are the results of bilingual listeners while and red dots show the results of the monolingual listeners. The stimuli pairs such as “sha” and “sa” in the plot are presented by Pinyin. Each plot has 5 dots which represent each fricative and vowel combination. Figure 3 shows the result of bilingual participants listening to the bilingual talker. These bilingual listeners perceived "sa" to be very similar to "sha". In other words, they perceived /s/ to be similar to /ʃ/ because the two dots overlapped. Interestingly, they also perceived the "ha" and "fa" produced by the bilingual talker similar to each other, which means they perceived /x/ to be close to /f/. However, when the bilingual subjects

listened to the monolingual speaker, they still perceived "sa" to be close to "sha" but not the "fa" and "ha" pair. The production of sa-sha contrast is clearer when produced by a Putonghua speaker. However, bilingual of Taiwanese and Guoyu still perceived these two pairs to be very similar to each other. On the other hand, the monolingual listeners can distinguish the five fricatives very well. Unlike bilingual listeners, they can separate the "sa" "sha" pair and "fa" "ha" pair very well for both talkers.

5. Discussion

The general pattern for the two groups of listeners is very similar to each other. In Figure 2, although bilingual listeners seem less certain in rating identical pairs than monolingual listeners, the statistical analysis shows that this variation is not statistically different ($p > .05$). Moreover, in Figure 2, listeners' rating of the sa-sha pair also drops which is different from all other pairs. A paired T test was conducted to test if this dipping is truly different from all other pairs. The statistic test indicated that the rating of the sa-sha pair is significant different from all other pairs. The rating of the ha-fa pair also differs from many pairs but it does not differ from all pairs like the sa-sha contrast. However, on the one-dimensional plot, it shows that the fa-ha pair is rated similarly only by bilingual listeners when they listened to the bilingual talker. It may be the reason why on the general pattern it did not show a clear dipping like the sa-sha pair.

To answer the research questions, the most difficult fricative pair for a bilingual of Guoyu and Taiwanese is the sa-sha pair. The result is further confirmed by the one dimensional plot because the perceptual distance is very close for /s/ and /ʃ/ and it is true no matter which talker they listened to. One possible explanation for the results may be the influence from Taiwanese. Due to the lack of retroflex fricatives in Taiwanese, these speakers had more difficulty in discriminating the two fricatives. The result of this experiment is similar to Johnson and Babel (2010)'s study. In their study, their Dutch listeners perceiving English fricatives not in their native language to be closer to the neighboring sounds. On the other hand, the bilingual listeners also have difficulty in discriminating the ha-fa pair which is the contrast between /x/ and /f/. However, the bilingual listeners only have difficulty distinguishing when the stimuli were produced by the bilingual talker. This is probably because the bilingual talker has a less distinctive fricative category when compared with the monolingual Putonghua talker as shown in figure 7 and 8. The result of this perception study also provides some evidence for Peng (1993)'s production study. The bilingual listeners of Guoyu and Taiwanese perceive /x/ to be close to /f/ due to the lack of a labiodental fricative in Taiwanese. Therefore, they use labialized /x/ to mimic the labial feature of /f/.

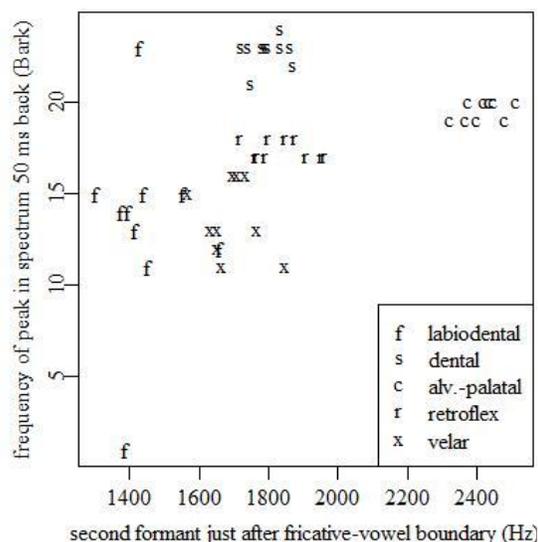


Figure 7. Plot of Guoyu talker's second formant after the fricative-vowel boundary.

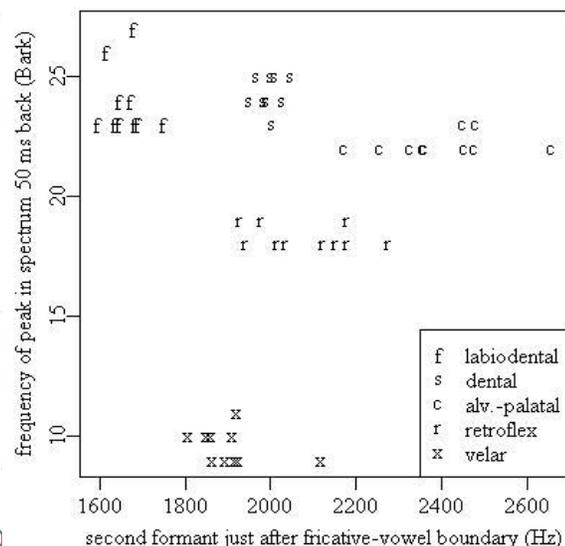


Figure 8. Plot of Putonghua talker's second formant after the fricative-vowel boundary.

6. Conclusion

The current experiment aims to investigate bilingual Guoyu and Taiwanese speakers' perception of one of their native language, Guoyu, and see if there is any influence from Taiwanese. The research questions are (1) which Mandarin fricative contrast is most difficult for the Guoyu and Taiwanese speakers and (2) whether talker's dialectal difference will influence listeners' rating of the Mandarin fricatives. Two groups of listeners were recruited. One group of Mainland Chinese/Putonghua speakers were recruited to serve as control group. The results show that the most difficult Mandarin fricative pair is the sa-sha pair. The monolingual Putonghua listeners can distinguish the contrast well but not the bilingual Guoyu and Taiwanese listeners. On the other hand, for the ha-fa contrast, bilingual listeners have difficulty in discriminating the contrast produced by the bilingual talker only. The findings match the results of previous perception studies that listeners have more difficulty in discriminating fricatives that are not phonemic in their native languages. They will also perceive those “non-phonemic” fricatives to be similar to the neighboring fricatives. Future research can examine these bilingual talkers' fricative space to see if or how they make a distinction between the fricatives and if there are any differences when they produce Guoyu and Taiwanese fricatives.

7. Limitations

There are some limitations of this study. First, the bilingual data was collected in southern Taiwan. Although the participants are more homogeneous, it may not reflect the larger population in Taiwan. On the other hand, the monolingual data was collected in the States. The subjects are less homogeneous and they also speak English. Lastly, the blocks were not counter-balanced when presented to the listeners. It may be better if they are counter-balanced to eliminate any potential order effects.

REFERENCES

- AOYAMA, KATSURA; GUION, SUSAN; FLEGE, JAMES; YAMADA, TSUNENO and AKAHANE-YAMADA, REIKO. 2008. The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. *International Journal of Applied Linguistics* 46. 61-90.
- BOERSMA, PAUL and WEENINK, DAVID. 2011. Praat: doing phonetics by computer [Computer program]. Version 5.2.25, retrieved 11 May 2011. <http://www.praat.org/>
- JOHNSON, KEITH and BABEL, MOLLY. 2010. On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* 38. 127-136.
- LADEFOGED, PETER. 2005. Vowels and consonants: An introduction to the sounds of languages. 2nd Edition. Malden, MA: Blackwell Pub.
- LI, FANGFANG. 2008. The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese. Columbus, OH: The Ohio State University dissertation.
- LIN, CHENG. 2008. A study of [f] as an allophone of /x/ in Taiwan mandarin. Tai-chung, PENG, SHU-HUI. 1993. Cross-language influence on the production of Mandarin /h/ by native speakers of Taiwanese Amoy. *Phonetica* 50. 245-260.
- TSAO, FENG-MING; LIU, HUEI-MEI and KUHL, PATRICIA. 2006. Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America* 120. 2285-2294.
- YANG, CHEN-YI. 2008. Sociolinguistic variation of mandarin Chinese /f/ and /hw/ in Taiwan. Tai-chung, Taiwan: Providence University thesis.

Investigating Taiwan Southern Min subsyllabic structure using maximum entropy models and wordlikeness judgments

Ying-Shing Li

Institute of Linguistics, Academia Sinica

To resolve the issue of Taiwan Southern Min syllabic structure, we first investigated the probabilistic co-occurrence of segments using maximum entropy models to simulate phonotactic learning processes. The algorithm constructed the constraint-based grammars that fitted the gradient phonotactic patterns of the input by yielding the numeric weights of the learned constraints. In addition to the baseline defined by the feature matrices, the constraints were augmented with the subsyllabic tier to express the hierarchical relationship of segments. For a comparison with the modeling consequences under different subsyllabic hypotheses, we further conducted a wordlikeness judgment experiment on nonsense syllables on a continuous scale. It was revealed that the body-coda model distinguished systematic gaps, accidental gaps, and attested syllables over a continuum of violation scores and obtained a significant correlation between violation scores and wordlikeness judgments. This study thus has provided evidence supporting Taiwan Southern Min as having body-coda structure that not only enhanced the phonotactic learning but also confirmed native speakers' phonotactic intuitions.

1. Introduction

Subsyllabic structure across languages has been proposed as the flat model, the onset-rhyme model, or the body-coda model (e.g., Clements & Keyser, 1983; Davis, 1989; Fudge, 1969, 1987; Godsmith, 1990; Hayes, 1989; Hyman, 1985; Kahn, 1976; Kiparsky, 1979; Levin, 1985; McCarthy & Prince, 1990; Selkirk, 1982; Treiman, 1988). As shown in Figure 1 (a), the flat model claims that segmental nodes are concatenated sequentially, without the branching structure within a syllable. Figure 1 (b) shows the onset-rhyme model that posits an intermediate rhyme node under which a vowel and a postvocalic consonant are associated closely. Figure 1 (c) exhibits the body-coda model in which a prevocalic consonant and a vowel comprise an intermediate body level. As in Figure 1 (d), lastly, within a syllabic structure, a prevocalic consonant and a postvocalic consonant constitute a discontinuous subpart before joining with a vowel.

LI: MODELING SUBSYLLABIC STRUCTURE

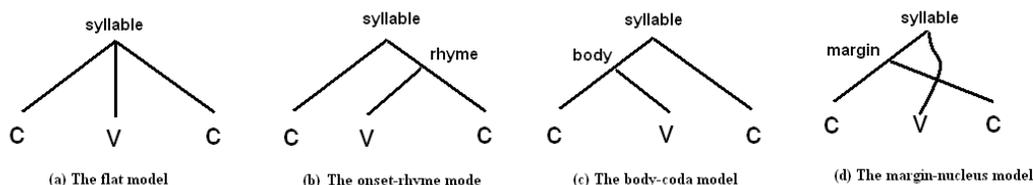


Figure 1. Four types of subsyllabic structure

Subsyllabic structure across languages has been presented along multiple threads of evidence: (1) stress or tone assignment based on syllable weight of the rhyme (e.g., Blevins, 1995; van der Hulst & Ritter, 1999) (2) co-occurrence restrictions (e.g., Fudge, 1987; Treiman, 1988) (3) spontaneous or experimentally elicited speech errors (e.g., Fowler, 1987; Mackay, 1972; Shattuck-Hufnagel, 1983) (4) naturally occurring or experimentally contrived word games (e.g., Fudge, 1987; Hockett, 1973; Pierrehumbert & Nair, 1995; Treiman, 1983) (5) other experimental techniques such as word-blending, unit substitution, sound similarity judgment, concept formation, unit reduplication, and nonce-word recall (e.g., Derwing, 2007; Wang, 1995; Yoon & Derwing, 2001).

In the pursuit of syllabic constituency in languages, the rhyme-coda model has been most widely accepted across languages (e.g., Blevins, 1995; der Hulst & Ritter, 1999; Fudge, 1969, 1987; Goldsmith, 1990; Kiparsky, 1979; Levins, 1985; Selkirk, 1982; Treiman, 1988). Crucially, a rhyme as the constituent that carries syllable weight can affect the distribution of primary stresses or contour tones in a language, leaving the prevocalic consonant not involving in the prosodic assignment. Citing evidence from English distributional constraints, speech errors, and word games, Fudge (1987) further contended that rhyme as a universal phonological unit.

Alternatively, in the moraic theory a mora is a hypothetical unit governing the metric weight or length, with a light syllable having a single mora and a heavy syllable having two morae (e.g., Hyman, 1985). Within a syllable, the prevocalic consonant and the nucleus converge under a moraic node, whereas the postvocalic consonant is placed under the other moraic node. This moraic theory expects a similar subsyllabic structure as the body-coda model in Figure 1 (c).

In the other version of the moraic theory (e.g., Hayes, 1989), the prevocalic consonant directly links with the syllabic node, yet the nucleus and the postvocalic consonant associated with two separate moraic nodes under a syllabic node. This moraic theory espoused the property of no intermediate branching node between the syllable and the segments, as show in Fig. 1 (a). Similar proposals were also found in the ternary branching model (e.g., Davis, 1989) and the flat model (Kahn, 1976). Unlike the other three models, however, the margin-nucleus model lacks theoretic modeling and direct evidence from languages (but cf. Fudge, 1973, 1987), even though it is logically

plausible.

Following this research line, the study pursues the question as to Taiwan Southern Min subsyllabic structure. Which subsyllabic structure does Taiwan Southern Min syllables contain? Taiwan Southern Min lacks evidence of the rhyme as the locus of syllable weight (Cheng & Tseng, 1997). Taiwan Southern Min syllables have long been proposed as having the onset-rhyme construction (Bao, 1990; Cheng & Cheng, 1977; Chung, 1996); for instance, Bao adopted evidence of *Fanquie*, originally used to specify the pronunciation of a novel character in terms of two known ones in traditional Chinese philological literature, to conclude onset-rhyme structure in Taiwan Southern Min. By contrast, the contestation that Taiwan Southern Min syllables are moraicly represented has been provided by other linguists (e.g., Chung, 1999; Cheng, 2002). Adopting experimental evidence (e.g., the body-sharing novel compounds were more easily memorized than the rhyme-sharing novel compounds), Derwing and his colleague (Derwing, 2007; Wang & Derwing, 1993) supported the body-coda construction as well.

Given the above contradictory evidence, the question regarding Taiwan Southern Min subsyllabic structure has not been fully settled down. In this article, we attempted to resolve the issue by computing the strength of coherence between adjacent segments within a syllable. Particularly, we attempted to simulate Taiwan Southern Min phonotactic learning based on the four subsyllabic hypotheses using maximum entropy models of phonotactic learning. For a reexamination of the modeling results, we also conducted a wordlikeness judgment experiment following the maximum entropy models of phonotactic learning. The working hypothesis adopted here is that proper subsyllabic structure not only fosters the phonotactic learning (i.e., making learning viable and efficient) but also corresponds with native speakers' phonotactic intuitions.

This article will be organized as follows. In the second section will be to present further literature review of Taiwan Southern Min subsyllabic structure. The third section will be to delicately introduce maximum entropy models of phonotactic learning on four types of subsyllabic hypotheses in Taiwan Southern Min. The fourth section will be to describe a wordlikeness judgment experiment and examine both of modeling and experimental results. In the last section will arrive at the conclusion on Taiwan Southern Min subsyllabic structure.

2. Some other issues on Taiwan Southern Min syllable structure

Traditionally, Taiwan Southern Min, like most of other Chinese languages (e.g., Mandarin), has no more than four segments in a syllable; see Table 1 (a)-(l). The maximum syllable structure can be CGVG or CGVX (Cheng & Cheng, 1977; Chung, 1996). Within a syllable, the onset consonant (C) is called as an *initial* whereas the rest elements of a syllable constitute a *final*. Within the *final* are a prevocalic glide (G), a nucleus vowel (V), and a postvocalic glide (G) or a postvocalic obstruent (X) that all construct a *rhyme* as well. Notwithstanding arguably a prevocalic glide may form a consonant cluster or a coarticulatory component with a preceding consonant (Bao, 1990;

LI: MODELING SUBSYLLABIC STRUCTURE

Duanmu, 1990), this study would adopt the widely accepted view that a prevocalic glide belongs to part of a nucleus vowel. Although there are also debates as to the status of a postvocalic glide or obstruent (Chung, 1996; Lin, 1989), but this study would tentatively determine a postvocalic glide as subordinate to a nucleus vowel, but a postvocalic obstruent as independent from a nucleus vowel. In the maximum entropy modeling, this study would thus schematize Taiwan Southern Min syllable structure as C-V-C, in which C- is a pre-glide consonant, V is an integral part that comprises maximum three segments (i.e., a prenuclear glide, a nucleus vowel and a postnuclear glide), and -C is a postvocalic obstruent.

A number of syllable types attested in a contemporary Taiwan Southern Min dictionary (Tung, 2001) will bring about a few further considerations of syllable structure. The example (m) in Table 1 that comprises five segments occurs when the postnuclear consonant is a glottal stop. A glottal stop in Taiwan Southern Min has been argued as a segmental or tonal element (e.g., Chung, 1995; Hung, 1994; Li, 1989). Although a glottal stop viewed as a toneme holds a maximum syllable size of four segments, a glottal stop in the following modeling would rather be treated as a distinctive segment in order to technically distinguish such attested syllables as /ka⁵³/, /kaʔ⁵³/, and /kak⁵³/ in Taiwan Southern Min. In addition, the examples (n) and (l) exhibit the syllables containing the segment sequences that involve the nucleic nasal consonants [m] or [ŋ] and the other consonants. Given these, the current maximum entropy models would also be required to acquire syllable structure of maximum five segments as well as consonant clusters consequently.

Table 1. The attested syllable types in Taiwan Southern Min

	Syllable types	Examples	Glosses
(a)	V	i ⁵³	'chair'
(b)	VX	ap ⁵³	'box'
(c)	VG	au ⁵³	'press'
(d)	GV	io ⁵⁵	'waist'
(e)	GVG	uai ⁵⁵	'awry'
(f)	GVX	ian ⁵⁵	'smoke'
(g)	CV	li ³³	'profits'
(h)	CGV	sia ⁵³	'write'
(i)	CVX	kun ⁵⁵	'troops'
(j)	CVG	tai ¹³	'bury'
(k)	CGVG	ziau ²¹	'scratch'
(l)	CGVX	puan ¹³	'plate'
(m)	CGVGX	ŋiauʔ ⁵³	'squirm'
(n)	CX	sŋ ⁵⁵	'sour'
(l)	CXX	ts ^h ŋʔ ³¹	'blow the nose'

LI: MODELING SUBSYLLABIC STRUCTURE

According to Chung (1996), seven distributional constraints of syllable structure have been proposed in Taiwan Southern Min: (1) the N-Constraint requires diphthongs to have at least one high vowel, e.g., *eo, *ao; (2) the Dissimilatory Constraint prohibits [α back](...)[α back] in a di- or triphthong, e.g., *ie, *uo, *uei, *iou; (3) the Fall Constraint disallows a falling diphthong from preceding a coda, e.g., *aip/m, *ait/n, *aik/ŋ; (4) the Branching-R Constraint bans [+high][+high] in the VC-structure, e.g., *ik, *iŋ, *uk, *uŋ; (5) the Branching-N Constraint prevents a prevocalic *u* from co-occurring with a velar coda, *uak, *uaŋ, *uek, *ueŋ; (6) the Labial Constraint prohibits [+labial](...)[+labial] within the syllable unless the two labials are onset and nucleus, e.g., *um, *op, *iop, *uam; (7) the Nasal Constraint requires that a maximum of one nasal autosegment may occur in a syllable (given that C(G)(V)(G)- and -X are two domains for nasality percolation), e.g., *mãŋ, *bãŋ. The first two distributional constraints are carried out within the nucleus, the next four distributional constraints within the rhyme, and the seventh distributional constraint within the body. A raw probability of constraints within certain subsyllabic component probably cannot determine syllable structure of a language (Fudge, 1987). Instead, the interaction with the other constraints and the application scope over the lexicon of a particular constraint will play a promising role of evaluating its effectiveness in influencing syllable structure of a language.

3. The application of maximum entropy models in Taiwan Southern Min

Maximum entropy can be traced back to Biblical times and have been widely applied outside the linguistic domains (e.g., natural language processing (NLP) in Berger, Della Pietra, & Della Pietra, 1996). A specific maximum entropy model (Hays & Wilson, 2008) is a stochastic model used to simulate the behaviors of phonotactic learning processes. The learning aims at constructing the constraint-based grammar (Prince & Smolensky, 2004) that best fits the learning data. Technically speaking, the model is meant to calculate the maximum mathematical fitness of the random-sampling inputs in each learning cycle. Given a constraint-seeking algorithm, the last goal of modeling is to make the summation of the violation scores of the weighted constraints minimized for all the attested data.

Some principles approaching such a goal are inherently determined in the modeling: First, the most accurate constraints are prioritized for adopting as the learned phonotactic constraints. The accuracy is defined as the O/E value, that is, the number of violations of the constraint observed in the learning data (O), divided by the number of violations expected from the current grammar (E). Second, if there is a tie of accuracy, generality is the secondary consideration. Shorter constraints (fewer feature matrices) are favored over longer ones (more feature matrices). Simpler constraints (involving fewer feature specifications) are also favored over more complex ones (more feature specifications). Constraints involving more segments are preferred to those with few segments.

3.1 Learning data

The learning data as the input were the lexical entries from a contemporary Taiwan Southern Min dictionary (Tung, 2001). The lexical entries were adopted from a variety of spoken resources with a few written texts as a supplement. There were totally 6,893 lexical entries (i.e., monosyllabic morphemes or characters) listed in the dictionary. For assessing syllable type frequency, all of the corresponding morphemes or characters were summed up. Segments involved were encoded as the machine-readable symbols in terms of the SPE-style feature specifications. Maximum entropy models subsequently can yield the feature-defined constraints by internalizing the learning data.

3.2 Setting parameters

Maximum entropy models must proceed after the parameters were set in advance. These values were used to extract more effective constraints but block less useful constraints. The maximum O/E value for constraints was 0.3 and the accuracy schedule was at [.001, .01, .1]. The maximum number of constraints to discover was unlimited in that the model could acquire as many constraints as possible. The maximum gram size of constraints was 4 (defined by the number of feature matrices), based on the observation that the majority of Taiwan Southern Min syllables were no longer than four segments, and the model eventually converged as the maximal gram size was only three. In terms of the constraint definitions, a complementation operator \wedge was used to address the logical implication between a segment sequence, that is, as *if a particular segment has a feature matrix, then any preceding or following segment must have another feature matrix*. According to Hayes & Wilson (2008), the use of a complementation operator improved the performance of the modeling as well as the interpretation of the learned grammars.

Crucially, four types of projections were considered to distinguish the four hypotheses of Taiwan Southern Min syllable structure. A projection worked like inserting a boundary at the projected domain edges that enhanced the cohesiveness of segments within the projected domain but blocked segments outside the projected domain. Analogically, a projection worked like constructing hierarchical structure as well. The Flat model was the default baseline in which no internal branching structure within a syllable is posited. For the syllable /#pan#/, syllable tier was marked as (i.e., [\pm boundary]) to anchor the syllabic boundaries (i.e., #), corresponding with the segmental features in the segment tier. In addition to the syllable tier and segment tier, the Body model was projected with the intermediate tier in which an onset and a nucleus were more closely combined. The body tier was marked as subsyllabic boundaries (i.e., [\pm body]) corresponding with the segment sequence /#pan#/. The Rhyme model was used to demonstrate the intermediate branching node under which a nucleus and a coda converged. The rhyme tier was marked as the subsyllabic boundaries (i.e., [\pm rhyme]) to specify the segment sequence /#pan#/ in the segment tier. Finally, the margin tier was marked as the subsyllabic boundaries (i.e., [\pm margin]) corresponding with the segment sequence /#pan#/ as hypothesized in the Margin model. Contrasting a sequential order by default, the Margin model was used to address a novel assumption in which a prevocalic

consonant and a postvocalic consonant were grouped before joining a nucleus.

3.3 Evaluation

The testing data consisted of all the toneless combinations of onsets (p, p^h, b, m, t, t^h, l, n, k, k^h, g, ŋ, h, ts, ts^h, s, z), nuclei (a, ã, ai, ãĩ, au, ãũ, e, ẽ, i, ï, ia, ãã, iau, ããũ, io, io, ãõ, iu, ãũ, o, õ, ɔ, u, ua, ãã, uai, ããĩ, ue, ãẽ, ui, ãĩ), and codas (p, t, k, m, n, ŋ, ʔ) in Taiwan Southern Min. There were totally possible 4,608 syllable types. Of all the syllable types, 836 were attested in the dictionary.

The testing data was then input to the learned grammars to generate the violation score for each syllable. The violation score of a syllable x , denoted $h(x)$, is

$$h(x) = \sum_{i=1}^N w_i C_i(x),$$

where

w_i is the weight of the i th constraint,

$C_i(x)$ is the number of times that x violates the i th constraint, and

$\sum_{i=1}^N$ denotes summation over all constraints (C_1, C_2, \dots, C_N).

The transformation from the violation score to the value representing degree of wellformedness of a syllable can be conducted by calculating the maxent value of x , denoted $P^*(x)$.

$$P^*(x) = \exp(-h(x))$$

The violation score is negated, and e (the base of the natural logarithm) is raised to the result. The probability of x in the modeling is calculated by determining its share in the total maxent values of all possible syllables in Ω , a quantity designed as Z . In effect, the probability makes a lot of sense for determining an optimal constraint by comparing its contribution to the model with other competing constraints in the modeling.

$$P(x) = P^*(x) / Z$$

$$\text{where } Z = \sum_{y \in \Omega} P^*(y)$$

The evaluation can begin with whether any of the four models can predict categorical lexical status of Taiwan Southern Min. Attested syllables were supposed to obtain 0 violation score as composed to nonsense syllables. As exhibited in Table 2(a), the Body model and the Rhyme model were more capable of assigning attested syllables

LI: MODELING SUBSYLLABIC STRUCTURE

0 violation score¹ (i.e., nearly 100%) and assigned nonsense syllables non-zero violation scores, whereas the other two models were to a larger degree away from the accurate lexicality distinction. As for the other three indexes, attested syllables were expected to be significantly distinguishable from nonsense syllables. As shown in Table 2 (b)-(c), average violation scores of nonsense syllables were higher than average violation scores of attested syllables, average maxent values of attested syllables were higher than average maxent values of nonsense syllables, and average probability of attested syllables were higher than average probability of nonsense syllables. In terms of the three indexes, quantitatively, the difference size between attested syllables and nonsense syllables was larger in the Body and Rhyme models than in the other two models. The comparisons across the four models supports one thing: Only the Body and Rhyme models are able to learn Taiwan Southern Min phonotactic grammar (i.e., distinguishing attested and nonsense syllables). Neither the Flat model nor the Margin model can do phonotactic learning as well as the Body and Rhyme models.

Table 2. Results of the maximum entropy models

(a)

Accuracy rates	Body	Rhyme	Flat	Margin
Attested	99.88%	99.88%	59.14%	59.14%
Nonsense	78.54%	83.88%	89.44%	85.94%

(b)

Average violation scores	Body	Rhyme	Flat	Margin
Attested	0.007	0.004	1.133	1.060
Nonsense	6.072	7.582	3.463	2.525

(c)

Average maxent values	Body	Rhyme	Flat	Margin
Attested	0.999	0.999	0.640	0.633
Nonsense	0.219	0.163	0.161	0.210

(d)

Average probability	Body	Rhyme	Flat	Margin
Attested	0.06%	0.07%	0.06%	0.05%
Nonsense	0.01%	0.01%	0.01%	0.02%

One additional thought regarding the results is that wellformedness of certain nonsense syllables provided evidence for accidental gaps which were considered

¹ The only syllable that the Body and Rhyme models failed to learn is an onomatopoeic morpheme [ŋh] 'the sound of having a bowel movement'.

LI: MODELING SUBSYLLABIC STRUCTURE

well-formed but underrepresented in the lexicon. Conversely, a couple of lexical exceptions (e.g., [ŋh]) which were considered ill-formed but overrepresented in the lexicon were found although the size was relatively limited.

Before moving on, we should examine the learned grammars from the Body and Rhyme models. The constraints of both learned grammars are exhibited in Table 3 and Table 4, respectively. Frequencies of constraints between the Body and Rhyme models are almost equal and the learned constraints were almost comparable.

Table 3. The learned constraints in the Body model (24 constraints)

	Constraints	Tiers	Weights	Examples
1	*[^-voice][-syl]	Body	1.192	tɪ (*dɪ)
2	*[-syl][^+son,+body]	default	5.458	ta, tɪ (*tp)
3	*[+nasal][-nasal]	Body	2.438	*na, *ãu
4	*[^-nasal,+body][-nasal,+body]	default	3.778	ba, au (*ma, *ãu)
5	*[+approx,-nasal][^+syl,-nasal]	Body	5.880	au (*aũ)
6	*[+syl,-nasal][+cor]	default	4.714	at, an (*ãt, *ãn)
7	*[-high][^+high]	Body	5.144	ai, au (*ae, *ao)
8	*[+boundary][-body]	default	2.501	*[-p
9	*[+back][^-back]	Body	4.695	ui, ua (*uo)
10	*[-body][-boundary]	default	5.012	*-pt
11	*[-high,-low][-boundary]	Body	4.073	*ei, *ou, *ɔa
12	*[^-nasal,+body][+lab]	default	4.016	ap (*ãp)
13	*[-low,-round][-low,-back]	Body	4.816	*ei, *oi, *oe, *ie
14	*[+back][+lab]	default	4.725	*up, *ɔm
15	*[+cont,+spread][+lab]	Body	2.928	hm (*k ^h m)
16	*[+syl,+nasal][-syl,+son]	default	1.772	*ãl, *ãm
17	*[+voice][^+syl,-nasal]	Body	4.287	ba (*bã, *bt)
18	*[+syl,+nasal][+velar]	default	4.819	*ɔk, ãk
19	*[-boundary][-syl][-boundary]	Body	4.852	*[...C...]
20	*[-high,-low][+cor]	default	3.863	*et, *en, *ot, *on
21	*[+syl][-boundary][^+high]	Body	1.371	iau (*uae, *iuo)
22	*[-high,-low,-round][^-cont,+glottal]	default	4.689	eʔ, oʔ (*ok, *ek)
23	*[+syl][^+low][-boundary]	Body	1.670	iau (*iou, *ieu)
24	*[+syl][^-high][+son]	default	4.267	ian (*uin, *aun)

LI: MODELING SUBSYLLABIC STRUCTURE

Table 4. The learned constraints in the rhyme model (23 constraints)

	Constraints	Tiers	Weights	Examples
1	*[[^] +high][-high]	Rhyme	5.768	ia (*ea)
2	*[-syl][-syl,-nasal]	default	3.056	*tk (tɿ)
3	*[-approx][-boundary]	Rhyme	2.394	*t]
4	*[[^] -voice,-rime][-rime]	default	7.495	tɿ (*dɿ)
5	*[[^] +syl,-nasal][+syl,-nasal]	Rhyme	2.205	ai (*āi)
6	*[[^] +syl,-nasal][+cor]	default	4.723	at, an (*āt, *ān)
7	*[[^] +syl,+nasal][+syl,+nasal]	Rhyme	5.695	āi (*āi)
8	*[+nasal][+approx,-nasal]	default	5.460	*na, *āi
9	*[[^] -back][+back]	Rhyme	5.094	au, io (*ou, *uo)
10	*[+nasal,+rime][-syl,+son]	default	1.689	*ām, *āl
11	*[-high,-low][+approx]	Rhyme	3.843	*eu, *ei, *ou, *oa
12	*[+nasal][+lab]	default	2.499	*āp, *nm
13	*[-low,-round][-low,-back]	Rhyme	4.919	*ei, *oi, *oe, *ie
14	*[-cont][+lab]	default	3.624	*km (hm)
15	*[-high,-low,-round][[^] -cont,+glottal]	Rhyme	5.781	eʔ, oʔ (*ek, *ok)
16	*[+nasal,+rime][+velar]	default	4.859	*āk, āŋ
17	*[[^] -nasal,-back][+lab]	Rhyme	4.102	em, ep (*ēm, *ēp, *om, *op)
18	*[+voice][[^] +syl,-nasal]	default	4.867	ga (*gā, *gt)
19	*[[^] +high][-boundary][+son]	Rhyme	1.283	uan (*aun)
20	*[-boundary][-syl][[^] -cont,+glottal]	default	6.764	ts ^h ŋʔ (*ts ^h ŋk)
21	*[-nasal,-high,+round][[^] +velar]	Rhyme	4.769	ɔk, ɔŋ (*ɔt, *ɔn)
22	*[+rime][[^] -high][+son]	default	2.551	uai, uen (*uin, *iun)
23	*[-boundary][+high][[^] -cont,+glottal]	Rhyme	5.360	uiʔ (*uik)

In Table 3, the Body model learned twenty-four constraints totally. Provided the seven constraints (Chung 1996), we are able to compare the learned constraints against Chung's proposal: the N-Constraint is represented as constraint (7), the Dissimilatory Constraint as constraint (9, 13), the Fall Constraint as constraint (24), the Labial Constraint as constraint (14), and the Nasal Constraint as constraints (3, 4, 5, 6, 16). In Table 4, twenty-three constraints were learned in the Rhyme model. Analogously, the Rhyme model learned the N-Constraint represented as constraint (1), the Dissimilatory Constraint as constraint (9, 13), the Fall Constraint as constraint (19, 22), the Labial Constraint as constraint (17), and the Nasal Constraint as constraints (5, 7, 8, 10).

The nonoccurrence of the Branching-R Constraint and Branching-N Constraint was expected since the following syllable types (i.e., [tsuaŋ], [uaŋ], [siŋ], [p^hiŋ], [piŋ], [kiŋ], [hiŋ], [ts^hiŋ], [liŋ], [tsiŋ], [biŋ], [t^hiŋ], [sik], [pik], [lik], [kik], [hik], [gik], [tsik], [bik]) were already attested in the Taiwan Southern Min dictionary. Such attested syllables might decrease the effectiveness of the Branching-R Constraint and Branching-N Constraint, therefore causing unsystematic predictions of lexicality of the associated syllable types.

4. Testing learned grammars using wordlikeness judgments

As indicated in literature (e.g., Greenberg & Jenkins, 1964; Ohala & Ohala, 1986; Vitevitch, Luce, & Charles-Luce, & Kemmerer, 1997; Wang, 1998; Frisch, Large, & Pisoni, 2000; Treiman, Kessler, Knewasser, Tincoff, & Bowman, 2000; Bailey & Hahn, 2001; Hay, Pierrehumbert, & Beckman, 2003; Hammond, 2004; Myers & Tsay, 2004; Kirby & Yu, 2007), nonsense syllables of a language may exhibit a continuum of wellformedness shown in native speakers' wordlikeness judgments. For nonsense syllables, accident gaps are judged no worse than attested syllables and some systematic gaps are judged no worse than accidental gaps.

The main goal of the present experiment was not to investigate the gradience property of phonotactic grammars elicited from Taiwan Southern Min native speakers, but to examine the predictability of the three maximum entropy models of phonotactic learning. The comparison based on lexical status across the models in previous section could only account for certain categorical phonotactic patterns. In this section will be to report an experiment designed to collect wordlikeness judgments from Taiwan Southern Min native speakers. Both Body model and the Rhyme model were capable of distinguishing attested syllables from nonsense ones but they constructed the constraints assigned with different feature definitions and quantitative weights. The inherent difference between the two models will thus be used to contrast the wordlikeness judgments of native speakers in order to determine Taiwan Southern Min syllable structure.

Nonsense syllables were most crucial to the present experimental study. Not only did they avoid the confounding properties of real words like semantic representations, age of acquisition, imagineability in wordlikeness judgments, but did also the wordlikeness of nonsense syllables serve as a useful counterpart to contrast the violations scores and maxent values generated from the maximal entropy models of phonotactic learning.

4.1 Methods

4.1.1 Participants

Seventeen undergraduates (7 males and 10 females) in National Chung Cheng University in Chiayi Taiwan were recruited in the experiment. All of them acquired Taiwan Southern Min as the first native language and were speaking Taiwan Southern Min frequently at home. No hearing or speech disorder was self-reported by the participants.

LI: MODELING SUBSYLLABIC STRUCTURE

4.1.2 Materials

Because of the realistic consideration of experimentation, we could not use all the gigantic amount of testing data (a total of 4,607 items) in a single experimental task. We shrank down the size of the potential materials, yet still selective, balanced, and representative, following the method in Bailey & Hahn's (2001).

Bailey & Hahn's (2001) criterion was originally used to select the materials with the greatest variety of neighborhood density as well as phoneme transition probability of all the potential materials. All of the testing data were divided into three categories: attested syllables, near-misses, and isolates. Attested syllables were those reported in Tsai (2000). Near-misses were the nonsense syllables that differed from the nearest attested neighbors by exactly one phoneme. Isolates were also the nonsense syllables but differed from the nearest attested neighbors by exactly two phonemes. At the onset of selecting the materials, twenty-two isolates were chosen at random. For each isolate, we then identified all the neighboring near-misses, that is, those that differed from the isolates by one phoneme, as well as differed from the nearest attested syllables by one phoneme. The process resulted into 259 syllable types, including 22 isolates and 237 near-misses. Thanks to the *cba* package (Buchta & Hahsler, 2006) in R, we were able to automate the process of selecting the materials.

In addition to isolates and near-misses, 68 attested syllables were randomly chosen as fillers. The complete set of 307 syllables (plus 20 other syllables for practice) was prepared by a female Taiwan Southern Min native speaker. For all the stimuli, the speaker read them in a falling tone consistently. Since syllables with obstruent codas in Taiwan Southern Min were pronounced with shorter duration and falling pitch contour, the speaker customized the syllables with obstruent codas with “typical” entering tones but the other syllables types with normal falling tones. Fairly speaking, tonotactics (i.e., the constraints of co-occurrence between segments and tones) should be a potential issue and awaits future research. The auditory stimuli prepared by the speaker were simultaneously recorded and digitized in 22k Hz using Praat (Boersma & Weeninks, 2008).

4.1.3 Procedure

After the recruitment, participants were seated in a sound-proof booth to perform the task. They were then instructed to put on a pair of headphones and to make wordlikeness judgments to the auditory stimuli by pressing the labeled keys on the keyboard. The numbers on the labels (from 1 to 9 in a continuous scale) denoted the meanings from *very unlike Taiwan Southern Min* to *very like Taiwan Southern Min*, as participants were notified in advance.

Before the real trials began, twenty practice trials whose stimuli were not included in the real trials were used to familiarize participants with the judgments. Each trial initialized with a visual warning of 300 ms, followed by an auditory stimulus from the headphones and a visual cue on the monitor. Without time restriction, participants were able to judge wordlikeness of the auditory stimulus using their Taiwan Southern Min

intuitions. The next trial proceeded once the judgment was made. The order of all the trials was randomized with the aid of E-prime (Schneider, Eschman, & Zuccolotto, 2002). The complete experimental procedure took nearly twenty-five minutes.

4.2 Results

Prior to analyses, judgment scores of the stimuli were averaged across participants. Judgment scores of real syllables were significantly higher than those of nonsense syllable (average: 5.95 vs. 4.49; $t(324) = -7.567$, $p < .01$ by item; $t(32) = -4.234$, $p < .01$ by participant). Only judgment scores of nonsense syllables would be relevant to the present study. As judgment scores of 259 nonsense syllables served as the predictor, violation scores, maxent values and probability served as the dependent measures in the Spearman correlation analyses. A significant correlation effect between judgment scores and other measures was obtained in the Body model ($\rho = -.15$, $S = 3766604$, $p < .01$), but not in the Rhyme model ($\rho = -.10$, $S = 3608572$, $p > .01$), the Flat model ($\rho < -.01$, $S = 3270060$, $p > .01$) and the Margin model ($\rho = -.09$, $S = 3593767$, $p > .01$)². Table 2 illustrates correlations between judgment scores and violation scores across the four models. The statistics thus supported that the grammars from the Body model better predict Taiwan Southern Min native speakers' intuitions more than those from the other models.

² Maxent values and probability were algorithmically derived from violation scores. Three measures thus yielded identical correlation results.

LI: MODELING SUBSYLLABIC STRUCUTURE

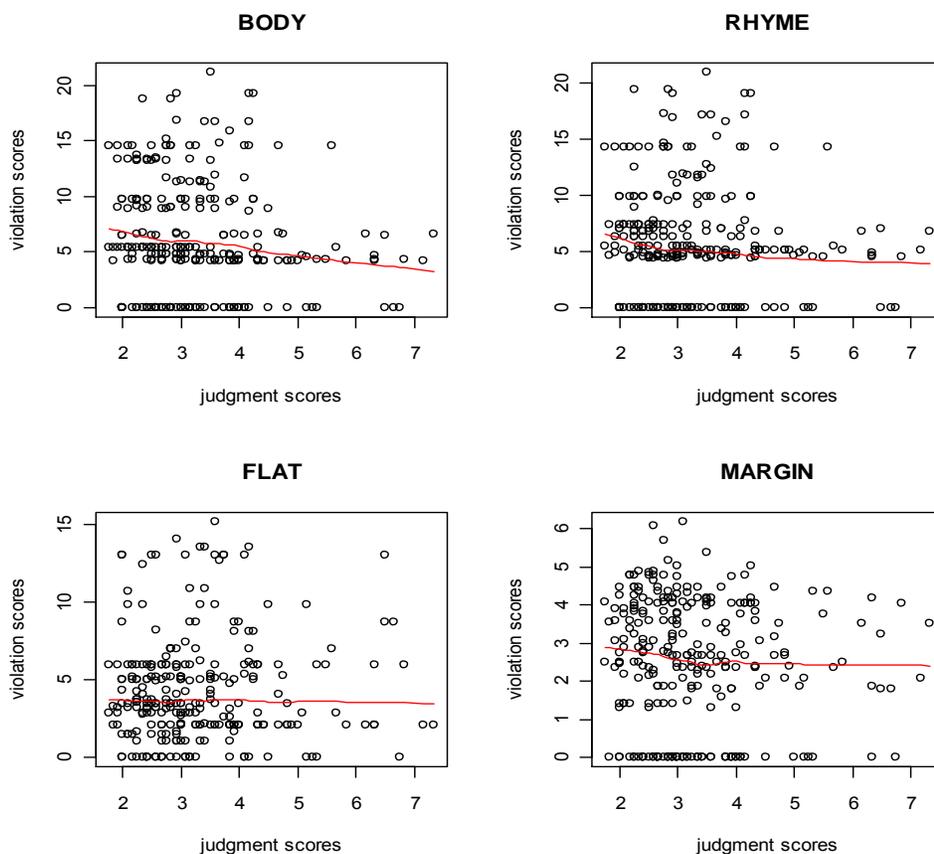


Figure 2. Correlations of judgment scores with violation scores across models

5. Conclusion

The issue of subsyllabic structure in Taiwan Southern Min had been unsettled since different sources of evidence came up with contradictory conclusions. The present study attempted to resolve the issue using a novel methodology. By applying maximum entropy models to assess the relative cohesiveness of segments across different subsyllabic domains, we discovered the Body model not only better predicted lexicality of attested and nonsense syllables but also confirmed the continuous wordlikeness judgments that reflected native speakers' phonotactic grammars.

The present results contradicted previous linguistic claims (Bao, 1990; Cheng & Cheng, 1977; Chung, 1996) espousing onset and rhyme as subsyllabic constituents in Taiwan Southern Min, but followed up Derwing and his colleague's (Derwing, 2007; Wang & Derwing, 1993) experimental finding that body-coda structure was processed in Taiwan Southern Min linguistic performance. Moreover, the lesson from which the Flat model and the Margin failed to simulate Taiwan Southern Min phonotactic behaviors was

LI: MODELING SUBSYLLABIC STRUCTURE

that hierarchical and sequential subsyllabic constituency were the essential linguistic mechanisms in Taiwan Southern Min.

This research line awaits future work on certain issues. First, maximum entropy models would be equivalently suitable for investigating the controversial status of the prevocalic or postvocalic glides. Second, tonotactics would be an issue inviting us to consider the interaction of segments and tones in dealing with syllable structure. Third, an extension of this study to other body-coda languages (e.g., Korean) would help reexamine the present conclusion.

REFERENCES

- Anderson, J. (1988). More on slips and syllable structure. *Phonology*, 5, 157-159.
- Bao, Z. (1990). Fanqie languages and reduplication. *Linguistic Inquiry*, 21(3), 317-350.
- Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory & Language*, 44, 569-591.
- Berger, A. L. Della Pietra, S. A., & Della Pietra, V. J. (1996). A maximum entropy approach to natural language processing. *Computational Linguistics*, 22, 39-71.
- Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206-244). Cambridge: Blackwell Publishers.
- Boersma, P. & Weenink, D. (2008). Praat: Doing phonetics by computer (Version 5.0.25 [Computer Program]. Retrieved May 31, 2008, from <http://www.praat.org/>
- Buchta, C. & Hahsler, M. (2006). cba: Clustering for Business Analytics. R package version 0.2-1.
- Chomsky, N. & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of Linguistics*, 1(2), 97-214.
- Cheng, M.-C. (2002). Mora and Southern Min. *Studies in Language Teaching, Linguistics, and Literature*, 4, 71-108.
- Cheng, R. L. & Cheng, S. S. (1977). *Phonological Structure and Romanization of Taiwanese Hokkien*. Taipei: Student.
- Cheng, R. L. & Tseng, C.-C. (1997). Form, function, interaction, and integration of stress in tonal languages. [Zhongyin cai shengdiao yuyeng zhong de xingshi kongneng hudong ji zhengh] [重音在聲調語言中的形式、功能、互動及整合]. In C.-T. Cheng (Ed.), *Chinese Languages & Linguistics IV: Typological Studies of Language in China* (pp. 225-263). Taipei: Academia Sinica.
- Chung, R.-F. (1996). *The Segmental Phonology of Southern Min in Taiwan*. Taipei: Crane.
- Chung, R.-F. (1999). Mora and Taiwanese syllable structure. *The Tsing Hua Journal of Chinese Studies, New Series*, 29(4), 533-549.
- Clements, G. N. & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the*

LI: MODELING SUBSYLLABIC STRUCTURE

- Syllable*. Cambridge, Massachusetts: MIT.
- Davis, S. (1989). On a non-argument for the rhyme. *Journal of Linguistics*, 25, 211-217.
- Derwing, B. L. (2007). What's in CVC-like things? Ways and means to look at phonological units across languages. M.-J. Solé, P. S. Beddor, & M. Ohala (Eds.), *Experimental Approaches to Phonology* (pp. 325-338). Oxford: Oxford University Press.
- Fowler, C. A. (1987). Consonant-vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. *Speech Communication*, 6, 231-244.
- Frisch, S. A., Large, N. A., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, 42, 481-496.
- Fudge, E. C. (1969). Syllable. *Journal of Linguistics*, 5, 253-286.
- Fudge, E. C. (1973). *Phonology*. Harmondsworth: Penguin Books.
- Fudge, E. C. (1987). Branching structure within the syllable. *Journal of Linguistics*, 23, 359-377.
- Frisch, S. & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics*, 30, 139-162.
- Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Cambridge, MA: Blackwell.
- Greenberg, J. H. & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, 20, 157-177.
- Hammond, M. (2004). Gradience, phonotactics, and the lexicon in English phonology. *International Journal of English Studies*, 4: 1-24.
- Hay, J., Pierrehumbert, J. B. & Beckman, M. (2003). Speech perception, well-formedness, and the statistics of the lexicon. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in Laboratory Phonology VI* (pp. 58-74). Cambridge: Cambridge University Press.
- Hayes, B. P. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253-306.
- Hayes, B. & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3), 379-440.
- Hockett, C. F. (1973). Where the tongue slips, there slip I. In Fromkin (Ed.), *Speech Errors as Linguistic Evidence* (pp. 93-119). The Hague: Mouton.
- Hung, W.-J. (1994). The Southern Min dialects of Kauhsiung. Ms. Academia Sinica, Taipei.
- Hyman, L. M. (1985). *A Theory of Phonological Weight*. Dordrecht: Foris Publications.
- Kahn, D. (1976). *Syllable-based Generalization in English Phonology*. Bloomington: Indiana University Linguistic Club.
- Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10(3), 421-441.
- Kirby, J. P. & Yu, C. L. (2007). Lexical and phonotactic effects on wordlikeness

LI: MODELING SUBSYLLABIC STRUCTURE

- Judgments in Cantonese. Paper presented ICPHS XVI, Saarbrücken, 6-10 August 2007.
- Lavor, J. (1979). Slips of the tongue as neuromuscular evidence for a model of speech production. In H. W. Doehert & M. Raupach (Eds.), *Temporal Variables in Speech. Studies in Honor of Frieda Goldman-Eisler* (pp. 21-26). The Hague: Mouton.
- Levin, J. (1985). A Metrical Theory of Syllabicity. PhD dissertation, MIT.
- Li, J. K. P. (1989). On the nature of syllabic final-ʔ in Southern Min dialects. *The Bulletin of the History & Philology Academic Sinica*, 60(3), 482-492.
- Lin, Y. H. (1989). Autosegmental treatment of segmental process in Chinese phonology. PhD dissertation. University of Texas, Austin.
- McCarthy, J. J. & Prince, A. (1990). Word and foot in prosodic morphology: The Arabic broken plural. *Natural Language & Linguistic Theory*, 8, 209-283.
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210-277.
- Ohala, J. J. & Ohala, M. (1986). Testing hypotheses regarding the psychological reality of morpheme structure constraints. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental Phonology* (pp. 239-252). San Diego, CA: Academic Press.
- Pierrehumbert, J. & Nair, R. (1995). Word games and syllable structure. *Language & Speech*, 38(1), 77-114.
- Prince, A. & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, Mass.: Blackwell.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime reference guide. Pittsburgh: Psychology Software Tools Inc.
- Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. I P. F. MacNeige (Ed.), *The Production of Speech* (pp. 109-136). New York: Springer.
- Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The Structure of Phonological Representations. Part II* (pp. 107-137). Dordrecht: Foris Publications.
- Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, 15, 49-74.
- Treiman, R. (1988). Distributional constraints and syllable structure in English. *Journal of Phonetics*, 16, 221-229.
- Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M. B. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 269-282). Cambridge: Cambridge University Press.
- Tung, C. S. (2001), editor. *Taiwan Southern Min Dictionary* [taiwan minnanyu zidian] [台灣閩南語辭典]. Taipei: Wu-Nan Books [wunan tushu] [五南圖書].
- van der Hulst, H. & Ritter, N. A. (2002). Theories of the syllable. In H. van der Hulst & N. A. Ritter (Eds.), *The Syllable: Views and Facts* (pp. 13-52). Berlin: Mouton de

LI: MODELING SUBSYLLABIC STRUCTURE

- Gruyter.
- Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech, 40*, 47-62.
- Wang, H. S. (1995). *Experimental Studies in Taiwanese Phonology*. Taipei: Crane.
- Wang, H. S. (1998). An experimental study on the phonotactic constraints of Mandarin Chinese. In B. K. T'sou (Ed.), *Studia Linguistica Serica* (pp. 259-268). Hong Kong: Language Information Sciences Research Center, City University of Hong Kong.
- Wang, H. S. & Chagn, C.-L. (2001). On the status of the prenucleus glide in Mandarin Chinese. *Language & Linguistics, 2*(2), 243-260.
- Wang, H. S. & Derwing, B. L. (1993). Is Taiwanese a 'body' language? In C. Dyck (Ed.), *Proceedings of the 1993 Annual Conference of the Canadian Linguistic Association* (pp. 679-694). Toronto Working Paper in Linguistics, University of Toronto.
- Yoon, Y. B. & Derwing, B. L. (2001). A language without a rhyme: Syllable structure experiments in Korean. *Canadian Journal of Linguistics, 46*(3), 187-237.

Variation in Tone 3 Sandhi: The case of prepositions and pronouns*

Chiung-Yao Wang and Yen-Hwei Lin
Michigan State University

This paper examines variation in Tone 3 Sandhi (henceforth T3S) with a focus on the case of prepositions and pronouns. Prepositions have been observed to produce an additional T3S surface pattern. Pronouns are found to behave differently from nouns in T3S application. The major approach for analyzing these cases is primarily the pre-OT (Optimality Theory) derivational process through which multiple T3S patterns surface in the output. Zhang (1997) analyzes examples with prepositions and pronouns within the OT framework. However, her approach requires a two-step two-tableau process. We adopt Coetzee's (2006) OT variation model to account for (i) the behavior of prepositions and pronouns in T3S and (ii) the variability in T3S surface forms, two of the challenging aspects in analyzing T3S application in sentences. The advantages of our re-analysis are twofold: multiple surface patterns are obtained in one step, and the relative frequency of surface patterns is indicated.

1. Introduction

Mandarin Tone 3 Sandhi (henceforth T3S) has been the most extensively studied tone sandhi phenomenon in Mandarin Chinese. T3S changes a Tone 3 (T3) to a Tone 2 (T2) before another T3. This simplified description of T3S is given in (1).

(1) Mandarin T3S:

T3 → T2/___ T3

(T2 = mid rising tone (35))

T3 = dipping tone (214) phrase finally and low tone (21) elsewhere)

The rule describe in (1) is deceptively simple as how T3S applies becomes rather complicated in cases where there are more than two T3s in a word or phrase (Lin 2007:204). In phrases and sentences where there are multiple adjacent T3's, T3S often exhibits variation by allowing more than one surface pattern depending on syntactic structure and how it is prosodified as shown in (2).

* We thank the participants of NACCL-23 for their comments and questions.

- (2) [Lao Li] [mai [hao jiu]]
 Lao Li buy good wine
 ‘Old Li buys good wine.’
 T3 T3 T3 T3 T3
 → (T2 T3) (T3) (T2 T3)
 → (T2 T2 T3) (T2 T3)
 → (T2 T2 T2 T2 T3)

In accounting for variation in T3S, the primary approach has been the pre-OT derivational analysis (Cheng 1973, Shih1986, 1997, Chen 2000, Duanmu 2000/2007, Lin 2007). The additional patterns are often considered to be present in faster or casual speech (e.g. Cheng 1973, Shih1986, 1997, Zhang 1997, Chen 2000, Lin 2007), but are considered to be just possible variants unrelated to speech rates in some studies (e.g. Duanmu 2000/2007). In either case, the question is how variation in Mandarin T3S can be modeled in Optimality Theory (OT).

In this paper, we analyze simple examples containing a prepositional phrase (PP) with or without a pronoun by adopting Coetzee's (2006) OT variation model. The next section (§2) presents the data of T3S and previous analyses, and §3 provides background for modeling variation in OT. Our proposed OT analysis is presented in §4. The concluding section (§5) offers remarks on theoretical implications and suggests what can be done in future T3S studies.

2. T3S: the case of prepositions and pronouns

In this section, we first look at the T3S data, focusing on examples containing prepositions and pronouns. Previous analyses are presented, both a pre-OT analysis (Shih 1986, 1997, Chen 2000, Lin 2007) and an OT analysis (Zhang 1997).

2.1. The data

Consider the examples in (3) first. ST3¹ is grammatical in (3b) but it is not in (3a) although the sentences have the same branching structure. The fact that ST3 surfaces in (3b) is not because the lexicon *bi* ‘than’ is a special case. The same pattern is found in other prepositions (e.g. *wang* ‘toward’ in [*Gou* [[*wang bei*] *zou*]] ‘The dog walks toward the north’). The fact that prepositions tend to be prosodically weak is responsible for the additional pattern.

¹ ST3 is ungrammatical for some audience at NACCL-23. One participant pointed out that the judgment of grammaticality may result from regional differences (which we agree), and the absence of the syllable “*gou*” in T2 (sandhi tone for *gou* ‘dog’) may play a role in speakers’ judgment. However, 苟且偷生 *gou qie tou sheng* ‘to drift and live without purpose’ T3T3T1T1 → T2T3T1T1 is found, so the lack of an independent word for the syllable *gou* with T2 does not prohibit a sandhi tone to appear on the surface.

(3) Same branching structure with or without a preposition (cf. Zhang 1997)

a. Without a preposition					b. With a preposition				
Ma	hen	shao	hou		Gou	bi	ma	xiao	
horse	very	rarely	roar		dog	than	horse	small	
'A horse rarely roars.'					'A dog is smaller than a horse.'				
3	3	3	3	UT	3	3	3	3	UT
(3)	(2	2	3)	ST1	(3)	(2	2	3)	ST1
(2	2	2	3)	ST2	(2	2	2	3)	ST2
*(2	3)	(2	3)		(2	3)	(2	3)	ST3

(UT= Underlying tones; ST1= surface tones; Surface pattern 1; ST2= Surface pattern 2...)

(4) Same branching structure containing a PP with or without a pronoun

a. Without a pronoun					b. PP containing a pronoun				
Gou	bi	ma	xiao		Gou	bi	wo	xiao	
dog	than	horse	small		dog	than	I	small	
'A dog is smaller than a horse.'					'A dog is smaller than me.'				
3	3	3	3	UT	3	3	3	3	UT
(3)	(2	2	3)	ST1	(3)	(2	2	3)	ST1
(2	2	2	3)	ST2	(2	2	2	3)	ST2
(2	3)	(2	3)	ST3	*(2	3)	(2	3)	

Like prepositions, pronouns are prosodically weak and their behavior differs from nouns. The sentences in (4a) and (4b) have the same branching structure containing a PP; the only difference lies in what follows the preposition: a noun or a pronoun. ST3 is grammatical in (4a) where a noun follows the preposition, but the same surface pattern is banned in (4b) where a pronoun follows the preposition.

The data in (3) and (4) show two challenges in analyzing T3S. First, any T3S analysis has to account for the variation as we see that one input is mapped onto more than one output. Second, sentences of the same syntactic branching structure do not always exhibit the same T3S surface patterns. That is, syntax alone is insufficient in accounting for the differences.

2.2. Pre-OT analysis

The contrasting patterns in sentences in (3) and (4) have been accounted for through different syntax-based prosodic parses. Surface variants are obtained by derivational steps (e.g. Shih 1986, 1997, Chen 2000: Ch.9, Duanmu 2000/2007: Ch.11, Lin 2007: Ch.9), as shown in (5) – (7). The derivational process follows that in Lin (2007).

(5) Without a preposition

[<i>Gou</i>	[[<i>hen</i>	<i>shao</i>]	<i>hou</i> .]]	
dog	very	rarely	roar	‘A dog rarely roars.’
T3	T3	T3	T3	UT
Normal speech variant				
T3	T3	T3	T3	Word level: Not applicable
T3	(T2	T3)	T3	Phrase level: Disyllabic foot, T3S
T3	(T2	T2	T3)	Phrase level: Incorporation, T3S
(T3	T2	T2	T3)	Phrase level: Incorporation, no T3S; ST1
Fast speech or alternative variant				
(T2	T2	T2	T3)	One prosodic domain, T3S; ST2

In (5), at the Word level, T3S is not applicable. At the Phrase level, a disyllabic foot is parsed for the smallest domain *hen shao* ‘very seldom’, and T3S applies. When the verb *hou* ‘roar’ is incorporated, T3S applies again. Finally, when *ma* ‘horse’ is incorporated, T3S does not apply because there are no adjacent T3’s. The first surface pattern derived is (T3**T2**T2T3). ST2 is a pattern in fast speech where a larger domain is parsed and T3S applies from left to right (Shih 1986, 1997, Chen 2000, Lin 2007), or it is an alternative pattern unrelated to fast speech (Duanmu 2000/2007). In (6), the branching structure is the same as that in (5), but in (6) there is a preposition which can, but does not have to cliticize.

(6) With a preposition, PP containing a noun

[<i>Gou</i>	[[<i>bi</i>	<i>ma</i>]	<i>xiao</i> .]]	
dog	than	horse	small	‘Dogs are smaller than horses.’
T3	T3	T3	T3	UT
No cliticization				
T3	T3	T3	T3	Word level: Not applicable
T3	(T2	T3)	T3	Phrase level: Disyllabic foot, T3S
T3	(T2	T2	T3)	Phrase level: Incorporation, T3S
(T3	T2	T2	T3)	Phrase level: Incorporation, no T3S; ST1

Fast Speech or alternative variant

(T2	T2	T2	T3)	One prosodic domain, T3S; ST2
-------------	-----------	-----------	-----	----------------------------------

Cliticization

T3	T3	T3	T3	Word level: Not applicable
(T2	T3)	T3	T3	Cliticization, T3S
(T2	T3)	(T2	T3)	Phrase level: disyllabic foot, T3S; ST3

In (6), when the preposition does not cliticize, the derivation is the same as that in normal speech variant in (5). In the case of ST1 in (6), the preposition is not treated differently from other word categories. ST2 is the fast speech or alternative variant when a larger domain is parsed and T3S applies from left to right. When the preposition cliticizes, ST3 surfaces. At the Word level, T3S is not applicable. Next, the preposition cliticizes to the host on its left, and T3S applies. Finally, a disyllabic foot is formed for the rest of the syllables, and T3S applies again. ST3 (**T2T3**)(**T2T3**) is thus derived. This same surface pattern is ungrammatical in (7) where a pronoun follows the preposition.

(7) With a preposition, PP containing a pronoun

[<i>Gou</i>	[[<i>bi</i>	<i>wo</i>]	<i>xiao</i>]]	
dog	than	I	small	‘Dogs are smaller than me.’
T3	T3	T3	T3	UT

No cliticization

T3	T3	T3	T3	Word level: Not applicable
T3	(T2	T3)	T3	Phrase level: Disyllabic foot, T3S
T3	(T2	T2	T3)	Phrase level: Incorporation, T3S
(T3	T2	T2	T3)	Phrase level: Incorporation, no T3S; ST1

Fast Speech or alternative variant

(T2	T2	T2	T3)	One prosodic domain in fast speech, T3S; ST2
-------------	-----------	-----------	-----	---

Cliticization

T3	T3	T3	T3	Word level: Not applicable
T3	(T2	T3)	T3	Phrase level: Cliticization, T3S
T2	(T2	T2	T3)	Phrase level: Incorporation, T3S
(T3	T2	T2	T3)	Phrase level: Incorporation, no T3S; ST3 (= ST1)

In (7), the derivational steps for ST1 and ST2 are the same as those in (6). For ST3, at the Word level, T3S is not applicable. At the Phrase level, the pronoun *wo* ‘I’ cliticizes

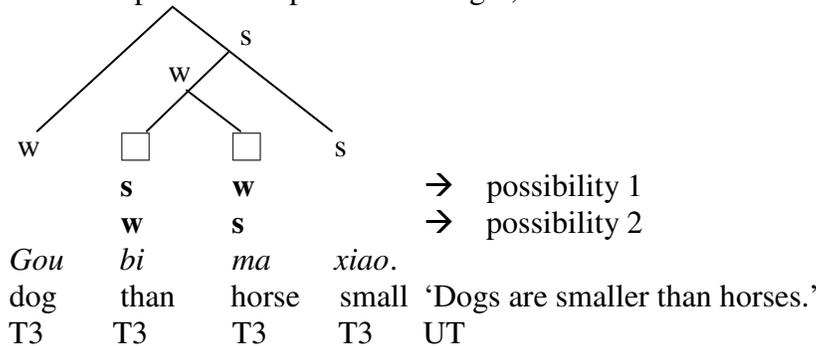
to the preposition *bi* ‘than’, and T3S applies. According to Shih (1997:110), a pronoun is weaker than a preposition and will cliticize leftwards onto the preposition, and as a two-syllable structure, it will not cliticize further because it is heavy enough. When the last syllable *xiao* ‘small’ is incorporated in the final step, T3S does not apply as there are no adjacent T3’s. The derived pattern ST3 is T3T2T2T3, identical to the pattern in ST1, even though the derivational processes of ST1 and ST3 differ.

As shown in (5) – (7), the pre-OT analysis accounts for all the surface patterns in the contrasting data derivationally. We now turn to Zhang (1997) who offered an OT account for the data.

2.3. Zhang’s (1997) two-step two-tableau OT analysis

A major difference in Zhang’s analysis (1997) from the previous studies is her use of unspecified *strong/weak* in the metrical structure for preposition phrases in accounting for multiple T3S outputs. Zhang adopts Cinque’s (1993) Null Theory of Phrase Stress and she takes strong/weak stress in Cinque’s theory as the strong/weak constituent strength (Zhang 1997:304). According to Zhang, a prepositional phrase is unspecified for its strength, and it can be either strong or weak (Zhang 1997:304-305). The sentence in (8) shows unspecified strength for the PP.

(8) PP is unspecified for prosodic strength, so it can be either strong or weak.



A sentence that has a PP, unspecified for its prosodic strength, then enters the two inputs in two separate OT tableaux for evaluation. It should be noted that the optimal outputs for PP-strong-weak and PP-weak-strong may completely overlap, partially overlap, or do not overlap. Six constraints used in Zhang’s (1997) analysis are given in (9).

- (9) Constraints used in Zhang (1997:306-308)
 - a. *33 (no sequential third tones): No adjacent third tones are allowed.
 - b. C1 (Clitic Dependency): A clitic cannot be separated from the TS domain of the preceding verb or preposition head.

- c. Max (Maximal Domain): The maximal TS domain is two syllables in normal speaking rate, but larger in more casual or faster style.
- d. Align-Di-L (Disyllabic Constituent Alignment): Align the left side of a TS domain with the left side of a disyllabic constituent when two or more TS domains occur.
- e. PTAS (Parse UT of an Absolutely Strong Node): The underlying tone of a strong constituent which is not dominated by any *w* node must be parsed.
- f. PTRS (Parse UT of a Relatively Strong Node): The underlying tone of a strong constituent which is dominated by at least one *w* node must be parsed.

Max maintains that the ideal length of a prosodic domain is two syllables at the normal speech rate (Zhang 1997:308). PTAS and PTRS ensure a strong element to be faithful to its underlying tone (Zhang 1997:306). The constraint Align-Di-L, according to Zhang’s definition, is irrelevant when there is only one T3S domain. Zhang uses the symbol □ to denote the unspecified constituent strength, which could be either *strong* or *weak*. We now turn to two examples of her analysis in (10) – (11).

(10) Zhang's two-input-two-tableau analysis: PP containing a noun

[*Gou* [[*bi ma*] *xiao*]]

dog than horse small

‘Dogs are smaller than horses.’

a. PP is *ws* (Zhang 1997:314)

<u>wws</u> S 3333	PTAS	*33	PTRS	Align-Di-L	Max
a. (223)(3)		*!		*	*
b. (23)(22)	*!		*	*	
☞ c. (3)(223)			*		*
☞ d. (2223)			*		*
e. (23)(23)			*	*!	

b. PP is *sw* (Zhang 1997:315)

<u>ws</u> wS 3333	PTAS	*33	PTRS	Align-Di-L	Max
a. (223)(3)		*!	*	*	*
b. (3)(223)			*		*!
c. (2223)			*		*!
☞ d. (23)(23)				*	

In (10a) where PP is *ws*, Candidate (c) (T3)(**T2T2T3**) and Candidate (d) (**T2T2T2T3**) are optimal outputs. In (10b) where PP is *sw*, Candidate (d) (**T2T3**)(**T2T3**) is the optimal output. The optimal outputs from both tableaux are final winners. All three surface patterns for *Gou bi ma xiao* ‘Dogs are smaller than horses’ are accounted for.

(11) Zhang's two-input-two-tableau analysis: PP containing a pronoun
 [Gou [[bi wo] xiao]]
 dog than I small 'Dogs are smaller than me.'

a. PP is *ws* (Zhang 1997:316)

w _w S 3333	PTAS	*33	Cl	PTRS	Align-Di-L	Max
a. (223)(3)		*!			*	*
b. (23)(23)			*!	*	*	
☞c. (3)(223)				*		*
☞d. (2223)				*		*

b. PP is *sw* (Zhang 1997:316)

w _w S 3333	PTAS	*33	Cl	PTRS	Align-Di-L	Max
a. (223)(3)		*!		*	*	*
b. (23)(23)			*!		*	
☞c. (3)(223)				*		*
☞d. (2223)				*		*

In (11a) where PP is *ws*, Candidate (c) (T3)(T2T2T3) and Candidate (d) (T2T2T2T3) are optimal outputs. In (11b) where PP is *sw*, Candidate (c) (T3)(T2T2T3) and Candidate (d) (T2T2T2T3) are optimal outputs. Notice that the two optimal outputs are identical in the two tableaux. Nevertheless, the union of the winners from the two tableaux does account for the two surface patterns in the sentence *Gou bi wo xiao* 'Dogs are smaller than me', which has a PP containing a pronoun.

In sum, the variant patterns are produced by two inputs with different prosodic strength for PP and two evaluation tableaux. The outputs are combinations of the winners from both tableaux.

3. Background for modeling variation in OT

A question we raised earlier was: how do we model variation in OT? There have been different approaches: (i) Partially ordered constraints (Kiparsky 1993, Reynolds 1994, Anttila 1997 *et seq.*, Anttila & Cho 1998), (ii) Unranked constraints plus two inputs and two tableaux (Zhang 1997), and (iii) Rank-ordering of EVAL (Coetzee 2006). In what follows, these approaches² are briefly introduced.

² Probabilistic OT models with numerically valued constraints, such as Stochastic OT and Harmonic Grammar are not discussed here (see Coetzee & Pater in press, for a review and references).

Partially ordered constraints (Kiparsky 1993, Reynolds 1994, Anttila 1997 et seq., Anttila & Cho 1998)

In standard OT, a grammar of a language is a total ordering of a ranked set of constraints, typically yielding a single output from each input. However, in a model with partially ordered constraints, a grammar is a partial ordering of the constraint set. Each time when the grammar evaluates a candidate set, one of the rankings consistent with the partial order is randomly chosen. Variation results when some of these rankings select different outputs. The illustration in (12) shows a partial ordering where C is ranked above C1, and C is ranked above C2, but C1 and C2 are crucially unranked.

(12) Partial ordering: $C \gg C1$ and $C \gg C2$

	C	C1	C2
☞ Candidate (a)			*
Candidate (b)		*!	
Candidate (c)	*!		

	C	C2	C1
Candidate (a)		*!	
☞ Candidate (b)			*
Candidate (c)	*!		

In (12), since C1 and C2 are crucially unranked (partially ordered) and either $C \gg C1 \gg C2$ or $C \gg C2 \gg C1$ is consistent with the partial ordering. Both Candidates (a) and (b) in (12) are possible outputs of the grammar.

Unranked constraints plus two inputs and two tableaux (Zhang 1997)

Zhang's (1997) analysis with unranked constraints plus two inputs and two tableaux has been presented in § 2.3. The approach is summarized in (13).

(13) Unranked constraints plus two inputs and two tableaux (Zhang 1997)

- a. Unranking of some constraints may generate more than one output since some candidates may have a tie under the evaluation.
- b. Two inputs are fed into two separate tableaux, potentially producing additional variants.
- c. Grammatical outputs are the union of the optimal outputs from the two tableaux.

Rank-ordering of EVAL (Coetzee 2006)

In Coetzee's (2006) rank-ordering approach, EVAL not only distinguishes winners from losers, but also imposes a well-formedness ranking-ordering on the whole candidate set. Variation does not arise as a result of variation in grammar (i.e. ranking) as in the previous models; rather, variation depends on how EVAL imposes on a well-formedness rank-ordering on the candidates. In most cases, only the topmost candidate on the rank-ordering is well-formed and grammatical, but under some circumstances, two or more candidates can be well-formed enough to be considered grammatical. Coetzee (2006:338) suggests, "What needs to be added to the grammar is a mechanism that will allow, in some circumstances, more than one of the already generated possible output forms to become actual outputs."

Importantly, there is only one consistent ranking, and added to the evaluation tableau is a critical cut-off line, below which constraint evaluation does not rule out candidates. That is, variation occurs when two or more candidates are not ruled out by higher ranked constraints upon reaching the cut-off. The relative degree of well-formedness of those candidates that pass through the cut-off indicates the relative frequency of the variants. Let us now look at the tableau which illustrates the Rank-ordering Model.

(14) Illustration of the Rank-ordering Model Coetzee (2006:343)

Cut-off

	C1	C2	C3	C4
☞ a. Cand-1				*
☞ b. Cand-2			*	
c. Cand-3		*!		
d. Cand-4	*!			

The constraint set is divided into two strata, separated by the critical cut-off. The stratum above the cut-off is the same as classic OT where non-optimal candidates do not surface. The candidates that survive upon reaching the cut-off are grammatical outputs. As we see in (14), Candidates 1 and 2 are winners, i.e. grammatical outputs. Candidates 3 and 4 have fatal violations above the cut-off, hence are ungrammatical. Violations of the constraints in the stratum below the cut-off are not fatal. Variation arises when two or three candidates are disfavored only by constraints below the cut-off, and the extent of constraint violation (i.e. the degree of well-formedness) suggests relative frequency. Candidate 1 and Candidate 2 both have one violation below the cut-off. However, Candidate 1 occurs more frequently than Candidate 2 because it has a higher degree of well-formedness by violating a lower ranked constraint.

4. Proposed OT analysis of Mandarin T3S

In this section, we present our proposed OT analysis of Mandarin T3S based on Coetzee's (2006) Model.

4.1 Constraints

The constraints, along with their definitions, that we used in our analysis are given in (15).

(15) Constraints

- a. ID (T): The tone in the output is faithful to its tone in the input.
- b. ID (T, PREP): The tone of a preposition is faithful to its tone in the input.
- c. *(33): No adjacent T3's within the same prosodic domain (based on Zhang (1997) with a slight revision).
- d. *33: No adjacent T3's, even if they belong to different prosodic domains.
- e. *DEG-FT (Degenerate Foot): Monosyllabic feet are banned.
- f. BOUND (BOUNDNESS): For every prosodic domain that has more than two syllables, assign a violation mark (based on Chen (2000) with a slight revision).
- g. SUBJ-PRED. BOUNDARY: There is a strong subject-predicate boundary. No prosodic domain should go across this boundary.
- h. MATCH (PROS, SYN): For each prosodic domain, there must be some syntactic constituent that matches with it. If a prosodic domain does not have a syntactic domain that matches with it, assign a violation mark.

e.g. [Gou [[bi ma] xiao.]]

dog than horse small

'A dog is smaller than a horse.'

3	3	3	3	UT
(3)	(2	2	3)	→ no violation
(2	2	2	3)	→ no violation
(2	3)	(2	3)	→ two violations

- i. WRAP (FC-FC)³: Adjacent prosodically weak functional elements in the same PP belong to the same prosodic domain (based on the phonology-syntax interface constraint, WRAP-XP, which prevents XPs from splitting up into multiple p-phrases in Truckenbrodt (1999, 2007)).

³ A possible motivation for WRAP (Fc, Fc) might be that consistency of directionality of cliticization is preferred. That is, this constraint prevents a situation where the first element cliticizes to the left and the second element to the right.

e.g.	[Gou	[[bi	wo	xiao.]]	
	dog	than	I	small	
	'A dog is smaller than me.'				
	3	3	3	3	UT
	(3)	(2	2	3)	→ no violation
	(2	2	2	3)	→ no violation
	*(2	3)	(2	3)	→ violation

4.2 The cut-off point and the ranking process

An essential question to ask is: how should these constraints be ranked and how is the cut-off point determined? Coetzee proceeds by listing *Violation Profiles of Observed Outputs* first (2006:350) and suggests the following procedure of determining where to place the cut-off point. Suppose we have two variant outputs. The input, the two variant outputs, and the constraints each output violates are listed. The constraints that both variant outputs violate are underlined. The underlined constraints cannot distinguish the variants since they are violated by both outputs. Crucially, these constraints must be placed below the cut-off to ensure the two variants do surface. It should be kept in mind that variant outputs must not be disfavored by a constraint above the cut-off. When the frequencies of two variants are available, and we know that one occurs more frequently than the other, we will need some constraint that can distinguish the two. For instance, if *Variant A* that violates *Constraint 1* is higher in frequency than *Variant B* that violates *Constraint 2*, *Constraint 1* has to be ranked lower than *Constraint 2*. By ranking *Constraint 1* lower, a violation of it is less serious, which reflects that *Variant A* is more well-formed, and therefore, occurs more frequently. In addition, in the initial ranking, Coetzee (2006:350) points out that Faithfulness constraints should be ranked low, in accordance with the notion of ranking conservatism (Tesar & Smolensky 1998), unless we have contrary evidence.

Next, we have to ensure that the two variants mentioned above are indeed the best outputs, that is, more well-formed than the rest of the candidates in the competition. According to Coetzee (2006), to ensure that there are no additional grammatical outputs other than the two variant outputs, the rest of the candidates must be eliminated by some constraint above the cut-off. Placing those constraints that are fatally violated by the ungrammatical outputs above the cut-off will guarantee that these output candidates can no longer compete upon reaching the cut-off. That is, for these ungrammatical output candidates, the competition ends at the cut-off point. In this model, constraint ranking is established upon what is observed in the data— what are the variant outputs, and what are not. Then, the constraint ranking is adjusted step by step. We provide selected examples of how we rank our constraints, but omit the whole step-by-step procedure in the adjustment of the ranking argument due to the space limit.

We begin the procedure by listing the constraints that are violated by optimal outputs in five sentences. The violation profiles of the observed outputs are shown in

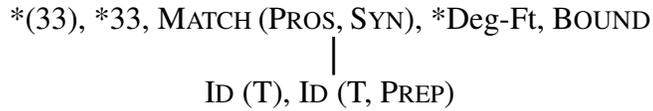
Table 1. Following Coetzee (2006:350), the constraints shared by the optimal outputs are underlined. This indicates that these constraints can neither distinguish the variants nor eliminate them.

Table 1: Violation profiles of the observed outputs:

Input	Variants in the Output	Constraints violated
a. [[Ma] [[<i>hen shao</i>] <i>hou</i> .]] (Zhang 1997:305) horse very rarely roar ‘Horses very rarely roar.’ [[σ] [[σ σ] σ]] 3 3 3 3 UT	(3) (223)	<u>BOUND</u> , *Deg-Ft, <u>ID</u> <u>(T)</u>
	(2223)	<u>BOUND</u> , <u>ID</u> (T)
b. [[Gou] [[<i>bi ma</i>] <i>xiao</i> .]] (Zhang 1997:293) dog than horse small ‘Dogs are smaller than horses.’ [[σ] [[σ σ] σ]] Clitic 3 3 3 3 UT	(3) (223)	ID (T, PREP), BOUND, <u>ID</u> (T), *Deg-Ft
	(2 2 2 3)	ID (T, PREP), BOUND, <u>ID</u> (T)
	(2 3) (2 3)	MATCH (PROS, SYN), <u>ID</u> (T)
c. [[Gou] [[<i>bi wo</i>] <i>xiao</i> .]] (Zhang 1997:307) dog than I small ‘Dogs are smaller than I.’ [[σ] [[σ σ] σ]] Clitic Clitic 3 3 3 3 UT	(3)(223)	<u>ID</u> (T, PREP), <u>BOUND</u> , <u>ID</u> (T), *Deg-Ft
	(2 2 2 3)	<u>ID</u> (T, PREP), <u>BOUND</u> , <u>ID</u> (T)
d. [[Lao Li] [<i>zhao xie</i> .]] (Lin 2007:211) Old Li look for shoes ‘Old Li looks for shoes.’ [[σ σ] [σ σ]] 3 3 3 2 UT	(23)(32)	<u>ID</u> (T), *33
	(2232)	<u>ID</u> (T), BOUND
e. [[wo] [<i>xiang [mai bi]</i> .]] (Lin 2007:215) I want buy pen ‘I want to buy pens.’ [[σ] [σ [σ σ]]] 3 3 3 3 UT	(23)(23)	MATCH (PROS, SYN), <u>ID</u> (T)
	(2223)	BOUND, <u>ID</u> (T)

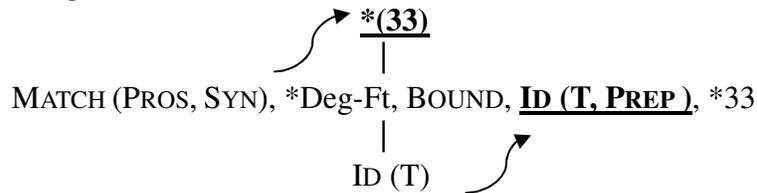
To rank the constraints conservatively, the initial ranking is to place the Markedness constraints over Faithfulness constraints, unless some evidence indicates otherwise. (16) shows the initial ranking.

(16) Initial ranking:



In (17), the arrows indicate from which strata the constraints that are underlined are moved. Adjustment of the ranking is made based on the violation information in Table 1. The constraint *(33) has been moved up to the highest in the ranking because it is never violated. ID (T, PREP) is sometimes violated, and sometimes not, so it is moved to the stratum that has constraints that are not always violated by the observed variants.

(17) Interim ranking



Interim ranking in (17) is not final and is used here for illustration purposes. Coetzee (2006) approaches his data by gradually ordering the constraints with the evidence of the variants. We follow the same procedure and make adjustments as evidence arises. These steps are critical in identifying exactly where the cut-off point is. In (18), we illustrate the partial ranking of our analysis with a simple sentence from Zhang (1997:295).

- (18) [You [[liang wan] mi.]
 there is two bowl (CL) rice 'There are two bowls of rice.'
 3 3 3 3 UT
 (3) (2 2 3) ST1
 (2 2 2 3) ST2

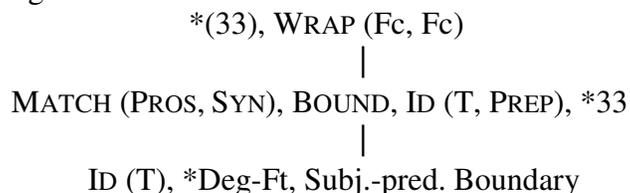
[[σ][[σσ]σ]]	*(33)	MATCH (PROS, SYN)	BOUND	*33	ID (T)	*Deg-Ft
☞ a. (3) (2 2 3)			*		**	*
☞ b. (2 2 2 3)			*		***	
c. (2 3) (2 3)		**!			**	
d. (2 2 3) (3)		*	*!	*	**	*
e. (3) (233)	*!		*	*	*	*
f. (33) (23)	*!	**		*	*	

Crucially, ST1 (T3)(T2T2T3) in (18) provides evidence that degenerate foot,

although disfavored prosodically, is in one of the surface patterns. This indicates that violation of the *Deg-Ft is not serious. It must be ranked below the cut-off. Otherwise, *Candidate (a)* would be ruled out by this constraint before reaching the cut-off, and that would be contradictory to the attested pattern ST1 in (18).

After following Coetzee’s procedure, the final ranking reached is shown in (19).

(19) Final ranking argument



In the next section, we use this final ranking to account for T3S variation in sentences given in (3) and (4).

4.3. The analysis

Consider first the sentence without a preposition, analyzed in (20).

(20) Without a preposition:

[Ma [[hen shao] hou.]]

horse very rarely roar

‘Horses very rarely roar.’

[σ [[σσ] σ]]	*(33)	WR AP	MAT CH (PROS, SYN)	BOU ND	ID (T, PREP)	*33	ID (T)	*DEG- FT	SUBJ.- PRED. BOUND ARY
☞ a. (3)(2 2 3)				*			**	*	
☞ b. (2 2 2 3)				*			***		*
c. (2 3)(2 3)			**!				**		*
d. (2 2 3)(3)			*	*!		*	**	*	*
e. (2 3)(3 3)	*!		**			**	*		*

Above the cut-off, the violation of *(33) rules out Candidate (e), MATCH (PROS, SYN) rules out Candidate (c), and both MATCH (PROS, SYN) and BOUND rule out Candidate (d). Candidates (a) and (b), both violating BOUND once, constitute a tie upon reaching the cut-off. Below the cut-off, no violation is fatal. Both Candidates (a) and (b) are grammatical. With respect to relative frequency, Candidate (a) has three violation marks and Candidate (b) has four, which suggests that Candidate (a) is more well-formed and hence occurs more frequently than Candidate (b).

The tableaux in (21) and (22) present our analysis of the sentences with a PP. In (21) where a noun follows the preposition, there are three variants.

(21) With a preposition:

[*Gou* [[*bi ma*] *xiao.*]]
 dog than horse small 'Dogs are smaller than horses.'

[σ [[σσ] σ]]	*(33)	WR AP	MAT CH (PROS, SYN)	BOU ND	ID (T, PREP)	*33	ID (T)	*DEG- FT	SUBJ.- PRED. BOUND ARY
☞ a. (3)(2 2 3)				*	*		**	*	
☞ b. (2 2 2 3)				*	*		***		*
☞ c. (2 3) (2 3)			**				**		*
d. (2 2 3) (3)			*	*	*!	*	**	*	*
e. (2 3) (3 3)	*!		**			**	*		*

Above the cut-off, Candidate (e) fatally violates *(33). Candidate (d) violates MATCH (PROS, SYN), BOUND, ID (T, PREP), and *33 and incur 4 violations upon reaching the cut-off and is out of the competition. Candidates (a-c) are tied as they all have two violation marks upon reaching the cut-off. Below the cut-off, no violation is fatal. Candidates (a), (b) and (c) are all grammatical outputs, although they have different numbers of violations (different degrees of well-formedness). Regarding relative frequency, Candidates (a) and (c) both have three violation marks, and Candidate (b) has four. This suggests that Candidates (a) and (c) are more well-formed, and therefore occur more frequently than Candidate (b). In contrast, as analyzed in (22), when a pronoun follows the preposition, there are two variants.

(22) PP with a pronoun

[*Gou* [[*bi wo*] *xiao.*]]
 dog than I small 'Dogs are smaller than I.'

[σ [[σσ] σ]]	*(33)	WR AP	MAT CH (PROS, SYN)	BOU ND	ID (T, PREP)	*33	ID (T)	*DEG- FT	SUBJ.- PRED. BOUN DARY
☞ a. (3)(2 2 3)				*	*		**	*	
☞ b. (2 2 2 3)				*	*		***		*
c. (2 3) (2 3)		*!	**				**		*
d. (2 2 3)(3)			*	*	*!	*	**	*	*
e. (2 3) (3 3)	*!	*	**			**	*		*

Above the cut-off, Candidate (c) fatally violates the WRAP constraint since the two elements of the PP are not parsed into the same prosodic domain. Candidate (e) violates both *(33) and WRAP in the top tier of the constraints, and is ruled out. Candidate (d) violates MATCH (PROS, SYN), BOUND, ID (T, PREP), and *33 and incurs 4 violations and is ruled out upon reaching the cut-off. Both Candidate (a) and Candidate (b) incur two violation marks upon reaching the cut-off. They are both grammatical outputs. Below the cut-off, Candidate (a) has three violation marks and Candidate (b) has four, which suggests that Candidate (a) is more well-formed and hence occurs more frequently than Candidate (b).

4.4. Summary

In sum, the one-to-many in the input-output mapping in T3S is a challenging aspect in T3S research and the variation is further complicated by the prepositions and pronouns. There has been little progress in analyzing these data in the OT framework. By adopting Coetzee's (2006) OT variation model, our proposed re-analysis of the T3S variation in the examples with prepositions and pronouns have the following merits: (i) This is a one-input and one-tableau analysis and multiple surface patterns are produced in one step, (ii) there is no need to assume that PP is unspecified for prosodic strength, and (iii) the relative frequency of the multiple surface forms is indicated, against which empirical data can be checked.

5. Concluding remarks

We start out with a few simple examples to illustrate that by adopting Coetzee's (2006) OT variation model, variation in Mandarin T3S can be analyzed with one input, one ranking, and one evaluation tableau. The encoding of relative well-formedness or relative frequency of the variants allows the proposal to provide the basis for future sociolinguistic and experimental studies. The notion of different degrees of well-formedness is also applicable to other domains. Preliminary results from one of our experiments show that children typically make the 'better' kinds of errors and the rank-ordering model provides a way to capture children's T3S error patterns.

Future studies will need to expand to additional examples with various syntactic structures, collect empirical data, and explore and compare several OT variation models. Careful sociolinguistic studies of the variation in Mandarin T3S are necessary to better understand the distribution of the surface variants, which can then be modeled formally. Lastly, little is known about how children acquire T3S in phrases and sentences, let alone the variation patterns through the developmental stages. It would be interesting to find out whether or not or the extent to which children produce the same variants attested in adults, and whether or not the relative frequency of children's patterns is similar to those of adults'.

REFERENCES

- ANTTILA, ARTO. 1997. Deriving variation from grammar. *Variation, change, and phonological theory*, ed. by Frans Hinskens, Roeland Van Hout, and W. Leo Wetzels, 35–68. Amsterdam: John Benjamins.
- ANTTILA, ARTO and YOUNG-MEE YU CHO. 1998. Variation and change in optimality theory. *Lingua* 104.31-56.
- CHEN, MATTHEW Y. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.
- CHENG, CHIN-CHUAN. 1973. *A synchronic phonology of Mandarin Chinese*. Mouton: The Hague.
- CINQUE, GUGLIELMO. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24.239-97.
- COETZEE, ANDRIES W. 2006. Variation as accessing ‘non-optimal’ candidates. *Phonology* 23.337-85.
- COETZEE, ANDRIES W. and JOE PATER. in press. The place of variation in phonological theory. *The handbook of phonological theory*, ed. by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu. Oxford: Wiley-Blackwell.
- DUANMU, SAN. 2000/2007. *The Phonology of Standard Chinese*. 1st edn./2nd edn. Oxford: Oxford University Press.
- KIPARSKY, PAUL. 1993. An OT perspective on phonological variation. Handout from *Rutgers Optimality Workshop 1993*, also presented at NWA 23 1994, Stanford University. On-line: <http://www.stanford.edu/~kiparsky/Papers/nwave94.pdf>.
- LIN, YEN-HWEI. 2007. *The sounds of Chinese*. Cambridge: Cambridge University Press.
- REYNOLDS, WILLIAM T. 1994. *Variation and phonological theory*. Philadelphia, PA: University of Pennsylvania dissertation.
- SHIH, CHILIN. 1986. *The prosodic domain of tone sandhi in Chinese*. La Jolla, CA: University of California San Diego dissertation.
- SHIH, CHILIN. 1997. Mandarin third tone sandhi and prosodic structure. *Studies in Chinese phonology*, ed. by Norval Smith and Jialing Wang, 81-123. Berlin: Mouton de Gruyter.
- TESAR, BRUCE and PAUL SMOLENSKY. 1998. Learnability in optimality theory. *Linguistic Inquiry* 29.229-68.
- TRUCKENBRODT, HUBERT. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30.219-55.
- TRUCKENBRODT, HUBERT. 2007. The syntax-phonology interface. *The Cambridge handbook of phonology*, ed. by Paul de Lacy, 435-56. Cambridge: Cambridge University Press.
- ZHANG, NING. 1997. The avoidance of the third tone sandhi in Mandarin Chinese. *Journal of East Asian Linguistics* 6.293-338.

Grass-Mud Horses to Victory: The Phonological Constraints of Subversive Puns

Seth Wiener¹

The Ohio State University

In 2008, Chinese netizens began creating *subversive puns*. These puns, including the well known “grass-mud horse,” were designed to engage in a satirical online movement against internet censorship of vulgar or politically sensitive words. By examining online subversive puns’ birth and development, this paper presents a phonological analysis of the growing Chinese internet lexicon. First, the relevant phonological features of the puns are identified, which underscore how the game plays with the inherent characteristics of Mandarin. Next, a series of rules or constraints are identified; these highlight both the formulaic nature of subversive puns as well as the flexibility of the language. Finally, using Optimality Theory as a descriptive tool, this paper explores the interaction of universal constraints with several possible new language game constraints. Through this examination, this paper identifies implications for Mandarin lexical access and Mandarin word form encoding.

1. Introduction

The Chinese language has a rich history of word play. Language game research dates back to Chao’s (1931) preliminary study in which he outlined a series of 反切语 *fanqieyu* ‘secret languages’ that made use of the syllable onset and fixed rime spelling system. Branner (2010) has suggested that these games are, in fact, rooted in an even earlier military *fanqie* cipher dating back to the 16th century. Furthermore, these Chinese language games are not restricted to one ‘regional dialect’ or 方言 *fangyan*. In addition to the secret languages and games Chao cites, research into Taiwanese (Li 1985), Hakka (Branner 2010), Shanxi dialect (Hou 1988), and Cantonese (Bolton and Hutton 1995) has underscored the ubiquitous creativity and metaphor inherent in speakers throughout China.

Language games and secret languages are by no means restricted to Chinese. Laycock (1972) first coined the term ludling by combining the Latin word for ‘game’

¹Several people have shared their time and suggestions in order to improve this paper. Rebecca Morley was especially helpful with regards to the OT framework. Heather Inwood was equally invaluable, reviewing earlier drafts and offering encouraging feedback on the Chinese internet. Any errors that remain are those of the author.

ludus with the word for ‘language’ *lingua*. Davis (1993:1980) defines this concept as “a widespread language play phenomenon in which phonological forms of words are systematically altered so as to disguise what they are.”

According to Bagemihl (1995), ludlings are found in nearly every human language. Yet, due to their small speaker population and restricted linguistic function, linguists were originally hesitant to use ludlings for insight into phonological processes. It was only when ludling data from English Pig Latin was used by Chomsky and Halle (1968) to argue for the necessity of rule ordering that ludlings began to gain legitimacy as a linguistic tool.

With the rise of nonlinear phonology and morphology, ludling studies were finally recognized as a worthy area of theoretical investigation (Bagemihl 1995); phonological theories could use ludling data as supporting evidence (Bagemihl 1987; Vago 1985). Furthermore, it was argued that using ludlings as a linguistic tool could help reveal phonological traits not otherwise accessible (Pierrehumbert and Nair 1995; Treiman 1983; Yip 1982).

While ludling data lacks the same significance that natural spoken data carries, there is a growing number of studies which show that ludling data can corroborate phonological theories. Recent research has suggested that ludlings may help uncover not only the *covert ranking* of constraints – rankings that play no role in the spoken language but which emerge via ludlings, loanwords and second-language acquisition (Davidson et al 2004) – but also the sonority and perceptibility of segments (Moreton et al 2006).

This paper connects the two aforementioned threads by first introducing the newest Chinese ludling – an online game designed to circumvent internet censorship – and then examining this data via a constraint-based approach. In doing so, this paper identifies implications for Mandarin lexical access and Mandarin word form encoding, and proves their importance to the teaching and learning of Mandarin as a second language. The remainder of this paper briefly introduces the online Chinese linguistic parameters, outlines the new internet ludling, and proposes the phonological constraints along with a proper framework by which to examine them. The final two sections examine the data through this constraint-based framework, concluding with the study’s implications, limitations and future direction.

2. Chinese internet regulations

As the number of Chinese netizens has increased, so too has the number of internet regulations and online censorship. Internet regulations started in China in 1993 with the passing of the “Temporary Regulation for the Management of Computer Information Network International Connection” at the 42nd Standing Convention of the State Council (Qiu 2000). In 1997 China had approximately 25 direct international network lines (Coale 1996). Shortly thereafter, China imposed what *Wired* called “the world’s largest firewall... the Great Firewall of China” (McKay 1998).

MacKinnon points to late 2004 and early 2005 as the beginning of a spike in Chinese internet use and blogging (2008). In response, online service providers began to implement regulated censorship software tools and business models that acquiesced to the government's demands. By 2005 the Chinese government decreed that anyone hosting a "non-commercial website" had until June 30th 2005 to obtain an official registration number to be displayed on their website (French 2005; OpenNet Initiative 2005). By fall of 2005 a much broader set of regulations governing any "Internet News Information Service" website was issued (MacKinnon 2008). This forbade any site from posting content which "violated the basic principles as they are confirmed in the Constitution... jeopardized the security of the nation... divulged state secrets... subverted the national regime or jeopardized the integrity of the nation's unity"; similarly, sites were forbidden from "inciting illegal assemblies, associations, marches, demonstrations or gatherings that disturb social order... or any other content prohibited by law or rules" (Hu 2002).

In 2005 this herculean task of monitoring and censoring Chinese web pages was outsourced to individual businesses (including foreign companies), implying that the Chinese government was ambivalent with respect to how these regulations were implemented, so long as the end result was successful (MacKinnon 2008). Most companies began to use blocking technologies such as IP address-based packet filtering, DNS poisoning and cache filtering (Zittrain & Edelman 2003). Some blog-hosting companies used keyword monitoring and filtering software which drew from lists of forbidden words (Qiang 2004). According to the *Washington Post*, one such list of forbidden words contained 236 items – the majority of which was political – including 18 words which were considered obscenities (Pan 2006). In 2011, the Chinese government established the State Internet Information Office, a new central agency designed to supervise and oversee online activity (Wines 2011).

In response to this prevalent online censorship, Chinese netizens devised their own new internet secret language. A ludling this paper calls *subversive puns*. These new lexical items were invented as a humorous way around the draconian censorship of Chinese internet forums, blogs and BBS's (electronic bulletin board systems). Seemingly innocuous homophones took the place of Chinese words which could not be typed due to their vulgar or politically sensitive nature. Over time these words have spread as online memes, gaining momentum in the form of online protests against the nature of censorship and the absurdity of online keyword filters. As more netizens take part in this ludling, Esarey and Qiang (2011) see this online word play as not just a game, but a form of "digital resistance."

3. Subversive puns

In 2006 China's President Hu Jintao formally called for the creation of a "harmonious society." This vision of China in the twenty-first century signaled a policy shift away from all-out economic growth and toward a policy designed to fix many of the

increasing social tensions in Chinese society (Fan 2006). Aside from fostering more democratic and financial opportunities for the Chinese people, the proposal called for a return to morality and social correctness (Geis and Holt 2009).

The reach of Hu's "harmonious society" quickly extended to the internet where it was used as the impetus for broad censorship. Any web content which was deemed inappropriate was now considered to go against "constructing a harmonious society" and thus the content was "harmonized" (Qiang 2007). The Chinese word, as seen in (1), is made up of two morphemes and typically glossed as 'peace' or 'harmony.'

(1) 和谐 hé xié 'harmonious'

This notion of "harmonizing" the internet may be considered the watershed moment in netizen attitudes towards online censorship. Before this 2007 campaign (leading up to the Beijing Olympics in 2008), very little online resistance to censorship had been documented. As a result of the intensified censorship, Chinese netizens responded with their own campaign – a subtle and nuanced response which took the form of an animal.

(2) 河蟹 hé xiè 'river crab'

The "river crab" in (2) is a understated play on the characters for 'harmonious.' This nearly identical homophone (which is comprised of wholly different characters) quickly became the new satirical slogan for what Chinese netizens called "River Crab Society" (Qiang 2007).

This clever invention of similarly sounding, albeit entirely different characters to take the place of banned words, was the beginning of a larger online movement that manifested itself in the form of online subversive puns. The so-called "river crab" meme can be thought of as the first major subversive pun invented strictly for online use in China. These satirical puns are used in exactly the same way the original banned word was used. (3) and (4) show how "river crab" can be used as a proxy for "harmonious" in the verb form.

(3) 网页被和谐了 wǎng yè bèi hé xié le
'the webpage has been harmonized'

(4) 网页被河蟹了 wǎng yè bèi hé xiè le
'the webpage has been river-crabbed'

Despite the seemingly nonsensical sounding English gloss, (4) is a common phrase on the internet. A cursory search for (4) yielded 1,630,000 web pages containing that exact phrase.²

Using this framework of building nearly homophonous words with different characters, Chinese netizens developed a fully flexible lexicon of subversive puns including nouns and verbs aimed at satirizing the absurdity of the keyword filters and protesting the exacerbation of online censorship. In 2009 the Chinese web portal Baidu became home to the most popular collection of these subversive puns. The Chinese

²Search done on <http://www.baidu.com>, 5/26/11

language collaborative web-based encyclopedia Baidu Baike (analogous to Wikipedia) was anonymously edited by users who began posting what they called “the ten mythical creatures” (Wines 2009). This spawned an internet movement which tied the new satirical lexical items to Baidu Baike and thus gave the impression that the words were real (and thus anchored in reality outside of the internet).

Recently this ludling has gained increasing academic attention. In 2010 the China Digital Times, a self-proclaimed “bilingual news website covering China’s social and political transition” run by a team at the University of California Berkeley, established the “Grass-Mud Horse Lexicon” (GMHL 2010). Additionally, Qiang (2008) has written extensively on the growing importance of these puns as part of Chinese netizen’s online voice and role in politics.

In short, these subversive puns are the next ludling in a long history of clever, playful utilizations of the Chinese language. By embracing modern technology, these puns are the logical heir to what began as *fanqie* games and now continues as online satire. These puns, however, are not haphazardly constructed but rather follow a set of rules or constraints which will be examined in detail in the next section.

4. The innovative lexicon and its phonological constraints

It is important to recognize the role the written language plays on the internet. Wilbur points out that despite the increasing prevalence of video and audio clips, it is still predominately “a text-based affair” (1996). For this reason the features of a language’s writing system such as the use of capitalization, spelling, punctuation and style (bold, italics, etc.) are emphasized more, whereas in a spoken language, phonetic features such as voice quality, vocal register and voice modality demonstrate a speaker’s individuality (Crystal 2001). In text-based environments these features are eliminated as the orthography represents a user’s voice. Presumably most online users do not read aloud, thus the pronunciation of a word is relegated to a lower tier online than it is offline. As a result the internet has placed a much stronger emphasis on the written form than the spoken form.

Mandarin is a phonologically unmarked language with a highly limited syllable structure CVX in which C is the onset, V is nucleus and X is the coda (Duanmu 2000). Considering the numerous restrictions placed on syllable formation – DeFrancis (1984) calculated 398 basic syllables the language could form – pitch contour is one way to differentiate syllables and provide speakers with a larger syllable pool. Spoken Mandarin differs from non-tonal languages like English in that any given syllable with a pitch contour can be produced in far more ways than its equal without a pitch contour. The text-based internet, however, evens this playing field and renders tonal languages toneless: chat rooms, blogs and email all reduce languages to their written form. And yet, this reduction of acoustic clues is what makes Mandarin so ripe for pun creation.

Consider the English word “see.” Written in this form it has only one meaning, but if we disassociate the orthography from the phonology [si], there are three possible

interpretations: the act of visual recognition, the ocean, or the third letter of the alphabet. Phonologically English is much more marked than Mandarin; spoken English allows up to three consonant clusters in the onset and up to four in the coda (Ladefoged 2001). This in turn is one of the reasons English has fewer homophones than Mandarin.

Returning to written Chinese, consider the character “课.” Written in this form it too has only one meaning (a class or lesson), but if we disassociate the orthographic from the pronunciation of the character [kǜ] and keep the same falling lexical tone – kè – then there are fifteen other possible interpretations. If we include the three other possible suprasegmental tones, suddenly the landscape on which subversive puns are painted becomes immense.

It is important, therefore, to note that the written forms of subversive puns are somewhat innocuous; only the phonological properties of the syllable are of value. It is when the words are said aloud that they resemble the target banned word. The characters are not entirely arbitrary since the proxy word is designed to be used online and therefore must maintain a modicum of meaning. To illustrate how this works, consider the banned obscenity in (5).

(5) 禽你妈 cào nǐ mā ‘fuck your mother’

This fairly common profanity is censored throughout the Chinese internet. As a result the most prolific and well known subversive pun, the namesake of the China Digital Times “Grass-Mud Horse Lexicon” (6), was created.

(6) 草泥马 cǎo ní mǎ ‘grass mud horse’

The “grass-mud horse” was one of the original mythical creatures to appear on Baidu Baike and easily the most popular. Videos, cartoons, songs, blog postings, mocumentary films and merchandise of this animal have appeared throughout the internet in both China and the West (Wu 2009).

On the surface, three rules appear to dictate the formation of this and all other similar lexical items. First, the orthographic representation must change. Crucially this is done to avoid keystroke monitoring and censorship. Thus the ludling at its core is a game designed to change the written form of banned words.

Second, the syllable must be preserved. Ostensibly this ensures phonological activation and without the preservation of syllable, the ludling would not work.

Third, the suprasegmental tone should be maintained, if possible, but can be modified. In (2) one of the two lexical tones is persevered, whereas in (6) all three original lexical tones are lost.

Given these three rules or constraints, how is an optimal form chosen? If one considers the combinatorics of the three syllables in (5), the math becomes fairly daunting³. The first syllable ‘cao’ orthographically has twelve discrete candidates - three first tone, six second tone, two third tone, and two fourth tone. The second syllable ‘ni’ has twenty-six orthographic candidates - one first tone, eleven second tone, four third

³Syllable count taken from online MDBG dictionary: <http://www.mdbg.net/chindict/>

tone, and ten fourth tone. The third syllable ‘ma’ has twenty-two orthographic candidates - two neutral tone, six first tone, five second tone, six third tone, and three fourth tone. Accounting for all possible orthographic representations, a total of 6,864 possible candidates are generated. If one takes into account characters that can be read with more than one tone (多音字 *duoyinzi*), the number of possible candidates reaches 7,436. Is it possible that all candidates were considered? How is it that those who know the rules of this ludling all seemingly converged on the same output?

Before reexamining the features of subversive puns, it is worth considering the architecture at work here. The selection of (2) and (6) out of the literally hundreds and thousands of candidates suggests that this is a comparative action: candidates are presented in some framework in which an optimal output is selected. Furthermore, whatever rules or constraints exist, they are clearly violable.

Given the violable nature of the rules or constraints and the comparative action of the candidates, it follows that a constraint-based approach like Optimality Theory (OT) may be an appropriate framework (Prince and Smolensky 1993). This Chinese ludling can be thought to start with the banned written form (the “input” in OT). Next, “GEN” generates all the possible candidates; similarly this ludling generates possible subversive puns in accordance with the constraints of the game (“EVAL”). In the end, an optimal “output” or winner is presented, which in turn becomes the new subversive pun.

Using a constraint-based approach allows for a clear delineation of each constraint and highlights how each constraint interacts with one another. The present analysis is not a true extension of OT; these are not universal constraints that OT would recognize. Rather, these are only universal in the sense that native, literate Mandarin speakers who understand the ludling possess them. By using OT as a descriptive tool, and utilizing the strictly ranked constraint aspect of OT, the phonological constraints of subversive puns can be analyzed. This analysis can be thought of as either a re-ranking specifically for the ludling or a covert ranking emerging from the specific subversive pun constraints (Davidson, Smolensky, & Jusczyk 2004).

Given the aforementioned proposed surface rules or constraints, it is clear that additional new constraints must be proposed. First a semantic anti-faithfulness constraint *SEM is required. This is the “game constraint” – ostensibly subversive puns at their linguistic core involve a strict domination of *SEM over all other constraints, which in turn alters the orthographic representation of the banned word.

Next a lower ranked corollary semantic faithfulness “game constraint” ANI (animal) is required. This constraint reflects the precedent that animals (river crab, grass-mud horse, etc.) are the medium of choice for this ludling.

Two phrase structure constraints are required. The head=noun and modifier=adj constraints ensure that the output will (at the minimum) be a noun (animal) with each additional syllable serving as a modifying adjective.

Additionally, the following analysis makes use of two faithfulness Ident constraints: Ident(seg) which requires input and output segments to be identical and

谷 /ku3/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
谷 [ku3]	*!							
[u]		*!						
[kun]			*!					

(9) Summary tableau one: 谷 /kǔ/ → 古 [kǔ]

Tableau (9) highlights the semantic anti-faithfulness constraint's dominance over all the lower ranked constraints. In (9) *SEM is required to eliminate the input form and effectively produce an output with a different semantic (or orthographic) value. As previously mentioned, this constraint serves as the “game constraint” which creates a new subversive pun. Both MAX and DEP are also violated in (9) showing how the loss (MAX) or addition (DEP) of a phoneme can not occur in this language game. Thus these output forms (though illegal syllables in standard Mandarin) are eliminated.

谷 /ku3/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
笔 [pi3]				*!				
股 [ku3]					*!			
骨 [ku3]					*!			

(10) Summary tableau two: 谷 /kǔ/ → 古 [kǔ]

Tableau (10) shows the need for Ident(seg) which ensures that the segmental input matches the segmental output. (10) also illustrates how the phrase structure constraint Mod=adj is required to eliminate outputs like 股 /kǔ/ ‘portion’ and 骨 /kǔ/ ‘bone’ which are both nouns.

WIENER: SUBVERSIVE PUNS

谷 /ku3/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
古 [ku3]						*		*
孤 [ku1]						*	*!	
固 [ku4]						*	*!	

(11) Summary tableau three: 谷 /kǔ/ → 古 [kǔ]

Tableau (11) shows the winning output which, along with the other outputs in (11) violates another phrase structure constraint, Head=n. Given that all three outputs violate the same constraint, the lower ranking constraint, Ident(tn) is considered. Only the winning output maintains the suprasegmental tone and consequently wins by default. Tableau (12) shows the first morpheme's full output.

谷 /ku3/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
古 [ku3]						*		*
谷 [ku3]	*!							
[u]		*!						
[kun]			*!					
笔 [pi3]				*!				
股 [ku3]					*!			
骨 [ku3]					*!			
孤 [ku1]						*	*!	
固 [ku4]						*	*!	

(12) Summary tableau four: 谷 /kǔ/ → 古 [kǔ]

The second morpheme repeats the same comparative action as the first morpheme. Tableau (13) shows the “game constraint” at work.

WIENER: SUBVERSIVE PUNS

歌 /kɿ1/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
歌 [kɿ1]	*!							
[ɿ]		*!						
[kɿn]			*!					

(13) Summary tableau five: 歌 /kɿ/ → 鸽 [kɿ]

Tableau (13) effectively eliminates the output from having the same input. MAX and DEP limit the output to segments identical to the input.

歌 /kɿ1/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
笔 [pi3]				*!				
脍 [kɿ1]					*!			
盖 [kɿ3]					*!			

(14) Summary tableau six: 歌 /kɿ/ → 鸽 [kɿ]

Tableau (14) again shows the need for the Ident(seg) constraint. Additionally both 脍 /kɿ/ ‘armit’ and 盖 /kɿ/ ‘family name Ge’ are nouns and therefore they are eliminated in this tableau (in a later tableau it will become apparent that these outputs suffer different fatal violations).

WIENER: SUBVERSIVE PUNS

歌/kɿ1/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
☞ 鸽 [kɿ1]					*			
哥 [kɿ1]					*			*!
蛤 [kɿ2]					*		*!	

(15) Summary tableau seven: 歌 /kɿ/ → 鸽 [kɿ]

Tableau (15) shows the winning output by virtue of the fewest violations. A few points must be made here. First, as mentioned with tableau (14), multiple outputs violate the phrase structure constraint, Mod=adj. As such, later constraints are fatal. In the case of 蛤 /kɿ/ ‘clam’ it is an Ident(tone) violation. In the case of 哥 /kɿ/ ‘elder brother’, however, something more interesting is happening. Here the lower ranked corollary “game constraint” ANI comes into play. Because a precedent was set by which subversive puns involve animals, netizens converged on ‘ancient dove’ rather than ‘ancient elder brother.’ Although both ‘dove’ and ‘elder brother’ perfectly maintain the segmental and suprasegmental features of the input, it appears that the ‘dove’ is merely following the tradition established by the river crab and grass-mud horse. Of course, the additional metaphor of Google redirecting its searches to its Hong Kong server much like a bird flying south may have played a role in the formation of (15). Tableau (16) summarizes the second morpheme (note the different fatal violations).

歌 /kɿ1/	*SEM	MAX	DEP	ID(seg)	Mod=aj	Head=n	ID(tn)	ANI
☞ 鸽 [kɿ1]					*			
歌[kɿ1]	*!							
[ɿ]		*!						
[kɿn]			*!					
笔[pi3]				*!				
脰[kɿ1]					*			*!
哥[kɿ1]					*			*!
蛤[kɿ2]					*		*!	
盖[kɿ3]					*		*!	

(16) Summary tableau eight: 歌 /kɿ/ → 鸽 [kɿ]

6. Implications, limitations and future directions

As the preceding analysis has shown, the creation of subversive puns reflects a comparative action constrained by specific rules to the game. Using the ranked constraint framework of Optimality Theory, a set of constraints was identified which then was ranked to produce the optimal output. This proposal has two important implications concerning the Mandarin lexicon.

First, as both the fixed hierarchy and phonological form of subversive puns have shown, tone is of secondary importance. Given the strict domination of Ident(seg) >> Ident(tone), it appears that the segment alone is enough for lexical activation. This finding supports Chen et al's (2002) study on word-form encoding of Mandarin, which showed that the syllable is the meaningful unit of information and that tone behaves more like metrical stress. The present framework continues to strengthen this claim by showing that reading a character seemingly activates phonological competitors with both identical segmental and suprasegmental information as well as those with only identical segmental information (but with different suprasegmental information).

If it is the case that the Mandarin lexicon is organized in such a way that the syllable is the useful unit for lexical storage, how should lexical tone be viewed? On one hand, as examples such as (2) and (6) have shown, tone is not required for native Mandarin speakers to activate the target (banned) word. However, on the other hand it is clearly the case that tone matters since Mandarin L2 speakers are continually misunderstood when first producing different tones. Is it the case that tone plays a different role between spoken lexical activation and written lexical activation? This is an empirical question that requires additional research, but the present findings certainly seem to suggest that the role of tone can be augmented or diminished given the language medium. Since text already presents the reader with a concrete segment with lexical tone, the reader may have a faster route to alternate tonal competitors as compared to a listener attempting to construct both the segmental and suprasegmental tiers. Related to this question is the role tonal competitors play in lexical activation; is it the case that within lexical activation dissimilar lexical tones are considered?

Second, the present study hints at an additional way in which the Chinese lexicon may be organized: concreteness versus non-concreteness. The proposed constraints seem to corroborate much the work done on processing concrete versus abstract words (Kiehl et al 1999). Native Mandarin speakers' ability to identify semantically relevant morphemes (in the case of subversive puns - animals) suggests that the lexicon may be organized in terms of concreteness. Nearly all of the subversive puns within the Grass-Mud Horse Lexicon are lexical items that can easily be pictured due to their concreteness (GMHL 2011). This may reflect the online tendency to design subversive puns that are animate or can be personified and thus drawn, made into cartoons, videos or webcomics as was the case with the grass-mud horse.

As the theoretical framework has shown, creating subversive puns is not an unsystematic process of choosing characters to make what at first may seem like drivel.

This framework highlights the means by which the Chinese writing system is reduced from its vast inventory to a more workable pool of characters. In doing so, a subversive pun is able to effectively maximize its playfulness and similarity to the banned word, while minimizing confusion and abstraction. Despite the present study's proposal, there are several limitations that must be addressed.

The constraint-based framework presented is a synchronic analysis based on the available data at the present time. This ludling is constantly evolving to reflect the political environment and current linguistic parameters of the internet. New subversive puns may challenge the proposed constraint ranking as the Chinese internet changes and censorship waxes and wanes. Furthermore, it may be the case that animals stop being preferred or that segments may be violated. Given the present data, neither ANI nor Ident(seg) has been violated, but it may be that a subversive pun with these traits has simply not yet emerged. For these reasons the relative ranking of some constraints such as ANI remains to be seen.

Additionally, the ludling examined in the present study is primarily a written ludling. A few words need to be said about the role the writing system plays in Mandarin. As DeFrancis (1989) has repeatedly shown, speech comes first; writing is secondary. This important point is often lost on scholars who place too much emphasis on the writing system. After all, subversive puns only work when the new lexical item is read and the phonology triggers the banned or censored lexical item. In this sense, the orthographic representation is merely a conduit to the phonology; the written form carries little importance.

Furthermore, as Mair (2011) has pointed out, it is often the case that the Chinese writing system further constrains and restricts Mandarin phonology. It is important to recognize the role orthography plays in this ludling but to not over exaggerate its significance. Could the ludling be played using pinyin? Arguably, yes, it could. The ludling is built upon the inherent homophony within Mandarin and that is a result of the phonology, not the orthography.

As the present study has shown, ludlings like Chinese subversive puns can play a key role in understanding a language's phonology, lexicon and word-form encoding. As this ludling continues to grow and new lexical items are invented, future studies will want to explore precisely how flexible these new innovations can be. Is it the case that segments can violate MAX or DEP? Is there a sense of how dissimilar a subversive pun can be from the targeted banned phrase? Is it possible to quantify segmental and suprasegmental units in a meaningful way to explain the activation of phonological competitors? In order to address these questions, future studies may want to design new subversive puns and test them on native speakers.

Additionally, using subversive puns and the Grass-Mud Horse Lexicon as a pedagogical tool for L2 learners could be very productive. Subversive puns incorporate both important cultural phenomena L2 learners may not be aware of, as well as advanced homophonic lexical knowledge that most L2 speakers lack. By integrating subversive

puns into the classroom, language teachers will be able to simultaneously teach L2 learners modern, humorous linguistic innovations while advancing their lexical knowledge of phonological and tonal competitors.

It is hoped that by outlining the background, genesis and constraints of subversive puns along with a framework by which to analyze them, future research can be carried out on their importance within Chinese linguistics. Subversive puns' popularity and growing academic legitimacy is well documented; as millions and millions of netizens can attest, this is not a temporary fad. Subversive puns are rapidly becoming a significant part of Chinese culture and language.

REFERENCES

- BAGEMIHL, BRUCE. 1987. Tigrinya Speech Disguise and Constraints on Spreading Rules'. *Proceedings of the West Coast Conference on Formal Linguistics* 6.1-15, Stanford, CA: Stanford Linguistics Association.
- BAGEMIHL, BRUCE. 1995. Language Games and Related Areas. *The Handbook of Phonological Theory*, ed. by John Goldsmith, 697-712. Cambridge: Blackwell.
- BOLTON, KINGSLEY, and CHRISTOPHER HUTTON. 1995. Bad and Banned Language- Triad Secret Societies, the Censorship of the Cantonese Vernacular, and Colonial Language Policy in Hong Kong. *Language in Society* 24(2).159-186.
- BRANNER, DAVID PRAGER. 2010. Motivation and Nonsense in Chinese Secret Languages. *Luó Jiérui xiānsheng qīzhì jìn sān shòuqìng lùnwénjí* 羅杰瑞先生七秩晉三壽慶論文集, ed. by Anne Yue-Hashimoto and W. South Coblin, 469-516. Hong Kong: University of Hong Kong Press.
- CHAO, YUEN REN. 1931. Fānqiēyǔ bāzhǒng 反切語八種. *Bulletin of the National Research Institute of History and Philology* 2(3).320-354. Academia Sinica.
- CHEN, JENN-YEU.; CHEN, TRAIN-MIN.; and DELL, GARY S. 2002. Word-form encoding in Mandarin Chinese as assessed by implicit priming task. *Journal of Memory and Languages* 46.751-781.
- CHOMSKY, NOAM and HALLE, MORRIS. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- COALE, KRISTI. 1996. China Shines Intranet Contract on Sun. *Wired* 4(12).
- CRYSTAL, DAVID. 2001. *Language and the Internet*. Cambridge, UK: Cambridge University Press.
- DAVIDSON, LISA; SMOLENSKY, PAUL; and JUSCZYK, PETER W. 2004. The initial and final states: Theoretical implications and experimental explorations of richness of the base. *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater and Wim Zonneveld, 321-368. Cambridge: Cambridge University Press.
- DAVIS, STUART. 1993. Language Games. *The Encyclopedia of Languages and Linguistics* 1980-1985. Oxford and New York: Pergamon Press.

- DEFRANCIS, JOHN. 1984. *Chinese Language Fact or Fantasy*. Honolulu: University of Hawaii Press.
- DUANMU, SAN. 2000. *The Phonology of Standard Chinese*. Oxford University Press.
- ESAREY, ASHELEY, and QIANG XIAO. 2011. Digital communication and political change in China. *International Journal of Communication* 5.298-319.
- FAN, MAUREEN. 2006. China's Party Leadership Declares New Priority: 'Harmonious Society.' *The Washington Post*. 12 October 2006.
- FRENCH, HOWARD. 2005. China tightens restrictions on bloggers and web owners. *The New York Times*. 8 June 2005.
- GEIS, JOHN, and HOLT, BLAINE. 2009. Harmonious Society: Rise of the New China. *Strategic Studies Quarterly* 3(4).75-94.
- GMHL (Grass Mud-Horse Lexicon). 2011. 28 May 2011. Online: http://chinadigitaltimes.net/w/index.php?title=GrassMud_Horse_Lexicon&oldid=1396.
- HE, KEKANG, and LI, DAKUI. 1987. *Xiandai hanyu san qian changyong ci biao* 现代汉语三千常用词表. Beijing: Beijing Shifan Daxue Chubanshe.
- HELFT, MIGUEL, and BARBOZA, DAVID. 2010. Google shuts China site in dispute over censorship. *New York Times*. 23 March 2011.
- HOU, JINGYI. 1988. Shānxī lǐfǎ shèqún hánguà de yánjiù bàogào 山西理髮社群行話的研究報告. *Zhōngguó Yǔwen* 中國語文 203(2).103-112.
- HU, JIM. 2002. Yahoo! Yields to Chinese Web Laws. *ZDNet*. 20 January 2010. Online: <http://news.zdnet.co.uk/internet/0,39020369,2120830,00.htm>
- KIEHL, KENT A.; PETER F. LIDDLE; ANDRA M. SMITH; ADRIANNA MENDREK; BRUCE B. FORSTER; and ROBERT D. HARE. 1999. Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping* 7.225-233.
- LADEFOGED, PETER. 2001. *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell.
- LAYCOCK DON. 1972. Towards a Typology of Ludlings, or Play-Languages. *Linguistic Communications: Working Papers of the Linguistic Society of Australia* 6.61-113.
- LI, PAUL JEN-KUEI. 1985. A Secret Language in Taiwan. *Journal of Chinese Linguistics* 13.91-127.
- MACKINNON, REBECCA. 2008. Flatter World and Thicker Walls? Blogs, Censorship and Civic Discourse. *Public Choice* 134.31-46.
- MAIR, VICTOR. 2011. Russian Loans in Northeast and Northwest Mandarin: The Power of Script to Influence Pronunciation. *Language Log*. 20 May 2011. Online: <http://languagelog.ldc.upenn.edu/nll/?p=2921>
- MCKAY, NIALL. 1998. China: the Great Firewall. *Wired* 6(12).
- MORETON, ELLIOTT; GARY FENG; and JENNIFER SMITH. 2006. Syllabification, sonority, and perception: new evidence from a language game. *Rutgers Optimality Archive* 829-0506.

- OPENNET INITIATIVE. 2005. Internet filtering in China in 2004–2005: A country study. 3 March 2010. Online: <http://opennet.net/research/profiles/china>
- PAN, PHILLIP. 2006. Keywords used to filter Web content. *The Washington Post*. 18 February 2006.
- PIERREHUMBERT, JANET, and RAMI, NAIR. 1995. Word Games and Syllable Structure, *Language and Speech* 38.78-116.
- PRINCE, ALAN, and SMOLENSKY, PAUL. 1993. *Optimality Theory: constraint interaction in generative grammar*. New Brunswick, NJ: Rutgers University
- QIANG, XIAO. 2004. A List of Censored Words in Chinese Cyberspace. *China Digital Times*. 8 February 2010. Online: http://chinadigitaltimes.net/2004/08/the_words_you_n.php
- QIANG, XIAO. 2007. Under the Internet Police's Radar. *China Digital Times*. 13 February 2010. Online: <http://chinadigitaltimes.net/2007/08/under-the-internet-polices-radar>
- QIU, JACK LINCHUAN. 2000. Virtual Censorship in China: Keeping the Gate Between the Cyberspaces. *International Journal of Communications Law and Policy* 4.1-25.
- TREIMAN, REBECCA. 1983. The structure of spoken syllables: evidence from novel word games. *Cognition*. 15.49-74.
- VAGO, ROBERT M. 1985. The treatment of long vowels in word games, *Phonology Yearbook* 2:329-342
- WILBUR, SHAWN. 1996. An archaeology of cyberspaces: virtuality, community, identity, *Internet Culture*, ed. by David Porter, 5-22. New York and London: Routledge.
- WINES, MICHAEL. 2009. A Dirty Pun Tweaks China's Online Censors. *New York Times*. 12 March 2009.
- WINES, MICHAEL. 2011. China creates new agency for patrolling the internet. *New York Times*. 4 May 2011.
- WU, VIVIAN. 2009. Censors strike at internet content after parody hit. *South China Morning Post* 3 April 2009.
- YIP, MORIA. 1982. Reduplication and the C-V Skeleta in Chinese Secret Languages. *Linguistic Inquiry* 13.637-661.
- ZHANG, JIE. 2001. *The Effects of Duration and Sonority on Contour Tone Distribution--Typological Survey and Formal Analysis*. Los Angeles, CA: UCLA Dissertation.
- ZITTRAIN, JONATHAN, and EDELMAN, BENJAMIN. 2003. Documentation of Internet Filtering Worldwide. *Spreading the Word on the Internet: reflections on freedom of the media and the Internet*, ed. by Christiane Hardy and Christian Moller, 137-148. Vienna: Organisation for Security and Cooperation in Europe.

Divergent places of articulation: [w] and [v] in modern spoken Mandarin

Seth Wiener and Ya-ting Shih¹
The Ohio State University

This sociophonetic study examines the production of the modern spoken Mandarin voiced labial-velar approximant /w/ in isolation and zero initial environments. Acoustic analyses of ten native Mandarin subjects' productions reveal that some speakers produce /w/ as both [w] and [v]. The environments in which [w] and [v] appear suggest that this variation may be allophonic, conditioned by the syllable nucleus. Furthermore, our results show evidence of regional and gender variation: northern female Mandarin speakers produce [v] in more contexts and more frequently than speakers from other regions.

1. Introduction

Throughout the world's languages approximants occupy a nebulous phonetic status. Beginning with Ladefoged's first use of the term in his *Phonetic Study of West African Languages* (1964:25), phoneticians have differed in how best to classify approximants within a phonetic inventory (Martinez-Celdran 2004). Even by IPA Handbook standards, approximants lack a well defined status reserved for seemingly more clear-cut categories such as stops and fricatives (International Phonetic Alphabet 1999). Because approximants are produced with one articulator close to another, yet without turbulent airflow, this "approximation" of manners further confounds how best to characterize the speech production, leading some phoneticians to prefer the name glides or semi-vowels (Stevens 1998).

The labial velar approximant /w/ is one of the more challenging approximants, primarily due to its co-articulation; some phonological descriptions do not classify /w/ as both a labial and velar even though both places of articulation are involved in the phonetic production (Ohala and Lorentz 1977). The problem of /w/ becomes even more complicated with respect to Standard Mandarin. Sinologists differ on everything from its place of articulation to its role within a syllable. Traditionally /w/ has been classified as part of the final (medial) position (Cheng 1973), but more recent views of the syllable structure place this approximant in the onset as a consonant-glide cluster (see Lin 2007

¹ The authors would like to thank the following individuals for their numerous helpful comments and suggestions: Marjorie Chan, Cynthia Clopper, Mary Beckman and Shawn Chang. Any errors that remain are the authors'.

and Duanmu 2007). Norman (1988) considers /w/ a labial voiced continuant, grouping it as part of the onset. Li and Thompson (1981) disagree and view /w/ not as an initial but rather as part of the vowel nucleus or the rime. Duanmu (2005) considers it a variation of one of three underlying high vowels which behaves as a glide in certain contexts.

Thus the disagreement of whether /w/ is part of the onset or rime in Chinese syllables depends on one's view of the Mandarin syllable. Yip (2003) and Wan and Jaeger (1998) have suggested that while the glide /w/ is generally considered part of the rime, secret languages, which break apart the syllable, actually treat the glide /w/ as part of the onset. The issue of classifying /w/ is actually a taxonomic assessment. As Ohala and Lorentz (1977) point out, there may be instances when one classification as a velar is appropriate, another instance when classification as a labial is appropriate and even a third instance when classification as both is the most appropriate choice. The same can be said for classification of /w/ as part of the onset or part of the rime.

Given the phonetic debate surrounding the Mandarin labial velar approximant, there is good reason to explore the phonological realization of /w/ across speakers. Previous studies have suggested that the labiodental approximant [ʋ] serves as a non-phonemic variant along with [w] (Shen 1987, Chan 1996). Shen's study examined a large group of over 400 speakers within Beijing and the surrounding areas. The findings strongly suggest that [ʋ] is more likely to be produced with a less rounded vowel nucleus and with the alveolar nasal [n] coda. Shen concluded that speakers who produced [ʋ] share three traits: they are younger, higher educated and female.

Given that over twenty years have passed since Shen's 1987 study, the present research aims to build on Shen's important findings in three ways. First, Shen relied exclusively on transcription and visual observation of the speakers' mouth shape (lip rounding). The present study utilizes acoustic analysis in order to present a more objective set of results less prone to human error. Second, Shen only analyzed spontaneous speech of eight Mandarin syllables. The present study examines additional modes of production as well as a previously omitted ninth syllable type. Finally, while Shen's study looked at over 400 speakers, all the speakers were concentrated within Beijing and its surrounding areas. The present study, while much smaller in the total number of speakers, examines speakers from throughout China.

Our study shows that despite the passing of time since Shen's (1987) study, females continue to produce [ʋ] more than males. Furthermore, the production context – reading versus spontaneous speech – does not affect the production of [ʋ]. Additionally, we find that [ʋ] is a northern regional, possible dialectal variation, which extends beyond the Beijing area. Finally our findings suggest that [ʋ] serves as a possible allophonic variant of /w/ for some speakers similar to Shen's findings but not across all of Shen's predicted vowel nuclei. This production can be thought of as a process of approximant dentalization conditioned by certain vowel contexts.

The paper is organized as follows: Section 2 highlights the relevant acoustic differences between [w] and [ʋ]; Section 3 outlines the methodology of the study; Section

4 presents the results; Section 5 offers a discussion of the findings; Section 6 concludes the paper.

2. Acoustic Characteristics

The articulation of /w/ varies depending on the following vowel (Ladefoged, 2006). In Mandarin /w/ appears only as the onset (or medial glide) in the following phonetic contexts:

Table 1. Examples of the 9 Legal Mandarin syllables with /w/

Pinyin	IPA	English Gloss	Character
wǒ	/wo/	I/me	我
wǔ	/wu/	five	五
wēng	/wəŋ/	Weng surname	翁
wèn	/wən/	to ask	问
wèi	/wei/	because of	为
wàng	/wɑŋ/	to forget	忘
wǎn	/wan/	night	晚
wài	/wai/	outside	外
wá	/wa/	doll	娃

Since approximants are often coarticulated, both Ladefoged (2006) and Stevens (1998) have characterized their formant structure as regularly changing. This change can be seen in numerous manifestations, but primarily approximants are responsible for lowering all formants. Stevens (1998:518) calculates the average F1 of male speakers producing /w/ before high vowels at 255 Hz (245 Hz for females) and at 293 Hz for males before non-high vowels (291 Hz for females). Given that approximants' formants are as difficult to quantify as approximants themselves – there are no steady states without a vowel – it is crucially the second formant which offers the most telling characteristics. Lower F2 frequencies are often indicative of approximants, especially in the case of the lip-rounding involved in the production of /w/ (Ladefoged 2006).

Figure 1 below illustrates both the malleability of the approximant /w/ in terms of its ability to resemble the subsequent vowel as well as the lowering of F1 and F2 (first and second formant line). All recordings were productions from a native Beijing Mandarin female speaker who produced /w/ as [w]. The formant measures were taken at the transitional point from onset to vowel using a Praat script (Boersma and Weenink 2011). As the formants show, /w/ significantly lowers the second formant when compared with both the bilabial plosive /p/ and the labiodental fricative /f/, which lack such lip rounding.

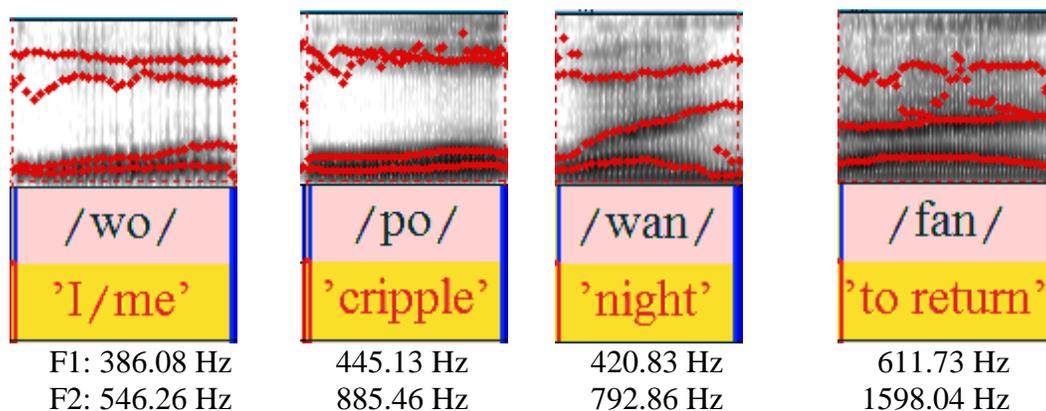


Figure 1. Spectrograms of Mandarin syllables juxtaposing the approximant onset /w/ with the bilabial plosive /p/ and the labiodental fricative /f/ (with F1 and F2 in Hz)

Furthermore, approximants are produced with a narrow vocal tract constriction. Here an important distinction must be made between the constriction of an approximant and a constriction which allows a turbulent airstream like a fricative (Martinez-Celdran 2004). As Stevens (1998:519) has pointed out, this results in a lowering of F1 frequency (as seen above in Figure 1) and reduced spectrum amplitude in F1. Bickley and Stevens (1986) and Espy-Wilson (1987) calculate a lowering of 5 to 10 dB for /w/. Figure 2 below illustrates this drop in spectral amplitude with the labial velar approximant /w/ clearly dipping in dB (left-hand side) when compared with labiodental fricative /f/ (right-hand side). This slight dip in intensity represents approximately a 5 dB difference between the approximant and the fricative from the onset transition to the vowel. Measurements were taken from production of the onset to the transition into the nucleus.

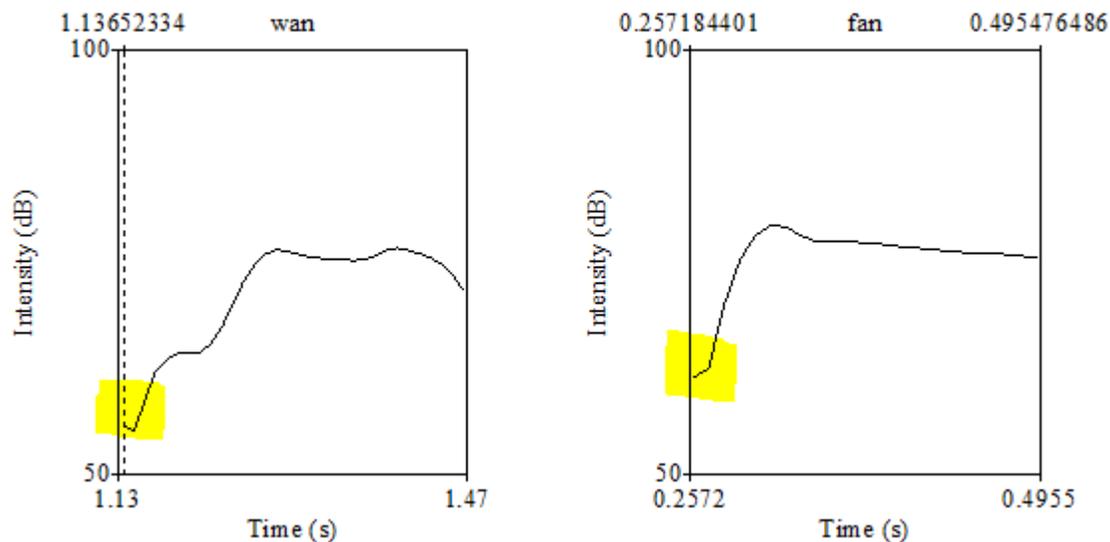


Figure 2. Spectral intensity measurements (in dB) of /wan/ from *wan* ‘evening’ on the left and /fan/ from *fan* ‘to return’ on the right

Since both [w] and [ʋ] are approximants, the lowering of spectral intensity can not alone be used as acoustic evidence to argue for one phoneme over the other. Therefore, the present study will only rely on the most salient acoustic cue: approximants with lip rounding result in F2 lowering. Dissimilarity in formant height (primarily F2) should be robust enough to suggest which approximant was produced. To test this hypothesis, the following production experiment was carried out.

3. Methodology

In order to make use of the aforementioned acoustic cue, the researchers designed a word list containing /w/ in the zero initial position. By combining all possible syllable and tone combinations, 32 target syllables were selected (see appendix for full research instrument). These targets were then placed within a fictional story. The story varied the location of the syllables across sentences; the 32 target syllables were embedded in Mandarin words and put into a paragraph for a total of 41 tokens. Nineteen follow-up content questions were created based on the paragraph. As a result, three specific speech tasks were used to elicit spontaneous speech and speech from reading. First, the subjects were asked to read aloud the paragraph containing the 41 targeted syllables. After reading the paragraph, participants were then asked to answer questions based on the paragraph. The questions were created to elicit the target syllables in a more informal and natural context. Subjects produced approximately 60 spontaneous tokens. Finally, the subjects were asked to read the 32 target syllables in isolation. The entire process took approximately ten minutes and yielded, on average, 120 tokens.

Each subject was tested individually in a quiet room. Recordings were made with a handheld digital Edirol recorder with a built-in microphone, which was placed on the desk in front of them. The sampling rate for the wav recorder was 44.1 kHz. Digital recordings were saved onto a PC and analyzed using the PC program Praat.

In total, 10 native Mandarin speakers (3 male, 7 female) were recruited from a major mid-west university in the United States. The subjects were drawn from northern, central and southern Mandarin speaking regions in China (see appendix for speaker information). The subjects were all graduate students who had been in the United States for less than four years at the time of the experiment.

In order to analyze the data, all of the recordings were first cut into three sound files based on their production contexts – paragraph, questions and wordlist. Each file was further cut into the individual tokens of interest to the study. Sounds which were unclear, produced incorrectly or too heavily influenced by the surrounding words were discarded. A total of 1,183 tokens were analyzed.

The tokens were first transcribed by a native Mandarin speaker. Each token was identified as having either an [w] or [v] initial. These transcriptions were also verified by an advanced non-native Mandarin speaker in order to ensure accuracy. The transcriptions were only used as a guide to organize the data analysis. In order to draw a more objective conclusion regarding the production of [w] or [v], each token was further annotated using Praat textgrids in order to identify the onset and nucleus. In doing so, the token was demarcated from the beginning of the /w/ onset to the transition into the nucleus, but before the coda. This approximately 0.3 second segment was used to extract the average F1 and F2 over the course of the token. A Praat script was used which performed short-term spectral analysis on approximately 40 frames. The average F1 and F2 of every token was output into a text file which was then analyzed using the statistical computing program R (Bell Laboratories 2010).

4. Results

The aforementioned linking assumption that the production of /w/, an approximant which involves lip rounding, will lower F2, was used as the primary acoustic cue to distinguish the production of [w] and [v] (Ladefoged 2005). Figure 3 below shows the clear difference between two speakers' production of the same segment /wən/ 'to ask'.



Figure 3. Spectrograms of [vən] (left) and [wən] (right) highlighting the lowering of F2.

Our results show that among the 10 speakers, only two female subjects – Subject 2 and Subject 9 – regularly produced /w/ as [v] in multiple vowel contexts. These two subjects’ productions of /w/ most often resulted in [v] as shown by the divergence in average F1 and F2 when compared with other speakers producing the same token. Two other female speakers – Subject 1 and Subject 4 – produced a restricted number of tokens as [v]; however, the majority of their production was [w].

Figure 4 shows the average F2 results of one of the [v] speaker’s tokens (Subject 2) collapsed across all three production contexts as compared with one of the [w] speaker’s tokens (Subject 8) collapsed across all three production contexts.

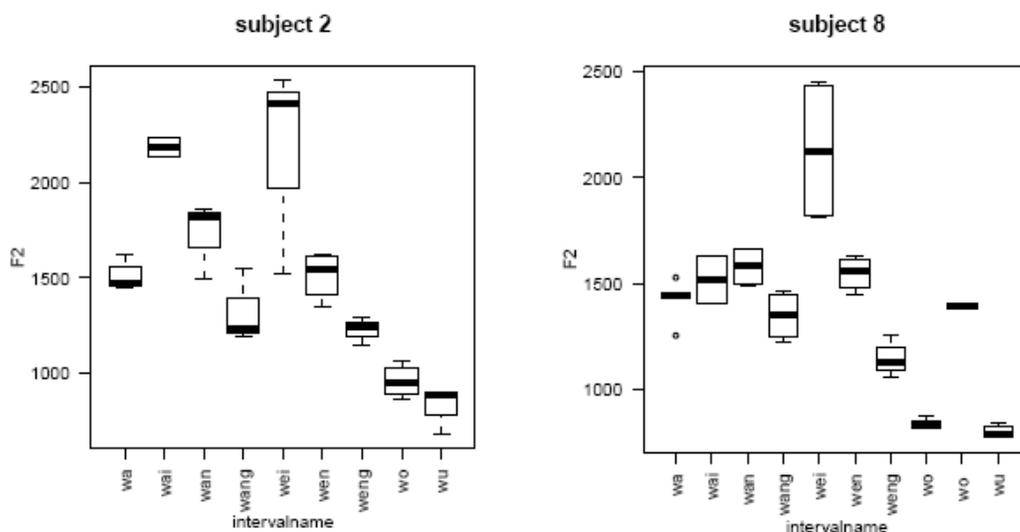


Figure 4. Average F2 (Hz) of Subject 2 and Subject 8’s production across all three contexts

Figure 4 clearly shows that Subject 8’s production generally trends towards a much lower F2, which is due to the lip rounding of [w], whereas subject two’s production shows a much higher F2 trend implying the non-lip rounded production of [v].

Two paired t-tests were run to examine the production of the two speakers (Subject 2 and Subject 9) who produced [v] as well as [w]. The results showed that the F2 of [w] was significantly different than that of the F2 of [v] ($t = 2.97$, $df = 19$, $p = .008$, $t = 3.91$, $df = 14$, $p = .002$ respectively). In order to address which phonological environments conditioned the production of [v], two chi square tests were conducted. The two chi square tests reported that vowel context did play a role in their production of [v] ($\chi^2=(8, N= 135)=57.11$, $p=.00.$, ($\chi^2=(8, N= 125)=57.1$, $p=.00$ respectively). Using a level of significance of .05, the standardized residual was calculated. Results were significantly different if they exceeded the range between -1.96 and 1.96. Thus, our findings suggest that when /w/ was followed by the vowels /ə/ and /ei/, the speakers produced /w/ as [v]. When /w/ was followed by a back rounded vowel such as /o/ and /u/, the speakers usually maintained the production as [w]. The results for the vowel /a/ were inconclusive; although our measurements indicate a divergence in mean F1 and F2 values for those speakers when producing syllables with an /a/ nucleus, our statistical analysis did not return a significant value.

Additionally we examined whether production context played a role in the variation between [w] and [v]. The Chi square tests showed that task, be it reading the paragraph, answering questions in natural speech or reading from a word list, did not play a role in the production of [v].

Finally, in order to investigate whether the production of [v] reflected a regional speech difference or a larger trend in spoken modern Mandarin, a logistic regression model was used. We used the “enter” method with the following four predictors in the model: gender, north/south region, vowel context and production context. The Hosmer-Lemeshow test for goodness of fit yielded $\chi^2(8)$ of 8.15 and was not significant ($p= .42$). As such, the model used fit the data well. Moreover, as shown in table 2, the model correctly predicted 94 percent of the production.

	Predicted		% Correct
Observed	[w]	[v]	
[w]	983	55	94.7
[v]	14	155	91.7
Overall % correct			94.3

In table 3, the statistical test of individual predictors indicated that among all four predictors, only the variable “region” was significant in predicting the production of [w] and [v] ($p=.00$): northern subjects are more likely to produce /w/ as [v]. Therefore, the

model suggests that the production of the labial dental approximant [v] is a regional difference rather than a more general trend of modern spoken Mandarin.

Table 3. Logistic Regression Analysis of 1207 tokens of the 10 subjects						
Predictor	β	S.E	Wald χ^2	df	Sig.	Exp B (odd ratio)
Constant	-46.18	3191.14	.00	1	.99	2.68E9
Gender	21.71	1822.40	.00	1	.99	117.92
Region	4.76	.35	188.20	1	.00	116.97
VowelCon			12.29	8	.14	
VowelCon(1)	20.10	2619.60	.00	1	.95	5.21E8
VowelCon(2)	20.29	2619.60	.00	1	.95	6.45E8
VowelCon(3)	19.72	2619.60	.00	1	.95	3.66E8
VowelCon(4)	19.95	2619.60	.00	1	.95	4.60E8
VowelCon(5)	20.10	2619.60	.00	1	.95	5.17E8
VowelCon(6)	21.43	2619.60	.00	1	.95	2.03E9
VowelCon(7)	19.21	2619.60	.00	1	.95	2.20E8
VowelCon(8)	-.20	4365.63	.00	1	1.00	.82
Proconx			.96	2	.61	
Proconx(1)	.24	.38	.42	1	.52	1.28
Proconx(2)	.37	.37	.99	1	.32	1.45
Test			χ^2	df	Sig.	
Overall model evaluation						
Likelihood ration test			622.01	13	.00	
Goodness-of-fit test						
Hosmer & Lemeshow			8.15	8	.42	

5. Discussion

Our results have shown that of the ten subjects tested, two consistently produced [v]. These two subjects were both female, which is unsurprising given that the labiodental [v] results in a much higher frequency than the labial [w]. As a result, the labiodental approximant may be perceived by many as feminine sounding. Our finding follows Shen's (1987) results that [v] was produced significantly more frequently by female speakers than by males. In sum, our study effectively corroborated Shen's gender-based observation nearly a quarter of a century later, suggesting that the production of [v] is still a widespread phenomenon.

Additionally, it has been observed that the production of [v] is often heard on news broadcasts where typically (although not exclusively) the labiodental approximant is produced by female reporters throughout China and Taiwan (Chan 1998). The

production of the labiodental variant by newscasters could be interpreted as a more prestigious variety of spoken Mandarin. The two subjects who produced [v] consistently were both Northern Mandarin speakers (Beijing and Inner Mongolia, respectively). Although our subject pool was restricted to only ten subjects (see appendix for additional subject information), our logistical regression findings conclude that this production is the result of regional variation and not a larger trend across Mandarin (i.e. a sociolinguistic phenomenon driven by newscasters' speech or another perceived prestigious dialect).

The most important finding of the present study is the statistically significant production of [v] when conditioned by the subsequent /ə/ and /ei/ nucleus. While /wəŋ/, /wən/, and /wei/ were not the only syllables in which [v] was produced, statistically these syllables were regularly shown to condition the labiodental approximant to be produced. These findings both support and contradict Shen's (1987) findings.

Both the present study and Shen's study concluded that /wən/ was produced as [vən] more often than any other syllable. Both studies also reported /wei/ as a syllable with a significant production of [v] tokens as [vei]. Shen, however, reported /wan/ and /wa/ as the second highest percentage of syllables containing [v]. In the present study these syllable types were not statistically significant. The difference in the present study and Shen's study may be accounted for by the considerably large discrepancy in number of subjects, the method of analysis (transcription versus acoustic analysis), or the passing of time since Shen's research. Future research will need to address this difference between the two studies.

Based on the findings from the ten speakers analyzed, we explain the production of [v] in the following ways. First, it may be the case that the production of [v] is the result of dentalization in close-mid vowel contexts. This would imply that the produced labiodental approximant is not necessarily due to an allophonic variation but rather due to a phonological process resulting in a change in place of articulation by certain speakers. This dentalization rule can be expressed as:

$$\text{Approximant dentalization rule: } /w/ \rightarrow [v] / _ \left(\begin{array}{c} \text{ə} \\ \text{ei} \end{array} \right)$$

This rule seemingly holds true for the two speakers in our study. Additionally, a rule such as this maintains that [v] is merely the surface form of /w/ and that speakers of this variety of Mandarin do not underlyingly possess /v/.

Alternatively, it may be the case that [v] is an allophone of /w/ and that speakers of this variety of Mandarin produce /w/ as both [v] and [w]. This allophonic variation was repeatedly observed during the acoustic analysis of Subject 9's sound files. Figure 5 shows an especially interesting production of /wei/ 'power'.

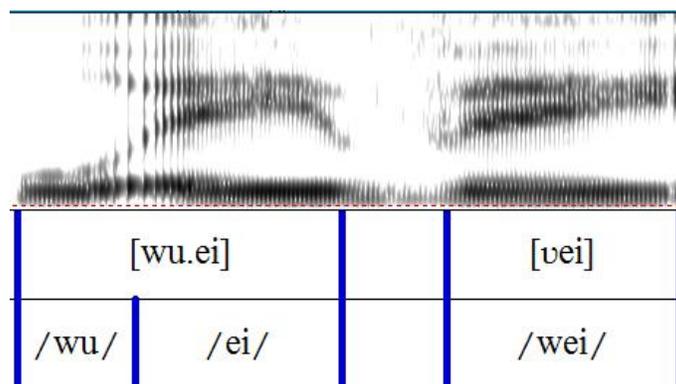


Figure 5. Production of a /wei/ syllable as [w] as the onset (with [u]), [ei] as the vowel and [vei] as the complete syllable

Figure 5 shows the speaker is able to produce the onset and nucleus as individual segments, but when combining them to form the complete syllable, the onset changes from /w/ to [v]. We see this as possible evidence for classifying [v] as an allophone of /w/. Future research will need to address this debate and begin identifying whether or not minimal pairs with [w] and [v] are produced in this variety of Mandarin and what role, if any, the perception of [v] plays. Future research may also want to examine children's production to see if this alternation is phonologized among northern speakers at a certain age.

Future studies should also explore the production of [v] in syllables with an /a/ nucleus. Though our data was statistically inconclusive, it may be that the production of [v] is spreading across other vowel contexts. Follow-up studies should explore the production of [v] across generations to see if the number of vowel contexts in which the labiodental approximant is produced varies across different age groups. Since the present study examined speakers all within the same age range, it is possible that the production of [v] is either increasing or decreasing among speakers.

Finally, future studies should continue to explore the perceived prestige of [v] and what, if any, role [v] plays among speaker's social dynamics. The production of the labiodental approximant may prove to be a salient cue of a certain variety of Mandarin in the same way that rhotacization does (Zhang 2005).

6. Conclusion

This paper has outlined some of the important features of the labial velar approximant /w/ in Standard Mandarin. Its narrow constriction and highly vowel-like formants result in a lowered first formant and a natural phonetic basis to treat it as a semi-vowel (Stevens 1998). Yet, the constriction, unlike fricatives, remains weak enough that no turbulent noise is produced (Laver 1994). The labial velar approximant /w/, especially in the onset position of Standard Mandarin, acts differently than the vowel alone. Due to

the inherent lip rounding involved in its production, a lowered second formant is visibly apparent in spectrograms.

These features were used to analyze the production of /w/ by ten native Mandarin speakers from different parts of China. It was found that the place of articulation of /w/ shifted from labial velar to labial dental [v] in four female speakers. Of those four, only two speakers – both Northern Mandarin speakers – produced [v] consistently and predictably. These findings are very much in line with the claim that the production of [v] is perceived as feminine, primarily found in Beijing Mandarin and thus produced by younger female speakers (Shen 1987). Our findings were further supported by a logistic regression model which showed that statistically a speaker's region (north) plays the most important role in the production of [v]. These findings uphold Shen's research nearly a quarter of a century later. The present study, taken in conjunction with Shen's groundbreaking work, proves that the labiodental approximant [v] is a robust phonetic production found in Northern Mandarin.

Our findings can be explained through a proposed approximant dentalization rule, which causes /w/ to become dentalized before /ə/ and /ei/. We also put forth the possibility of [v] serving as an allophone of /w/ for speakers of certain varieties of Northern Mandarin. This claim requires additional fieldwork and further phonetic studies to examine the extent of [v] production as well as its role across the speaker's language. Future studies may also want to consider the perception of [v] and explore whether or not minimal pairs exist within that variety of Mandarin.

7.1 Appendix: Research Instrument Word List

Sound in IPA	Tone of the target word	Target word	Target word in phrase	Pinyin	English gloss
[wu]	Tone 1	屋	屋子	Wuzi	House
	Tone 2	无	无穷	Wuqiong	Endless
	Tone 3	五	五月	Wuyue	May
	Tone 4	雾	雾气	Wuqi	Fog
[wa]	Tone 1	挖	挖土	Watu	Dig
	Tone 2	娃	娃娃	Wawa	Baby
	Tone 3	瓦	瓦片	Wapian	Tile
	Tone 4	袜	袜子	WaZi	Sox
[wo]	Tone 1	莴	莴苣	Woju	lettuce
	Tone 3	我	我们	Women	We/us
	Tone 4	握	握手	Woshou	Shake hands
[wan]	Tone 1	弯	弯曲	Wanqu	Winding
	Tone 2	顽	顽皮	Wanpi	Naughty
	Tone 3	晚	晚上	Wanshang	Evening
	Tone 4	惋	惋惜	Wanxi	Feel sorry for
[wai]	Tone 1	歪	歪曲	Waiqu	Misrepresent
	Tone 4	外	外面	Waimian	Outside
[wei]	Tone 1	威	威风	Weifeng	Mighty
	Tone 2	违	违规	Weigui	Violate rules
	Tone 3	纬	纬度	Weidu	Latitude
	Tone 4	味	味道	Weidao	Smell
[wən]	Tone 1	温	温暖	Wennuan	Warmth
	Tone 2	蚊	蚊子	Wenzi	Mosquito
	Tone 3	稳	稳健	Wenjian	Stability
	Tone 4	问	问题	Wenti	Question
[wɑŋ]	Tone 1	汪	汪洋	Wangyang	Boundless/vast
	Tone 2	王	王子	Wangzi	Prince
	Tone 3	网	网球	Wangqiu	Tennis
	Tone 4	忘	忘记	Wangji	Forget
[wəŋ]	Tone 1	嗡	嗡嗡	Wengweng	Humming sound
	Tone 3	蓊	蓊郁	Wengyu	Luxuriant
	Tone 4	瓮	瓮	Weng	Jar/pot

7.2. Appendix: Research Instrument Paragraph

英国的威廉王子今年五月到北欧的瑞典进行访问。瑞典因为纬度较高,因此即使是五月,天气还是很冷,不仅屋子里需要开暖气,早上外面的雾气也很浓。威廉王子魅力无穷,他穿上正式服装,看起来相当威风。沿途上很多民众争相跟他握手,他还顽皮地抱起路旁在玩挖土游戏的娃娃,他亲切的态度与他已故的母亲黛安娜王妃一样使人感觉很温暖,我们都很喜欢他。晚上,在晚宴之前,威廉王子先参加了记者会,他稳健地回答记者的问题并说明他接下来的公益活动,他同时也表达对于有一些记者报导他的私生活且歪曲事实,还说他利用贵族的身分违规停车,对此他感到很遗憾与惋惜。晚宴里,他表示餐点的味道相当好,其中还有他最喜欢的莴苣。

隔天早上,威廉王子先去打了网球,接着下午又到附近的山林走走,由于山里气温较低,他换上了较厚重的衣服和袜子以保暖。他对随行的人员说他已经忘记上一次到山里走走是什么时候了。他说山林里的小路弯曲蔓延,草木蓊郁,又有蜜蜂在耳边嗡嗡地叫,虽然偶而有蚊子,但是登上高处远眺汪洋的大海和山下的点点的屋瓦,让人心曠神怡。

7.3. Appendix: Research Instrument English translation

Prince William of the United Kingdom visited Sweden this May. Because of the high latitude of Sweden, even though it was already May, it was still cold. People there still needed a heater in the house and there was also a heavy fog outside. Prince William looks very charming, especially when he dresses up in formal attire. Many Swedes like him and they all wanted to shake hands with him. Prince William also mischievously picked up a child playing and digging on the playground. Prince William is very down-to-earth and nice. His warm attitude is just like his mother, Princess Diana. We all like him a lot. In the evening, Prince William had a press conference before the banquet. At the press conference, he answered all the questions and also talked about some charity activities he plans to do next. He also said that he felt deeply sorry that some paparazzi gave inaccurate reports of his personal life and also claimed that he violated some parking rules. At the banquet, he said that he liked the food very much and his favorite vegetable is lettuce.

The next day, Prince William played tennis in the morning and then went hiking in the afternoon. Due to the possible lower temperature in the mountains, he puts on a coat and socks to keep him warm. He told his entourage that he couldn't remember the last time he went hiking. In the mountain, there were winding trails, luxuriant forest and bees. Though, sometimes, there were mosquitoes, he felt great when he reached the top of the mountain and looked down at the sea and the small houses below.

7.4. Appendix: Research Instrument Questions

1. 威廉王子几月到瑞典访问?
2. 五月瑞典的天气怎样么? 还需要暖气吗? 早上雾气是否也很浓?
3. 瑞典为什么五月天气还是很冷?
4. 瑞典的民众喜欢威廉王子吗?
5. 在晚宴前的记者会,威廉王子表现得怎么样?
6. 威廉王子对总是报导他私生活的记者怎么样?
7. 威廉王子觉得晚宴的餐点怎么样? 他最喜欢吃什么?
8. 威廉王子隔天在登山前做了什么运动?
9. 在山里威廉王子看到哪些小昆虫?
10. 威廉王子从山顶往下看看到什么?
11. 威廉王子换上怎样的衣服去爬山?
12. 威廉王子记得上次去爬山是什么时候吗?
13. 山里的景色怎么样?
14. 威廉王子什么时候参加记者会?
15. 有很多人想跟威廉王子握手吗?
16. 威廉王子穿上正式的衣服看起来怎样?
17. 威廉王子抱起了谁?
18. 威廉王子很有魅力吗?
19. 威廉王子真的有违规停车吗?这是事实吗?

7.5. Appendix: Research Instrument English translation

1. When did Prince William visit Sweden?
2. How was the weather in Sweden in May? Did people there still need a heater? Was there heavy fog in the morning?
3. Why was it cold in Sweden in May?
4. Do Swedes like Prince William?
5. How did Prince William do in the press conference before the banquet?
6. How did Prince William feel about the paparazzi?
7. How did Prince William like the food at the banquet? What did he like the most?
8. What did Prince William do the next day before he went hiking?
9. What insects did Prince William see in the mountain?
10. What did Prince William see from the top of the mountain?
11. What did Prince William wear when he went hiking?
12. Did Prince William remember the last time he went hiking?
13. How was the scenery in the mountains?
14. When did Prince William have the press conference?
15. Did many people want to shake hands with Prince William?
16. How did Prince William look when he dressed up in formal attire?

17. What did Prince William hold in his arms?
 18. Is Prince William charming?
 19. Did Prince William violate parking rules? Was it true?

7.6. Appendix: Participants information

Subject ID	Gender	Age	Place, Province (In Chinese)	Place, Province (In English)	Categorized region in the study
1	Female	35	湖北省, 武汉,	Wuhan, Hubei,	South
2	Female	30	内蒙古自治区, 赤峰	Chifeng, Inner Mongolia	North
3	Male	27	福建省, 尤溪	Youxi, FuJian	South
4	Female	36	江苏省, 徐州	Xuzhou, Jiangsu	South
5	Female	26	江苏省, 南京	Nanjing, Jiangsu	South
6	Male	35	辽宁省, 沈阳	Shenyang, Liaoning	North
7	Female	29	湖南省, 水潭	Shuitan Hunan	South
8	Female	25	广东省, 新会	Xinhui, Guangdong	South
9	Female	30	北京市	Beijing	North
10	Male	30	陕西省, 西安	Xi'an, Shanxi	North

REFERENCES

- BELL LABORATORIES. 2010. The R project for statistical computing [Computer program]. Version 2.9.2, retrieved 10 October 2010. <http://www.r-project.org/>
- BICKLEY, CORINE; STEVENS, KEN. 1986. Effect of a vocal tract constriction on the glottal source: Experimental and modeling studies. *Journal of Phonetics* 14. 373-382.
- BOERSMA, PAUL ; WEENINK, DAVID. 2011. Praat: doing phonetics by computer [Computer program]. Version 5.2.25, retrieved 11 May 2011. <http://www.praat.org/>
- CHAN, MARJORIE K.M. 1996. Sound symbolism and the Chinese language. *Proceedings of the 7th North American Conference on Chinese Linguistics (NACCL) and the 4th International Conference on Chinese Linguistics (ICCL)*, ed. by Tsai-Fa Cheng; Yafei, Li; Hongming Zhang, 2. 17-34. Los Angeles, CA: GSIL Publications, University of Southern California.
- CHAN, MARJORIE K.M. 1998. Gender differences in the Chinese language: A preliminary report. *Proceedings of the Ninth North American Conference on Chinese Linguistics (NACCL- 9)*, ed. by Hua Lin, 2. 35-52. Los Angeles: GSIL Publications, University of California.
- CHAO, YUEN REN. 1941. "Distinctions within Ancient Chinese." *Harvard Journal of Asiatic Studies* 5. 203-233.
- CHENG, CHIN-CHUAN. 1973. *A Synchronic Phonology of Mandarin Chinese*. Berkely, CA: University of California Press.
- DUANMU, SAN. 2006. Chinese (Mandarin): phonology. *Encyclopedia of Language and Linguistics*, ed. by Keith Brown, 351-355. 2nd Edition. Oxford, UK: Elsevier Publishing House.
- DUANMU, SAN. 2007. *The Phonology of Standard Chinese*. 2nd Edition. Oxford, UK: Oxford U. Press.
- ESPY-WILSON, C. Y. 1987. *An acoustic-phonetic approach to speech recognition: Application to the semivowels*. Cambridge, MA: MIT dissertation.
- INTERNATIONAL PHONETIC ASSOCIATION (IPA). 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- LADEFOGED, PETER. 1964. *A Phonetic Study of West African Languages*. Cambridge: Cambridge University Press.
- LADEFOGED, PETER. 2005. *Vowels and consonants: An introduction to the sounds of languages*. 2nd Edition. Malden, MA: Blackwell Pub.
- LADEFOGED, PETER. 2006. *A Course in Phonetics*. 5th Edition. Boston: Thomson, Wadsworth.
- LAVER, JOHN. 1994. *Principles of Phonetics*. Cambridge, UK: Cambridge U. Press.
- LIN, YEN-HWEI. 2007. *The Sounds of Chinese*. Cambridge, UK: Cambridge U. Press.

- MARTÍNEZ-CELDRÁN, EUGENIO. 2004. Problems in the classification of approximants. *Journal of the International Phonetic Association* 34. 201–210.
- NORMAN, JERRY. 1988. *Chinese*. Cambridge University Press.
- OHALA, JOHN AND JAMES LORENTZ. 1977. The story of [w]: An exercise in the phonetic explanation for sound patterns. *Berkeley Linguistic Society* 3. 577-599.
- SHEN, JIONG. 1987. Beijinghua hekouhu ling shengmude yuyin fenqi. (北京話合口呼零聲母的語音分歧) (Phonetic differences of zero initial before finals beginning with u in the Beijing dialect) *Zhongguo Yuwen* 5. 352-362.
- STEVENS, KEN. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- WAN, I-PING ; JERI, JAEGER. 1998. Speech errors and the representation of tone in Mandarin Chinese. *Phonology* 15. 417-461.
- XINHUA NEWS. 2009. 印尼山体滑坡造成的死亡人数升至 1 4 人. Retrieved on 11 November, 2009. website: http://news.xinhuanet.com/world/2009-11/10/content_12420336.htm
- YIP, MOIRA. 2003. Casting doubt on the Onset/Rime distinction. *Lingua* 113. 779 - 816.
- ZHANG, QING. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34. 431-466.

Chinese and English Relative Clauses: Processing Constraints and Typological Consequences

Chien-Jer Charles Lin
Indiana University

Languages with distinct typological characteristics often present different challenges to language processing and thus lead to differences in linguistic performances across languages. This study investigated the relations between linguistic typology, constraints on language processing, and crosslinguistic differences in language use. Two studies on Chinese and English relative clauses were conducted. The first study looked at the semantic properties of the head nouns in Chinese and English relative clauses, showing that shared communicative functions lead to shared patterns in Chinese and English relative clauses: the heads of object relative clauses tend to be inanimate nouns, while those of subject relative clauses tend to be animate. The second study looked into the use of relative clauses in parallel texts, showing that the typological differences in terms of head positions in Chinese and English led to distinctive complexity patterns. Chinese relative clauses tend to be shorter (in terms of number of syllables and number of words) and less complex (in terms of embeddings other relative clauses) than English relative clauses. These results supported the Head-driven Constituent Complexity Hypothesis.

1. Introduction

Languages with distinct typological characteristics often present different challenges to language processing. Such differences lead to differences in linguistic performances across languages. Previous attempts to associate language processing with linguistic typology have focused on two themes: using universal processing constraints to account for crosslinguistic differences and finding functional contrasts in different word orders. The first approach was exemplified by the seminal typological study of relative clauses conducted by Keenan and Comrie (1977), who investigated the relativizability of noun phrases at different grammatical positions across languages and found that different grammatical positions showed differences on the ease of extraction across languages. Noun phrases in certain grammatical positions are more likely to be relativized than noun phrases in other positions. They proposed the well-known Keenan-Comrie Accessibility Hierarchy (i.e, Subject > Direct Object > Indirect Object > Oblique object > Genitive > Object Complement, abbreviated as the AH), which summarized the likelihood of relativization across languages. When a language can relativize noun phrases at a gram-

matical position, it can also relativize noun phrases at all higher positions on the hierarchy. For this generalization, they provided a processing account: noun phrases higher in the hierarchy are easier to access than those that are lower in the hierarchy. They explicitly attributed this accessibility hierarchy to “the psychological ease of comprehension”, namely, that “the lower a position is on the AH, the harder it is to understand relative clauses formed on that position (p.88).”

More recently, Langus and Naspor (2010) entertained the hypothesis that different cognitive functions motivate two popular word orders in human languages—SVO and SOV, which together accounted for 76% of human languages (Dryer, 2005). According to this hypothesis, SVO is a preferred word order based on the computational system of human syntax, while SOV is a preferred form for effective communication. The support for SOV being communicatively motivated came from their experiments asking Italian and Turkish speakers to describe events by using gestures, not the spoken language. They found that despite the different basic word orders in Italian (SVO) and Turkish (SOV), SOV was the preferred order in gesturing. In a comprehension study of different word orders using gesture inputs, they also found the SOV order to take the least time for comprehension. Interestingly, however, in an experiment where participants listened to words and sentences in their native languages presented with prosodically flat speech, regardless of the dominant word orders in their native languages, VO is preferred over OV (when the subject position was controlled for). The preference of VO over OV is taken to support the computational superiority of SVO over SOV.

The current research provides a different take on the relation between processing and crosslinguistic contrasts. Our study examined two processing-based hypotheses by looking at relative clauses in large corpora as well as those in parallel translated texts. The corpus study looked at the animacy of the head nouns in Mandarin Chinese and compared the findings with those of the processing studies of other languages. The study of parallel translated texts examined the complexity of relative clauses in head-initial (English) and head-final (Chinese) structures.

Before we delve into the data, a few notes about Chinese and English relative clauses are in order. Chinese and English are both languages that rely heavily on word orders for coding thematic relations. The dominant word orders for both languages are SVO, semantically interpreted as Agent-Action-Patient. In terms of head positions of nominal structures, however, the two languages contrast typologically. Chinese noun phrases are head-final; all modifying phrases, including adjectivals, prepositional adverbials, and relative clauses, precede the nouns. English noun phrases are head-initial; the preposition phrases and relative clauses follow rather than precede the head nouns. Such a contrast makes Chinese and English an ideal pair of languages for investigating the effect of head positions in processing as well as language use. They are typologically distinctive in terms of the head positions in noun phrases; yet, they are typologically similar in terms of the basic word orders. Such a contrast allows us to attribute differ-

ences in processing and language use to the position of the heads internal to noun phrases rather than to other factors.

In section 2 of this article, we report a corpus study of Chinese relative clauses, looking at the animacy of the head nouns. Section 3 investigated how the typological differences in terms of head positions in Chinese and English led to distinctive clausal complexity of Chinese and English relative clauses in parallel translated texts. Section 4 concludes the paper.

2. Study 1: Animacy of the head in Chinese and English relative clauses

Several recent processing studies of relative clauses have shown that the animacy of the head noun affects how easy relative clauses are for comprehension. Mak, Vonk, and Schriefers (2002) demonstrated that while subject relatives are generally easier to process than object relatives in Dutch, when the head nouns of the object relatives were inanimate, there were no processing differences between subject and object relatives. Mak, Vonk, Schriefers (2006) further demonstrated that animacy on both the head noun and the noun phrases inside the relative clauses is important in processing.

Gennari and MacDonald (2008) also investigated the effect of animacy on English relative clauses. They conducted a gated sentence completion task using animate and inanimate nouns as the head nouns. Only 15% of the sentences with animate head nouns were completed as object relative clauses, while 65% of the sentences with inanimate head nouns were completed as object relatives. These results were corroborated by reading time data.

Regarding the animacy of the head nouns of Chinese relative clauses, Wu, Kaiser, and Anderson (2009) extracted 331 relative clauses from the Chinese Treebank 5.0 corpus (Palmer, Chiou, Xue & Xia, 2005). They examined the animacy information on the head noun and the embedded noun in subject and object modifying relative clauses that involved subject and object extractions and found that subject extracted relative clauses tended to have animate heads and inanimate embedded nouns, while object extracted relative clauses tended to have inanimate heads and animate embedded nouns. The self-paced reading experiments they conducted showed the same preference for contrastive animacy on the head noun and the embedded noun. Wu, Kaiser, and Anderson (2011) further extracted 1218 relative clauses in a later study. They found that the heads of object relatives are overwhelmingly inanimate, but the animacy of the heads of subject relatives depended on whether the relative clauses modified the subject or the object of the main clause. If the relative clause modifies the subject, then its head noun tends to be animate. If it modifies the object, then its head noun is equally likely to be animate or inanimate.

The current study is interested in whether the animacy information on the head nouns of Chinese relative clauses would follow the same universal pattern; that is, whether we would find more inanimate head nouns in object relatives and more animate head nouns in subject relatives. In comparison with previous studies, our study collected

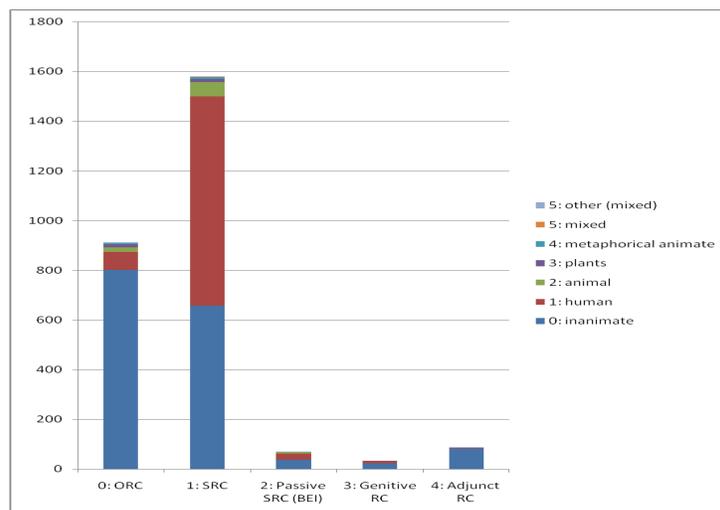
a greater number of relative clauses and classified these relative clauses into finer categories. 3075 relative clauses were extracted from Sinica Treebank 3.0 released by Academia Sinica in Taiwan (<http://godel.iis.sinica.edu.tw/CKIP/treebank.htm>). These relative clauses were manually coded on syntactic and semantic dimensions. In previous studies, the passive relatives, the possessive relatives, and the adjunct relatives were not coded separately. Our study makes it possible to look at animacy of these relative clauses as well. For the purpose of the current study, we categorized relative clauses into the following types for illustration:

(1) Classification of relative clauses extracted from Treebank 3.0:

Type	Example
0: ORC	張老師處罰的學生
1: SRC	沒有考過的人
2: Passive SRC	被處罰的學生
3: Possessive RC	家被颱風吹垮的居民
4: Adjunct RC	張老師處罰學生的原因

In terms of the overall distribution, more subject relatives were found than object relatives, confirming previous findings on the dominant frequencies of subject relatives over object relatives. All other types of relative clauses were relatively infrequent.

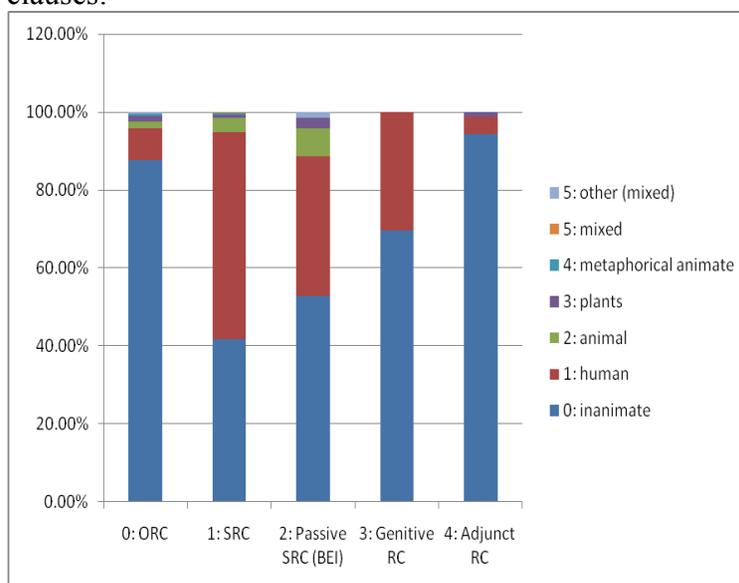
(2) Raw frequencies of different kinds of relative clauses and animacy of the head nouns:



Semantically, the head nouns were coded as inanimate, human, animal, plant, and metaphorically animate. Within each type of relative clauses, different distributions regarding head noun animacy were found. Close to 90% of object relatives in Chinese

have inanimate head nouns, while animate and inanimate head nouns were more equally distributed on subject relatives (though there were more animate head nouns than inanimate head nouns). The dominance of inanimate head nouns on object relatives is consistent with a similar dominance found in English and Dutch. These results reflect a universal tendency for the patient roles (located at the object positions) to be inanimate. A similar preference was found on passive relatives; even though the extracted head nouns of passive relative clauses are at the subject position in the embedded clauses, they usually hold the patient role thematically. Therefore, these head nouns are more prone to be inanimate and nonhuman.¹

(3) Distribution of animacy information on the head nouns of different kinds of relative clauses:



These findings confirmed a universal preference for inanimate head nouns in object relative clauses. The head nouns of subject relatives, however, are less biased in terms of animacy (similar to the equivalence of animate and inanimate head nouns in object modifying subject-extracted relatives obtained by Wu et al. 2011).

3. Study 2: Chinese and English relative clauses in parallel (translated) texts

The second study focused on how the differences of head positions may result in different degrees of complexity in the production of head-initial and head-final relative clauses. We propose the *head-driven constituent complexity hypothesis* (HCCH) based on the observations that typologically opposite head positions may induce distinct pro-

¹ Note, however, that the subjects of passives in Mandarin Chinese are not overwhelmingly inanimate. Close to 50% of these head nouns were actually animate nouns.

cessing strategies (Lin, 2011) and that the later the head noun is encountered, the greater temporary uncertainty exists in parsing, and therefore the more difficult for parsing. We therefore hypothesize that head-final structures are overall harder than head-initial structures due to the uncertainty prior to the appearance of the heads. The HCCH predicts that given the same contexts and similar textual contents, head-final constituents tend to be shorter and less complex than head-initial constituents.

To test this hypothesis, we compared the complexity of relative clauses in English and Chinese texts. This paper presents only the preliminary results based on our textual analysis of two sets of comparable essays. However, the contrasts were distinctive enough and therefore the results were quite suggestive even though the data were limited at the current phase. The complexity of relative clauses was measured by the length of the relative clauses and the structural complexity of embeddings. The lengths of relative clauses were measured by (a) the number of syllables before the head noun in an embedded clause and (b) the number of words (i.e., semantic units) before the head noun in an embedded clause. An embedded clause is structurally complex when it has another relative clause embedded in it.

Two comparable popular magazines were selected as the targeted texts--*Taiwan Panorama* (*Taiwan Guanghai Zazhi* 台灣光華雜誌) and *Scientific American* (*Kexueren* 科學人). An article from the *Taiwan Panorama*, which was originally written in Chinese and translated into English, was randomly selected for analysis. This article was made of 128 English sentences.² In the corresponding Chinese text, 82 relative clauses were identified.³ The average lengths of the Chinese relative clauses were 6.63 in terms of number of syllables and 3.78 in terms of number of words. In the translated English text, 48 relative clauses were identified, and the average length of these relative clauses was 11.28 syllables long corresponding to 6.83 words. Therefore, both in terms of phonological length and semantic complexity, head-initial relative clauses tend to be longer than head-final relative clauses. Among the Chinese relative clauses, no relative clauses were embedded in other relative clauses. Among the English relative clauses, 4 of them were embedded in other relative clauses. English relative clauses are more tolerant of structural complexity than Chinese relative clauses.

To be sure that the difference was not due to translated texts versus original texts, we also analyzed relative clauses in an article originally written in English (published in *Scientific American*) and its Chinese translation published in the magazine *Kexueren* (科學人). In the original English text, which contained 111 sentences, 39 relative clauses were identified. The average number of syllables in the English relative clauses was

² Number of sentences were measured based on the English texts because the definition of what a sentence is more definite in English than in Chinese.

³ Relative clauses in Chinese are defined as clauses embedded in DE-phrases with verbs (including stative verbs) in them. English relative clause are defined as clauses embedded in noun phrases with relativizers such as, *who*, *which*, *that*, *whom*, *whose*, *what*, *where*, *why* in them, and in noun phrases with embedded verbs in the forms of past participles and gerunds.

13.47, which included an average of 7.39 number of words. 77 relative clauses were identified in the Chinese text. The average length of the Chinese relative clauses was 7.32 syllables and 4.51 words long. Among the relative clauses, 3 relative clauses in English were embedded in relative clauses, and no relative clauses in Chinese were embedded in other relative clauses.⁴

- (4) Lengths of Chinese and English relative clauses in the parallel texts (number of syllables):

	Chinese to English	English to Chinese
Chinese relative clauses	6.63	7.32
English relative clauses	11.28	13.47

- (5) Lengths of Chinese and English relative clauses in the parallel texts (number of words):

	Chinese to English	English to Chinese
Chinese relative clauses	3.78	4.51
English relative clauses	6.83	7.39

- (6) Number of relative clauses embedded in other relative clauses in the Chinese-English parallel texts:

	Chinese to English	English to Chinese
Chinese relative clauses	0	0
English relative clauses	4	3

These results confirmed the HCCH. Regardless of the direction of translation, Chinese relative clauses are on average shorter than English relative clauses both in terms of phonological length (i.e., number of syllables) and in terms of semantic units (i.e., number of words). English relative clauses are more likely to have relative clauses embedded in them but no occurrences of multiply embedded relative clauses have been found in Chinese. Both in terms of semantic and phonological quantity and in terms of structural complexity, therefore, head-initial relative clauses tend to be more complex than head-final relative clauses.

4. Concluding remarks

In this paper, we investigated two aspects of Chinese and English relative clauses—animacy of the head nouns and clausal complexity. Previous studies of head-initial relative clauses (in Dutch and in English) showed that animacy of the head is important in determining whether an object relative clause is easy or difficult to comprehend. An

⁴ Note that 3 Chinese relative clauses were embedded in genitive DE-phrases, which, though not considered as relative clauses in our study, appeared to be structurally complex nevertheless.

object relative clause with an inanimate head noun and an animate embedded noun was no more difficult than a subject relative clause. Given shared functions of relative clauses across languages (i.e., modifying nouns and providing grounding information in a sentence), it is expected that animacy patterns on the head would be shared across languages. Our study of the head nouns in Chinese relative clauses in corpora indeed found overwhelming use of inanimate nouns as the head nouns of object relative clauses in Chinese. The universal pattern suggests that human languages tend to use animate nouns as the subject and inanimate nouns as the object.

The second study looked into the complexity of Chinese and English relative clauses. We tested the Head-Driven Constituent Complexity Hypothesis, which predicted that head-initial relative clauses are more tolerant of clausal complexity than head-final relative clauses. Given that Chinese and English shared the same word orders, our hypothesis is that any difference found between Chinese and English relative clauses may be due to the opposite positions of the heads inside noun phrases. Our comparisons of parallel texts of Chinese and English showed that in both texts translated from Chinese to English and those translated from English to Chinese, Chinese relative clauses were shorter and less tolerant of complexity than English relative clauses.

For future research, the corpus study on animacy should also consider the information status of these noun phrases and whether the relative clauses modified the subject or the object of the matrix clauses. The comparison between Chinese and English relative clauses can be better understood when the complexity of other aspects of Chinese and English structures are also contrasted.

References

- Gennari, Silvia P. and Maryellan C. MacDonald. 2008. Semantic indeterminacy and relative clause comprehension. *Journal of Memory and Language* 58.161-87.
- Dryer, Matthew. 2005. The order of subject, object and verb. *The world atlas of language structures*, ed. by M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie, 330-333. Oxford: Oxford University Press.
- Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.63-99.
- Langus, Alan and Marina Nespore. 2010. Cognitive systems struggling for word order. *Cognitive Psychology* 60.291-318.
- Lin, Chien-Jer C. 2011. Thematic templates and the processing asymmetry of Chinese relative clauses. Manuscript. Indiana University, Bloomington.
- Mak, W. M., W. Vonk and H. Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47.50-68.
- Mak, W. M., W. Vonk and H. Schriefers. 2006. Animacy in relative clauses: The hikers that rocks crush. *Journal of Memory and Language* 54.466-490.
- Palmer, M., F.-D. Chiou, N. Xue and F. Xia. 2005. The Penn Chinese Treebank Corpus. University of Pennsylvania.

- Wu, Fuyun, Elsi Kaiser and Elaine Andersen. 2009. Animacy effects in Chinese relative clause processing. *Proceedings of the 38th western conference on linguistics*, ed. by Michael Grosvald and Dionne Soares, 318-329. Davis: University of California, Davis.
- Wu, Fuyun, Elsi Kaiser and Elaine Andersen. 2011. Subject preference, head animacy, and lexical cues: A corpus study of relative clauses in Chinese. *Processing and producing head-final structures*, ed. by Hiroko Yamashita, Yuki Hirose, and Jerry Packard, 173-194. The Netherlands: Springer.

Locality versus Anti-locality Effects in Mandarin Sentence Comprehension

Yowyu Lin
National Chengchi University

Whether the distance between the filler and the gap increases or decreases the difficulty for comprehenders has always been subject to debate. Locality theory claims that the greater the distance between the filler and the gap, the more difficult it is. However, anti-locality theory predicts that the greater the distance, the easier it is. This study examined whether it was the position of the verbs that made the differences. Three kinds of structures, simple SVO, relative clauses, and BA constructions, were manipulated. If the locality theory is correct, increasing reading times at the critical words should be observed. If anti-locality theory is correct, decreasing reading times should be found. In both SVO and RCs, we observed increasing reading times at the critical nouns. In BA constructions, however, we observed decreasing reading times at the critical verbs. It is thus the critical word at the end that leads to differences.

1. Introduction

Whether the distance between the filler and the gap increases or decreases the difficulty for comprehenders have always been subject to debate in the field of sentence comprehension. According to locality theory, e.g., Gibson's Dependency Locality Theory (Gibson, 1998, 2000; Babyonyshev & Gibson, 1999), it attributes language processing difficulties to the distance between an argument and its head. Although a majority of studies have found evidence in support of the locality effect, there are some counterexamples (Konieczny, 2000; Vasishth & Lewis, 2006) from head-final languages such as German and Hindi. They found that the longer the distance, the more difficulty it is for comprehenders to interpret. The situation that the farther the argument from its head the faster the reading times has been called anti-locality effect.

Konieczny (2000) was among one of the first researchers to notice anti-locality effect. He compared two detailed metric calculation of processing difficulties, one being Dependency Locality Theory and the other, Hawkins' (1994) IC-to-Words calculation based on the concept of Early Immediate Constituent (EIC). According to Hawkins, "...words and constituents occur in the orders they do so that syntactic groupings and their immediate constituents (ICs) can be recognized (and produced) as rapidly and

LIN: LOCALITY VERSUS ANTILOCALITY

efficiently as possible in language performance” (p. 57). The ratio, IC-to-Words, was proposed to account for the processing difficulty by examining whether constituents that should be grouped together is immediately recognized or not. If the ratios are maximized, there will be efficient and faster recognition of the mother node and the examples are below:

- (1)
- a. I_{VP}[gave NP[the valuable book that was extremely difficult to find] PP[to Mary]]
 1 2 3 4 5 6 7 8 9 10 11
- b. I_{VP}[gave PP[to Mary] NP[the valuable book that was extremely difficult to find]]
 1 2 3 4

The verb “gave” in (1) is a dative verb, which requires three arguments. After processing “I” and “gave”, comprehenders would expect to see another two arguments. Take (1) for example, there are two possibilities to continue the sentence: one is to adjoin NP right after “gave” as in (1a). The other possibility is to adjoin the PP right after the verb “gave” as in (2b). Since three constituents (V, NP and PP) are required in this example, (1b) facilitates comprehenders’ understanding than (1a) because it only takes 4 words (at “the” position) to complete the requirement of searching for the other two arguments. On the other hand, construction of VP will be greatly delayed in (1a) since it not until the 11th word “to” that makes people realize the presence of NP and complete the search for arguments. Simply put, of all the twelve words, comprehenders need to scan 11 words for recognition in (1a) while only 4 words in (1b). Thus people should prefer (1b) over (1a). The metric IC-to-words calculation of the two sentences here are 3/11=27.3% for (1a) and 3/4=75% for (1b).

To further investigate locality versus antilocality issue, Konieczny (2000) examined the effect of relative clause position, relative clause length and extraposition distance In a 2 x 3 x 3 design. Examples are given below:

- (2)
- a. Er hat das Buch, *das Lisa gestern gekauft hatte*, **hingelegt**.
 he has the book that Lisa yesterday bought had laid_down
 (“He has laid down the book that Lisa had bought yesterday.”)
- b. Er hat das Buch **hingelegt**, *das Lisa gestern gekauft hatte*.
 he has the book laid_down that Lisa yesterday bought had
 (“He has laid down the book that Lisa had bought yesterday.”)

The italicized regions represent German relative clauses, which can be either directly adjoined to the main clause or extraposed to the right of the main clause. When the relative clause is extraposed, it crossed one word, three-four words, or five-six words, the latter two of which being prepositional phrases. The examples are again given below:

LIN: LOCALITY VERSUS ANTILOCALITY

(3)

a. One Word (no PP)

Er hat die Rose **hingelegt** + RC (extraposed)

He has the rose laid_down + RC

b. Three-four Words (short PP)

Er hat die Rose auf den Tisch **gelegt** + RC (extraposed)

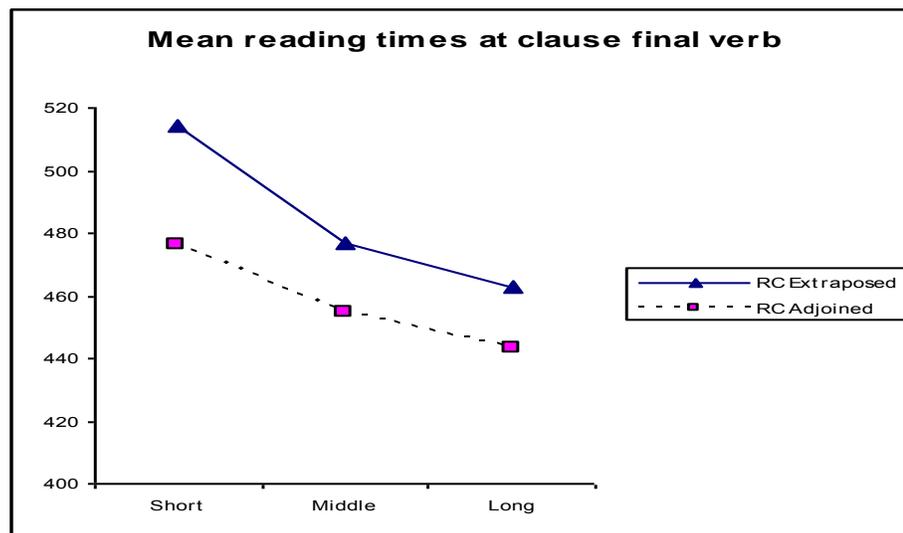
He has the rose on the table laid + RC

c. Five-six Words (long PP)

Er hat die Rose auf den Kleinen runden Tisch **gelegt** + RC (extraposed)

He has the rose on the small round table laid + RC

In terms of the RC length, it varied from three-five words, six-eight words to nine-eleven words. According to DLT and other locality-based theories, reaction times should increase at the position of “hingelegt” since it is computationally more costly when this verb has to be integrated with a farther previous structure. However, IC-to-Word has different predictions. When the relative clause is extraposed across a longer prepositional phrase, the percentage of IC-to-Word ration should drop sharply. This is because with the same number of constituents have to be divided by a larger number of total words. However, when the relative clause is adjoined to the main clause, the percentages remain similar across all 9 conditions. The actual mean reading times at the clause final verb by levels of RC is also reconstructed as follows:



The figure above shows that for regardless of the position where the relative clause

LIN: LOCALITY VERSUS ANTILOCALITY

is attached, reading times drop sharply when the relative clause is long. Basically, this result goes against the prediction of locality theory. Furthermore, in spite of the fact that Hawkins' EIC predicts that when RC is short, people should prefer it to be extraposed, the reading times shows faster reading times for the relative clause to be adjoined. However, except for this position, when RCs are of middle and long length, the results conform to Hawkins' predictions. Konieczny concluded that their results do not support the locality theory. Instead, it conforms more to the so-called "anticipation hypothesis, which states that the verb can be anticipated through preceding arguments" (p. 643).

Another piece of evidence came from Hindi, Vasishth and Lewis' (2006) example, which also found the anti-locality effect in Hindi, another head-final language, with facilitation at the inner most verb, *buy*, when different intervening phrases, including adverbs, prepositional phrase and relative clauses, were interposed. In addition, they also found facilitation in processing at the inner most verb with intervening phrases when compared with no intervening items.

Studies which found faster reading times at the verb position seem to be based in head final languages like German and Hindi, but not in head-initial languages like English. However, there is one exception. Jaeger et al. (2005) inserted one, two and three prepositional phrases in the English relative clause region to investigate the processing speed at the main clause verb "bought", which is getting farther and farther away from the main clause subject. Their examples are also given below:

(4)

- a. The player [that the coach met *at 6 o'clock*] **bought** the house..
- b. The player [that the coach met *by the river at 6 o'clock*] **bought** the house..
- c. The player [that the coach met *near the gym by the river at 6 o'clock*] **bought** the house...

If the locality effect is correct, longer reading times could be observed at the bought position. If the anti-locality effect is true, faster reading times could be observed. In this experiment, they argued to have found facilitation effect at the verb "bought" when more prepositional phrases were used in the relative clause region.

Even though most of the above-mentioned studies have claimed to found anti-locality effect, concerns arise regarding whether comprehenders need to process these adverbial or prepositional phrases at all when they process these stimuli sentences. For example, the comprehension question for (6) could be "Did the coach buy the house?" which does not require comprehenders to process the prepositional phrases at all. If subjects do not need to process the prepositional phrases, it will not be too surprising that people might speed up more and more just to get rid of the prepositional phrases. Thus the faster reading times at the inner most verb could be a spill over effect from the speed up of the prepositional phrases.

Since most anti-locality effect was found in head-final languages like German and

LIN: LOCALITY VERSUS ANTILOCALITY

Hindi, we decide to examine whether this effect appears in another head-final language like Mandarin. We also plan to embed one, two and three prepositional phrases into the stimuli sentences. Because of some similarities between prepositional phrase and relative clause, we need to illustrate the prepositional phrase first.

In Mandarin, if we need to use a prepositional phrase to modify a noun, we need to include DE in between, as the following example shows.

- (5)
攤位旁的女孩
Standnext DE girl
(the girl next to the stand)

When we embed the above structure into a simple SVO structure, we will get a stimuli sentence as in (8a). Note that this structure looks a lot like our relative clause since it includes a typical DE. If the ant-locality effect is correct, we should be able to observe faster reading times at the final main clause object, girl, as more and more prepositional phrases are included in the structure.

- (6)
a. SVO One Prepositional Phrase
同學 喜歡 攤位旁 的 女孩
Classmate like stand-near DE girl...
(*"The classmate likes the girl who is next to the stand."*)
b. SVO Two Prepositional Phrases
同學 喜歡 市場內 攤位旁 的 女孩
Classmate like market-inside stand-near DE girl
(*"The classmate likes the girl who is next to the stand inside the market."*)
c. SVO Three Prepositional Phrases
同學 喜歡 公園邊 市場內 攤位旁 的 女孩
Classmate like park-next market-inside stand-near DE girl
(*"The classmate likes the girl who is next to the stand inside the market next to the park."*)

Since the previous studies all include relative clauses as their stimuli sentence, we would like to include relative clauses with different numbers of prepositional phrases. If we want to turn (8) into relative clauses, we need to insert another DE after the verb. The stimuli sentences are like the following:

- (7)
a. RC One Prepositional Phrase
同學 喜歡 的 攤位旁 的 女孩

LIN: LOCALITY VERSUS ANTILOCALITY

Classmate like DE stand-near DE girl
(*“The classmate likes the girl who is next to the stand.”*)

b. RC Two Prepositional Phrases

同學 喜歡 的 市場內 攤位旁 的 女孩
Classmate like DE market-inside stand-near DE girl
(*“The classmate likes the girl who is next to the stand inside the market.”*)

c. RC Three Prepositional Phrases

同學 喜歡 的 公園邊 市場內 攤位旁 的 女孩
Classmate like DE park-next market-inside stand-near DE girl
(*“The classmate likes the girl who is next to the stand inside the market next to the park.”*)

(7) shows two DEs in the same sentence. The first DE represents relative clause while the second DE goes with the prepositional phrase. Therefore, the only difference between (8) and (9) lies in the first DE only. Again, if the locality theory or DLT is correct, the reading times for the final noun “girl” would be longer as more prepositional phrases are inserted into the stimuli. However, if anti-locality theory is correct, we should observe faster reading times for the final noun in (8c).

Even though we have constructed stimuli sentences with different numbers of prepositional phrases inserted, there is an important difference between our stimuli and those in the previous stimuli. All of the anti-locality effect was found in head-final languages where the verb appears after the arguments. Even though Mandarin has this head-final property, this property is mostly confined to relative clause constructions, where it is the head noun that appears at the end of the clause. We suspected that one of the reasons for the anti-locality effect to occur might be due to the verb, which appear at the final position. Verbs can be used to denote a large variety of events (McKoon & Ratcliff, 2003) and its importance has been found in many studies (Stowe, 1986; Garnsey, Tanenhaus & Chapman, 1989). In a simple Noun1 Verb Noun2 structure, we know that someone or something (Noun1) engages in some activity that affects someone or something else. The verb links together two entities and let people realize the relationship between these two entities. Boland, Tanenhaus, Garnsey and Carlson (1995) used stop-making sense paradigm to investigate wh-questions with respect to verb argument structure. For an implausible wh-phrases, there is a clear verb argument effect at the verb position. However, if the wh-question is plausible to be temporarily interpreted as the argument of the verb, the effect went away. These examples suggested that when arguments precede the verb, there is a tendency for the subjects to search for the verb as soon as possible so that the argument structure assignment or filler-gap assignment can be fulfilled.

Given the above findings, we have reason to believe that the reason for the antilocality effect to appear may be due to the fact the arguments all occur prior to the verb in German and Hindi. Speakers of these languages need to get to the verb as soon as

LIN: LOCALITY VERSUS ANTILOCALITY

possible to fulfill the assignment of verb argument structure. If it is the verb that facilitates the sentence comprehension, such facilitation will not appear in the head-final relative clause in Mandarin since the final word is a noun. In this situation, Mandarin is particularly suited for studying this locality versus anti-locality effect because of its flexibility in word order. By using BA constructions, it is possible to place nouns or arguments before the verb, as (8) shows. In these examples, the verb, hit, appears at the final position and is modified by one, two or three adverbial phrases.

(8)

a. BA One Adverbial Phrase

同學 把 壞人 狠狠地 打了 一頓

Classmate BA bad guy seriously beat once

(*"The classmate beat the bad guy seriously."*)

b. BA Two Adverbial Phrases

同學 把 壞人 狠狠地 結實地 打了 一頓

Classmate BA bad guy seriously fully beat once

(*"The classmate beat the bad guy fully and seriously."*)

c. BA Three Adverbial Phrases

同學 把 壞人 好好地 狠狠地 結實地 打了 一頓

Classmate BA bad guy completely seriously fully beat once

(*"The classmate beat the bad guy fully and seriously."*)

If it is the verb that facilitate the processing, we would be able to observe faster reading times at the verb in (8) but not in the noun in (7). However, if it is not the verb that leads to faster reading times, or locality theories are correct, we would expect to observe elevated reading times at both the verb and the noun position.

2. Materials, design and procedure

The experiment was conducted using a self-paced moving-window reading paradigm, like the one that we have used previously. Thirty native speakers of Mandarin, recruited from University of Illinois at Urbana-Champaign, participated in the experiment. Of the thirty subjects, 28 of them were graduate students while 2 of them were undergraduate students.

Three different kinds of structures in Mandarin were used as the stimuli: simple SVO, relative clauses and BA construction. A complete list of stimuli is given as follows:

(9)

a. SVO One Prepositional Phrase

同學 喜歡 攤位旁 的 女孩 。

Classmate like stand-near DE girl Period

(*"The classmate likes the girl who is next to the stand."*)

b. SVO Two Prepositional Phrase

LIN: LOCALITY VERSUS ANTILOCALITY

同學 喜歡 市場內 攤位旁 的 女孩 。

Classmate like market-inside stand-near DE girl Period

(*"The classmate likes the girl who is next to the stand inside the market."*)

c. SVO Three Prepositional Phrase

同學 喜歡 公園邊 市場內 攤位旁 的 女孩 。

Classmate like park-next market-inside stand-near DE girl Period

(*"The classmate likes the girl who is next to the stand inside the market next to the park."*)

d. RC One Prepositional Phrase

同學 喜歡 的 攤位旁 的 女孩 已經 來了 。

Classmate like DE stand-near DE girl already come Period

(*"The classmate likes the girl who is next to the stand."*)

e. RC Two Prepositional Phrase

同學 喜歡 的 市場內 攤位旁 的 女孩 已經 來了 。

Classmate like DE market-inside stand-near DE girl already come Period

(*"The classmate likes the girl who is next to the stand inside the market."*)

f. RC Three Prepositional Phrase

同學 喜歡 的 公園邊 市場內 攤位旁 的 女孩 已經 來了 。

Classmate like DE park-next market-inside stand-near DE girl already come Period

(*"The classmate likes the girl who is next to the stand inside the market next to the park."*)

g. BA Construction with One Adverbial Phrase

同學 把 壞人 狠狠地 打了 一頓 。

Classmate BA bad guy seriously beat once Period

(*"The classmate beat the bad guy seriously."*)

h. BA Construction with Two Adverbial Phrases

同學 把 壞人 狠狠地 結實地 打了 一頓 。

Classmate BA bad guy seriously fully beat once Period

(*"The classmate beat the bad guy seriously."*)

i. BA Construction with Three Adverbial Phrases

同學 把 壞人 好好地 狠狠地 結實地 打了 一頓 。

Classmate BA bad guy completely seriously fully beat once Period

(*"The classmate beat the bad guy seriously."*)

Seventy-two sets of stimuli were created and were separated into three lists by using Latin Square design. Each of the subjects saw one of the lists which contained 216 stimuli sentences and 84 fillers. The whole experiments took around 40 minutes to complete. In addition, comprehension questions were created in a way to make sure that subjects had to process the prepositional and adverbial phrases.

3. Results

The results of the comprehension questions are presented below:

	One Phrase	Two Phrases	Three Phrases
Simple SVO	94%	91%	86%
Relative	86%	81%	75%
Clauses			
BA	85%	84%	81%

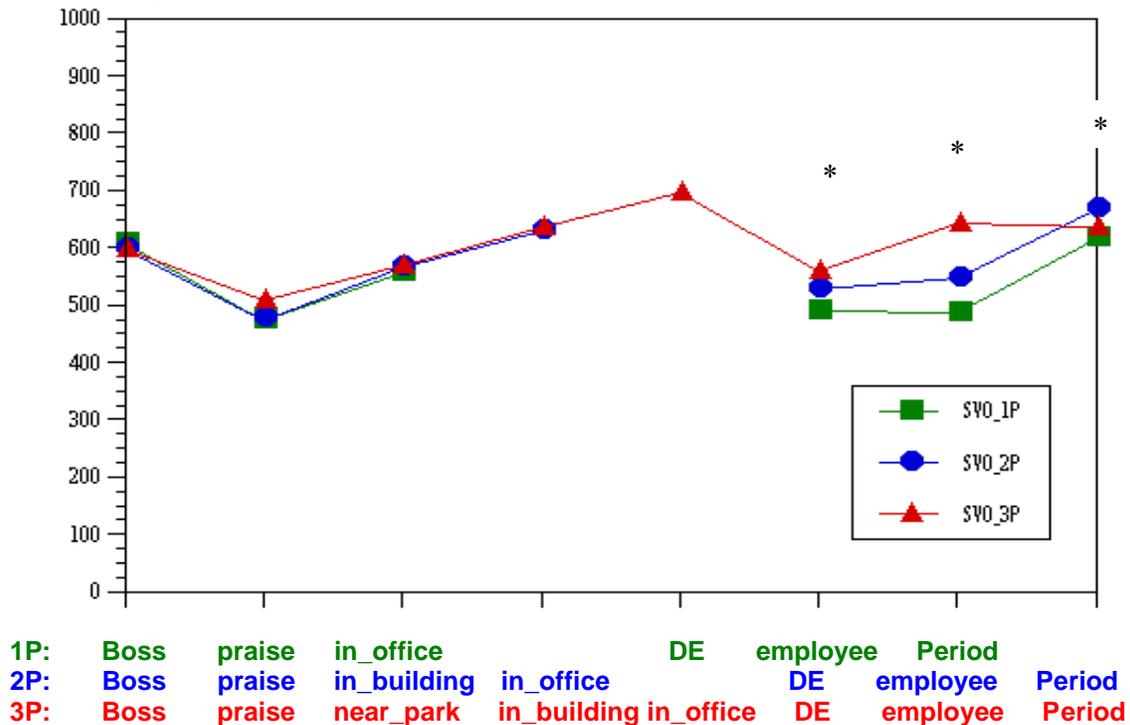
LIN: LOCALITY VERSUS ANTILOCALITY

construction			
--------------	--	--	--

Table 1: Correct Rate of Comprehension Questions

As can be seen from the table above, the proportion correct for comprehension questions drops in all three structures when more phrases are inserted. Relative clauses with three prepositional phrases dropped the most. This should not be too surprising since the more prepositional phrases there are, the more information people have to remember so as to answer the comprehension questions correctly. With respect to BA construction, there is no reliable difference between BA_1 AP and BA_2 AP ($\chi^2(1)=0.67, p>.05$). However, there are marginally reliable differences between BA_2P and BA_3P ($\chi^2(1)=3.48, p=0.062$) and reliable differences between BA_1P and BA_3P ($\chi^2(1)=6, p<.05$). With respect to SVO constructions, there are marginally reliable differences between SVO_1P and SVO_2P ($\chi^2(1)=3.27, p=0.0707$) and between SVO_2P and SVO_3P, ($\chi^2(1)=4.72, p<.05$) and between SVO_1P and SVO_3P ($\chi^2(1)=15.81, p<.01$). In relative clause comprehension questions, there are reliable differences across all three pairs (RC_1P vs RC_2P: $\chi^2(1)=4.48, p<.05$; RC_2P vs RC_3P: $\chi^2(1)=9.67, p<.01$; RC_1P vs RC_3P: $\chi^2(1)=15.72, p<.01$).

The results of the reading times of SVO, RC and BA constructions are presented consecutively as follows:



LIN: LOCALITY VERSUS ANTILOCALITY

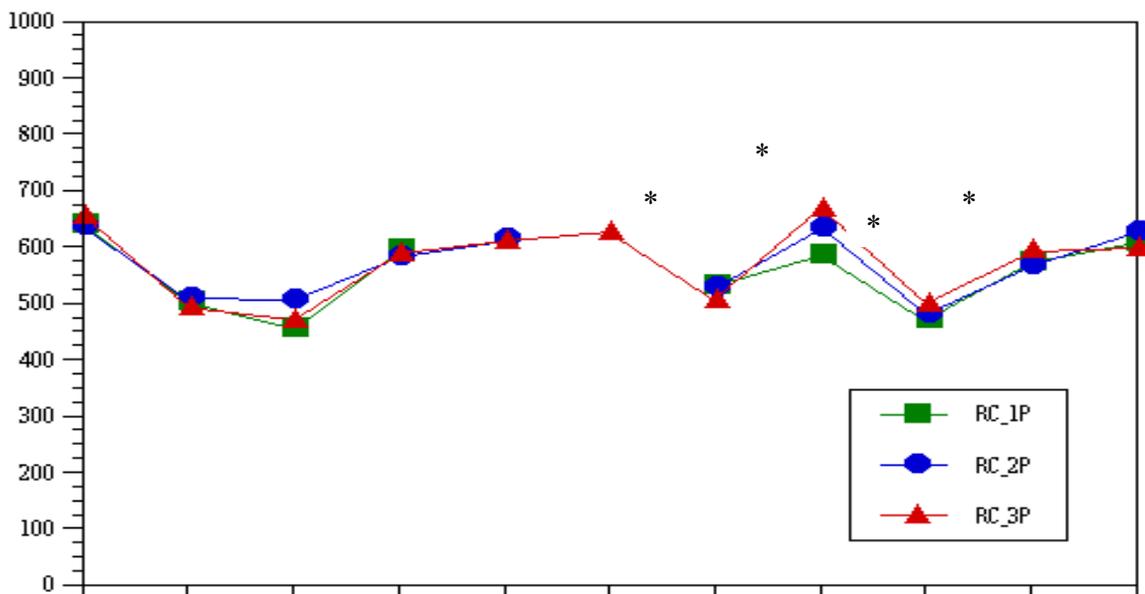
Figure 1: SVO with one, two and three prepositional phrase

The figure above presents the result from SVO with one prepositional phrase (green line), two prepositional phrases (blue line) and three prepositional phrases (red line). In order to align the final three positions since they are all the same *DE*, *head noun* and the *period*, the two lines, SVO with one and two prepositional phrases, actually contain a break. A closer look at the figure reveals that the more the prepositional phrases, the more elevated the reading times. Basically, the reading times for the first the second prepositional phrases almost lie on top of one another. Afterwards, there is a sharp drop in reading times at the DE position for all three lines. Despite the drop in reading times across three conditions, there were reliable differences for the pairwise comparisons (SVO_1P vs SVO_3P: $f(2128)=23.18$, $p<.01$; SVO_2P vs SVO_3P: $f(2128)=4.56$, $p<.05$); and SVO_1P vs SVO_2P: $f(2128)=7.23$, $p<.01$). This effect has got carried over to the next position, the main clause object position, where there were also reliable differences for the pairwise comparisons (SVO_1P vs SVO_3P: $f(2128)=26.65$, $p<.01$; SVO_2P vs SVO_3P: $f(2128)=78.30$, $p<.01$); and SVO_1P vs SVO_2P: $f(2128)=13.87$, $p<.01$).

At the period position, there is a main effect of the number of prepositional phrases ($f(2118)=5.01$, $p<.01$). In the pairwise comparisons, there is a reliable difference for the pair: SVO_1P vs SVO_3P ($f(2118)=7.03$, $p<.01$) and SVO_1P vs SVO_2P ($f(2118)=7.93$, $p<.01$). However, there is no difference between SVO_2P and SVO_3P ($p>.05$).

The results from both DE and main clause object suggest that the more intervening prepositional phrases there are, the more time it took people to process. The results from the SVO sentences support the predictions of locality theory, which predicts that more prepositional phrases actually lengthen the processing time.

LIN: LOCALITY VERSUS ANTILOCALITY



1P: Boss dislike DE in_office DE employee already go_home Period
2P: Boss dislike DE in_office in_building DE employee already go_home Period
3P: Boss dislike DE near_park in_office in_building DE employee already go_home Period

Figure 2: Relative Clauses with one, two and three prepositional phrases

The figure above presents relative clauses with one, two and three prepositional phrases. Again, there are gaps in RCs with one and two prepositional phrases since we aligned word positions after the second DE. As shown in the graph, reading times kept going up when more prepositional phrases were added into the structure. Then there is a drop at the DE position. The reading times at the DE position is not reliable different between the RC_1P vs RC_2P ($f(2128)=2.22, p>.05$). However, there were reliable differences between RC_1P vs RC_3P ($f(2128)=4.67, p<.05$) and RC_2P vs RC_3P ($f(2128)=13.28, p<.05$).

Even though the reading times were a little faster at the DE position for the RC with three prepositional phrases, reading times were most elevated for the relative clauses with three prepositional phrases at the head noun position. Relative clauses with two prepositional phrases at the head noun position also took longer for people to respond than the relative clause with one prepositional phrase. There is a reliable difference for the main effect of length of prepositions ($f(2128)=6.71, p<.01$). The pairwise comparisons all have shown reliable differences (RC_1P vs RC_3P: $f(2128)=26.12, p<.01$; RC_2P vs RC_3P: $f(2128)=5.54, p<.05$; RC_1P vs RC_2P: $f(2128)=7.72, p<.01$).

In the position following head noun, even though the three dots look to lie on top of each other, there is still main effect of prepositional phrases ($f(2128)=3.99, p<.05$). The pairwise comparisons reveals that there is reliable difference between RC_1P vs RC_3P ($f(2128)=6.18, p<.05$) and between RC_2P vs RC_3P ($f(2128)=5.79, p<.05$). However,

LIN: LOCALITY VERSUS ANTILOCALITY

there is no reliable difference between RC_1P vs RC_2P ($f(2128)=0.01, p>.05$). In the position before the period, there is no reliable difference of the main effect of the number of prepositional phrases ($f(2118)=2.18, p>.05$). However, in the pairwise comparisons, reliable difference only occurs in RC_2P vs RC_3P ($f(2118)=4.27, p<.05$). For the other two sets of comparisons, there is no reliable difference ($p>.05$). At the Period position, there is no main effect and the pairwise comparisons did not reach significance at all. It shows that the effect of prepositional phrases did not last to the final word.

The results from both the SVO and relative clauses supported the locality theory since our findings are in line with its predictions that more intervening phrases will lead to more reading times at the critical words.

The figure below presents BA construction with one, two and three adverbial phrases. Unlike the previous results where we found longer reading times when more prepositional phrases were inserted into the structure, we found reversed pattern in BA construction when more adverbial phrases were inserted into the structure. Decreasing reading times were observed across the three phrases. At the verb position, we found a main effect ($f(2128)=10.2, p<.01$). The pairwise comparisons also showed reliable difference across three pairs of comparisons: BA_1P vs BA_3P ($f(2128)=20.34, p<.01$), BA_2P vs BA_3P ($f(2128)=4.21, p<.05$) and BA_1P vs BA_2P ($f(2128)=6.05, p<.05$).

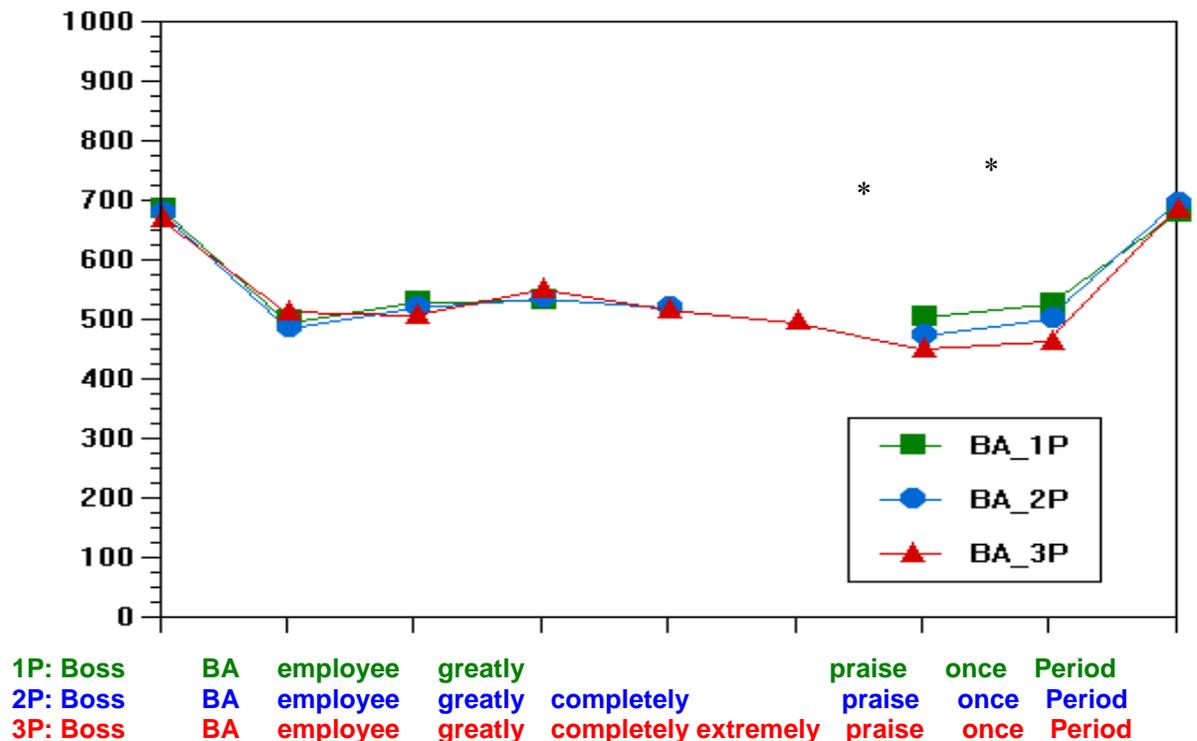


Figure 3: BA with one, two and three prepositions

LIN: LOCALITY VERSUS ANTILOCALITY

In the next position, we also found reliable difference for the main effect ($f(2128)=10.67$, $p<.01$). However, the pairwise comparisons revealed reliable differences only for two pairs: BA_1P vs BA_3P ($f(2128)=20.69$, $p<.01$) and BA_2P vs BA_3P ($f(2128)=8.85$, $p<.01$). There is no reliable difference for the third comparison: BA_1P vs BA_2P ($f(2128)=2.5$, $p>.05$). At the period position, we found no reliable difference for either the main effect. The results of the BA constructions actually supported the anti-locality theory, which claims that more intervening phrases will lead to more expectation, causing less reading times for the critical word.

4. Discussion

In this experiment, we set out to examine locality versus anti-locality theories. According to the locality theory, it claims that the greater the distance between the filler and the gap, the more difficult it is and it's possibly due to limitation of working memory. However, anti-locality theory actually claims the opposite. It predicts that the greater the distance between the filler and the gap, the easier it is and it's possibly due to increasing anticipation. Previous studies (Konieczny, 2000; Vasisht and Lewis, 2006; Jaeger et al., 2005) that have found support for anti-locality theory are mostly head-final languages like Hindi and German and have used structures with verbs as the critical words. They embedded one, two or three phrases into different kinds of syntactic structures to lengthen the distance between the verb and its preceding element. They found decreasing reading times for the critical word when more phrases are embedded in the structure. However, not many studies in the literature have found support for the anti-locality theory.

One thing to be noted in studies that have found support for antilocality theory is that their critical words were all verbs with the nouns occurring in advance as in Hindi and German. We suspected that it is the nature of the verb that made people wanted to speed up since they might be eager to know the relationship between these nouns that have occurred earlier. Adding more prepositional phrases or other kinds of phrases might just increase their anticipation to know what is going on with these nouns. However, if the final critical word is a noun, people know the relationship of at least one noun and the verb. All they need to find is another noun to complete the relationship. In other words, they would have less anticipation when the critical word is a noun.

In order to further look into the nature of locality versus antilocality theories, we decided to use Mandarin to find out why some studies support locality theory while others don't. At this juncture, Mandarin provides a great test ground for teasing apart these two theories. Given many different kinds of structures are allowed in Mandarin, we decided to make use of these structures to see if we can find out what contributes to the different claims between the two theories.

In our experiment, we used three kinds of structures, namely, simple SVO sentences, relative clauses, and BA constructions. The reason for us to use these structures is that the

LIN: LOCALITY VERSUS ANTILOCALITY

critical words for these structures differ. For simple SVO and relative clauses, the critical words are nouns. For simple SVO, the critical word is the object of the sentence while for the relative clauses, we were looking at the head noun which occur at the end of the clauses. For BA constructions, since they have N BA N V word order, the critical word is a verb. All together, we have nouns as critical word for SVO and RCs, and verbs as critical words for BA constructions. Of these three structures, we inserted one, two and three prepositional phrases into SVO and RCs and adverbial phrases into BA constructions, just like what previous studies have done.

If the locality theory is correct, we would observe increasing reading times at the critical words when more phrases are intervening regardless of whether the critical word is a noun or a verb. If anti-locality theory is correct, we could observe decreasing reading times at the critical words when more phrases are inserted. However, if our hypothesis is correct that it is due to whether the critical word is a noun or a verb that lead to difference in expectation, we would observe increasing reading times for the critical nouns in both SVO and RCs and decreasing reading times for the critical verbs in BA constructions. This may be due to the fact that in BA construction two nouns are presented in a row and it creates expectations in comprehenders to want to know what happened between the two nouns. In SVO and RCs, since a noun and a verb have been presented, all is left is the noun to complete the action. Our findings support our hypothesis. In both SVO and RCs, we observed increasing reading times at the critical nouns. In BA constructions, however, we observed decreasing reading times at the critical verbs. We thus have some evidence that it is the critical word at the end that lead to differences in reading times.

Despite the fact that our findings support our hypothesis, we need to point out one concern: the use of prepositional phrases and adverbial phrases in front of the critical words. Prepositional phrases are used in simple SVO and RCs while adverbial phrases are used in BA constructions. Since it is not possible for us to use adverbial phrases to modify the head noun or to use prepositional phrases to modify the verb, we did not have a counterbalance condition to rule out the possibility that results in our experiments are entirely due to the critical words. There is a possibility that the difference is due to the phrases used. However, since previous studies have found similar results using different kinds of phrases in front of the critical verbs, including adverbial and prepositional phrases and they still found decreasing reading times at the critical verb, we have reasons to believe that our results are due to verbs and nouns as the critical words, instead of due to prepositional or adverbial phrases.

REFERENCES

BABYONYSHEV, MARIA. AND GIBSON EDWARD. 1999. The complexity of nested structures in Japanese. *Language*, 75: 423-450.

LIN: LOCALITY VERSUS ANTILOCALITY

- CARLSON, GREG, AND TANENHAUS, MICHAEL. 1988. Thematic roles and language comprehension. In W. Wilkens (ed.). *Syntax and semantics* (Vol. 21). New York: Academic Press, Inc.
- FRAZIER, LYN. 1987. Sentence processing: A tutorial review. In M. Coltheart (ed.), *The Psychology of reading. Attention and performance XII* (pp. 559-586). Hove, London: Erlbaum.
- GARNSEY, SUSAN., PEARLMUTTER, NEAL., MYERS, ELIZABETH., AND LOTOCKY, MELANIE. 1997. The contribution of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37: 58-93.
- GARNSEY, SUSAN. AND TANENHAUS, MICHAEL, AND CHAPMAN, R. M. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*. 18: 51-60.
- GIBSON, EDWARD. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68: 1-76.
- GIBSON, EDWARD. 2000. The dependency locality theory: a distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, and W. O'Neil, (eds.), *Image, language, brain* (pp. 95-126). Cambridge, MA: MIT Press.
- GORDON, PETER, HENDRICK, RENDALL. AND JOHNSON, MARCUS. 2004. Effects of noun phrase type on sentence complexity. *Language AND Cognitive Processes*, 51: 97-114.
- GORDON, PETER. HENDRICK, RENDALL. JOHNSON, MARCUS. AND LEE, YOONHYOUNG. 2006. Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory AND Cognition*, 32: 1304-1321.
- HAWKINS, JOHN. 1994. A performance theory of order and constituency. Cambridge: Cambridge University Press.
- HAWKINS, JOHN. 2004. *Efficiency and complexity in grammar*. Oxford: Oxford University Press.
- LIN, YOWYU. 2005. Word order, animacy, and agreement cues in sentence processing by L1 Mandarin EFL learners. Presented at the *Midwestern Conference on Culture, Language, and Cognition*, Northwestern University, IL.
- LIN, YOWYU. AND GARNSEY, SUSAN. 2009. The contributions of classifiers and pronouns to Mandarin relative clause comprehension. Presented at the 22nd *Annual CUNY Conference on Human Sentence Processing*, Davis, CA.
- O'GRADY, WILLIAM. 1997. *Syntactic development: The acquisition of English*. Chicago: University of Chicago Press.
- TANENHAUS, MICHAEL, CARLSON, GREG, AND TRUESWELL, JOHN. 1989. The role fo thematic structures in interpretation and parsing. *Language and Cognitive Processes*, 4: 211-234.

The psychological reality of phonological representations: The case of Mandarin fricatives

Yu-an Lu
Stony Brook University

This study investigates the psychological reality of the phonological representations of two Mandarin fricatives [s] and [ç] by comparing these sounds with a contrastive sound [f], and with the same sounds in Korean in which the two sounds are allophonic. Three tasks were employed to test contrast/allophony in perceiving and processing [ç] and [s]: discrimination on a continuum, similarity rating, and semantic priming. Taken together, the categorical perception on a [s]-[ç] continuum, and the phonemic-like judgment on the similarity rating task suggest that the relationship between [s] and [ç] are more phonemic than allophonic. The different results from the semantic priming task are due to a different level of processing.

1. Introduction

Mandarin Chinese has three palatals [tç, tç^h, ç] that are in complementary distribution with the velars [k, k^h, x], the dentals [ts, ts^h, s], and the retroflexes [tʂ, tʂ^h, ʂ]. The palatals co-occur with the high-front-vowels [i, y] and glides [j, ɥ] while other series do not (Duanmu 2007), as shown in (1).

(1) COMPLEMENTARY DISTRIBUTION OF MANDARIN FRICATIVES

tç, tç ^h , ç	before [i, y] or [j, ɥ]
ts, ts ^h , s	
k, k ^h , x	before non-high-front vowels/glides
tʂ, tʂ ^h , ʂ	

Several hypotheses on the representations of the palatals are therefore proposed as to which series they should be identified with. Chao (1934) uses data from word games to argue that [tç, tç^h, ç] should be underlying /k, k^h, x/¹, while Hartman (1944) and

* I thank Ellen Broselow, Marie Huffman, and the participants in NACCL 2011 for all their comments and ideas. Special thanks to Yuwen Lai from National Chiao Tung University for making the experiment-running possible in Taiwan. This work is supported by the National Science Foundation grant BCS 07460227 to Ellen Broselow, Marie Huffman, and Nancy Squires, and by Chiang Ching-Kuo Foundation Dissertation Fellowship.

¹ Chao (1934) reports on a set of word games that appear to show [k]~[tç] alternation.

- (i) t^ha → t^hai ka
 k^huŋ → k^hwai kuŋ

Duanmu (2007) argue that [tɕ, tɕ^h, ɕ] should be identified with /ts, tsH, s/ from an etymological point of view and phonotactic restrictions, respectively. On the other hand, Cheng (1968) argues that some instances of [tɕ, tɕ^h, ɕ] should be identified with /ts, tsH, s/, and some instances should be identified with /k, k^h, x/ etymologically, but Cheng (1973) later argues that [tɕ, tɕ^h, ɕ] should be independent/underlying segments.

Wan (2010) conducted four experiments, onset similarity, sound contraction, sound similarity, and sound expansion, aiming to investigate the psychological status of palatals.² The results show that there was an asymmetrical response from Mandarin native speakers in terms of palatals and dentals with respect to the other series, and thus suggest that palatals should not be independent, and should be identified with dentals. However, two of these experiments used real words (words that are legal phonotactically) as stimuli, and employed tasks that directly compared the similarity among palatals, dentals, velars, and retroflexes. One can argue that the results only show that the palatals are perceptually more similar to the dentals than the other series.

The current study investigates the palatal [ɕ] and the dental [s], two sounds that are suggested to be underlyingly related in Wan (2010). To avoid direct comparison with the other series in complementary distribution, this study uses a contrastive sound, [f], as a comparison in Mandarin. If [ɕ] and [s] are perceived not different from those with [f], then [ɕ] and [s] should be considered different categories. This study also compares [ɕ] and [s] in Mandarin with the same pair of sounds in Korean, in which the two sounds are allophones. If Mandarin listeners' perception of [ɕ] and [s] are similar to Korean listeners' perception of the same sounds, then [ɕ] and [s] should be considered variants of the same category.

Three previously established methods of testing speakers' perception and processing of sounds are employed to test contrast/allophony in perceiving and processing [s] and [ɕ], discrimination on a continuum, similarity rating, and semantic priming. The results suggest that the perception of [s] and [ɕ] is similar to the perception of separate categories, and thus suggest that [s] and [ɕ] should be of different categories.

This paper reports the results from the three experiments that suggest [ɕ] and [s] are perceived as separate categories in sections 0, 0, and 0, respectively, and a conclusion is provided in section 0.

2. Experiment I – Discrimination on a continuum

Studies of discrimination on a continuum show that speakers discriminate better when sounds are categorically/phonemically different in their native language (Best et al. 1988; Lasky et al. 1975; MacKain et al. 1981; Werker & Lalonde 1988). Werker & Lalonde (1988) show that Hindi speakers identified three categories (two categorical boundaries) from an eight-step continuum from [p^ha] to [ba], and when presented with pairs from the same continuum, Hindi speakers discriminated better when the pairs

ɕjʊŋ	→	ɕjʊŋ tɕjʊŋ
liŋ	→	liɕ tɕiŋ

² The participants in these four tasks were instructed to choose one from the three items played acoustically to them per trial. The tasks were to see, in the absence of palatals, which series, dentals, velars, or retroflexes, the participants would choose to replace palatals.

crossed the categorical boundaries, consistent with the Hindi three-way contrast of stops (aspirated, unaspirated voiceless, voiced). On the other hand, when English speakers were presented with the same continuum, only two categories were identified (one categorical boundary), and only discriminate better when the sound pairs crossed the categorical boundary, reflecting the two-way contrast of stops in English (voiceless, voiced).

2.1. Methodology

Following up on the previous research, this experiment was designed to see how Mandarin listeners behave when presented with a continuum from [s] to [ç].³ I used an eight-step synthesized continuum from [s] to [ç] to test Mandarin listener's discrimination of pairs along the continuum.⁴ Another eight-step continuum from [f] to [s] was synthesized as a comparison. If there is a categorical boundary present only for the [f]-[s] continuum, and no boundary present for the [s]-[ç] continuum, it suggests that [s] and [ç] are perceived as variants of the same category. If there is a boundary present for both continua, it suggests that [s] and [ç], just like [s] and [f], are perceived as separate categories.

The experiment was in ABX discrimination paradigm. 12 two-step pairs (6 pairs from each continuum) in four orders (ABB, ABA, BAA, BAB) were presented in different random order for each participant, using E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA). 20 Mandarin listeners took part in the experiments in a group of up to four people, or individually, using a computer that was connected to a keyboard with two keys labeled '1' and '2.' In the experiment, participants were instructed in Mandarin that they would hear three sounds in one trial, and they were asked to judge if the 1st sound was the same as the 3rd sound, or the 2nd sound was the same as the 3rd sound. The participants completed a 10-trial practice, and had the opportunity to ask questions before proceeding to the experiment.

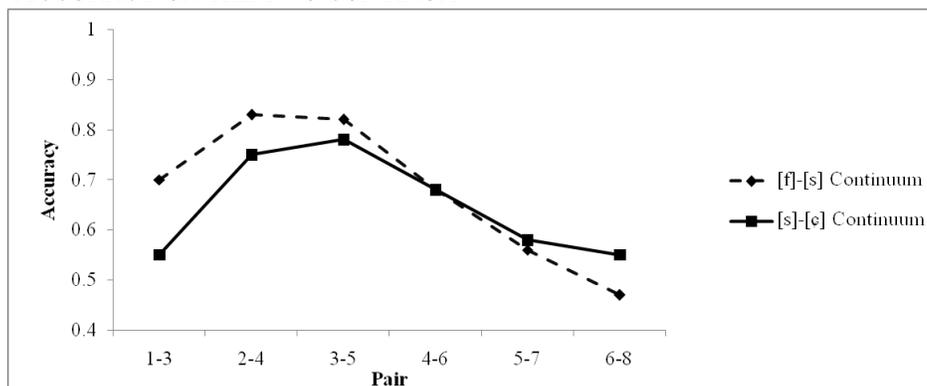
³ Due to the restrictions of the phonotactics of Mandarin, [s] and [ç] cannot be compared in identical vowel contexts. Therefore, I used just the frication part from syllables [si] and [çi].

⁴ The endpoints [s] and [ç] were spliced out using Praat software package (<http://www.praat.org>) from natural productions of [si] and [çi]. The endpoints were synthesized proportionally to create an eight-step continuum using Audacity (<http://audacity.sourceforge.net/>). Step 1 was created by overlap 7 tracks of the endpoint [s], and step 2 was created by overlapping 1 track of endpoint [ç] with 6 tracks of endpoint [s], etc. The intensity of the stimuli was scaled to 56 dB, the averaged intensity of the endpoints [s] and [ç]. The continuum of [f]-[s] was created similarly.

2.2. Results

The accuracy of discrimination on the two continua is illustrated in (2).

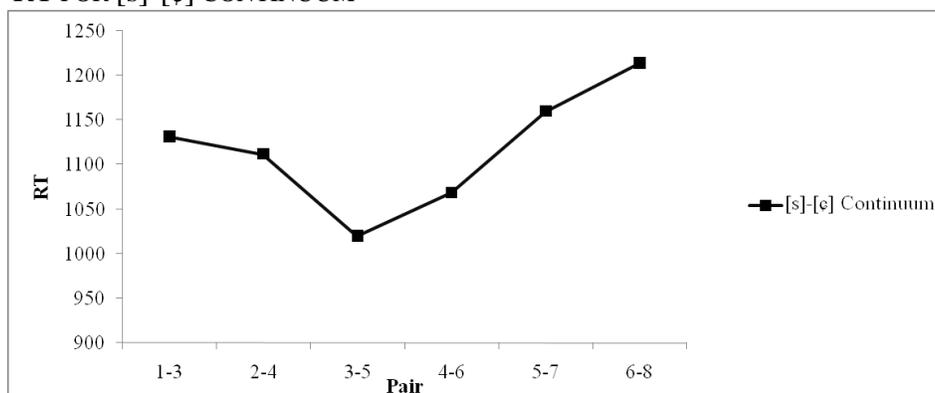
(2) ACCURACY ON THE TWO CONTINUA



As we can see from the line movements in (2), there was a boundary present in both continua around pairs 2-4 to 3-5 for [f]-[s] continuum, and around 2-4 to 4-6 for [s]-[ç] continuum. A repeated measure analysis confirmed this observation. For [f]-[s] continuum, there was a main effect of PAIR ($F(5,95)=22.149$, $p<.001$), which means that the accuracy was not the same for all the pairs. Pairwise comparisons showed that pairs 2-4 and 3-5 were not significantly different ($p=.859$), but pairs 1-2 and 2-4, and pairs 3-5 and 4-6 were significantly different (both $p<.001$). This suggests that the perceptual boundary is around pairs 2-4 and 3-5. Similarly, a repeated measure analysis was run on the [s]-[ç] continuum. There was also a main effect of PAIR ($F(5,95)=9.610$, $p<.001$). Pairwise comparisons among pairs 2-4, 3-5, and 4-6 were shown not significant (all $p>.05$). On the other hand, pairs 1-2 and 2-4, and 4-6 and 5-7 were significantly different (both $p<.05$).

The results from the response time (RT) also suggest that there was a perceptual boundary around pairs 2-4 to 4-6 on the [s]-[ç] continuum, as shown in (3).

(3) RT FOR [s]-[ç] CONTINUUM



The RTs showed a quadratic trend. The valley (pairs 2-4 to 4-6) corresponded nicely to the peak of the accuracy. In other words, Mandarin listeners took less time and responded more accurately around pairs 2-4 to 4-6. The results for the Mandarin group parallel the finding reported in other studies in which RT serves as a positive function of uncertainty (Pisoni & Tash 1974): when the comparisons of two sounds were across categorical boundary, the RTs were shorter; when the comparisons of two sounds were within a category, the RTs were longer.

The results from accuracy and RT showed that there was a perceptual boundary on the [s]-[ç] continuum, just like on the [f]-[s] continuum. This suggests that the perception of [s] and [ç] is not different from that of two contrastive sounds.

3. Experiment II – Similarity Rating

In similarity rating tasks, listeners tend to rate allophones as more similar than phonemes (Babel & Johnson 2010; Boomershine et al. 2008; Johnson & Babel 2010). Boomershine et al. (2008) tested native English and Spanish speakers' similarity judgments of [ð], [d], and [P] using an AX paradigm. [ð], and [d] are contrastive in English but allophonic in Spanish while [d] and [P] are contrastive in Spanish but allophonic in English. Participants were asked to rate the similarity between a pair of sounds that they just heard from the following VCV sequences, [ada], [aPa], [aða], [idi], [iPi], [iði], [udu], [uPu], and [uðu], on a scale of 1-5, in which 1 was 'very similar' and 5 was 'very different'. The results show a clear native language effect with English speakers rating [d] and [P] as most similar, but Spanish speakers rating [ð] and [d] as most similar.

3.1. Methodology

This set of experiments was designed to investigate how Mandarin and Korean listeners rate the target sounds, [s] and [ç]. Korean is chosen to be the comparing language because [s] and [ç] display complementary distribution and take part in rich morphological alternations. We expect a more similar rating between [s] and [ç] for Korean listeners due to their allophonic status. The goal of this experiment is to see how Mandarin listeners rate the similarity between the two target sounds. If the ratings are comparable to those of Korean listeners, then [s] and [ç] are allophonic, just like those in Korean. If the ratings are not comparable to those of Korean listeners, then [s] and [ç] are perceived as different categories.

12 disyllabic VCV stimuli were used in this set of experiments. They were composed of the target fricatives [s, ç] along with two other fricatives [f, h] as controls, embedded in three vowel contexts [a_a], [i_i], and [u_u].⁵ The tokens were produced by a native Mandarin speaking trained phonetician. The speaker recorded multiple examples of the stimuli with high tone on both syllables. One instance of each VCV was selected as the test item. In order to control the amplitude across tokens, the amplitude was scaled to 65 dB, the rough average of the amplitudes of all the tokens, for each of the tokens. All the tokens were approximately matched on intonation and duration.

20 Korean and 20 Mandarin speakers participated in this experiment. Participants in the Mandarin group were recruited in Taiwan. Participants in the Korean groups were all native speakers of Korean from South Korea, and were recruited in Stony Brook

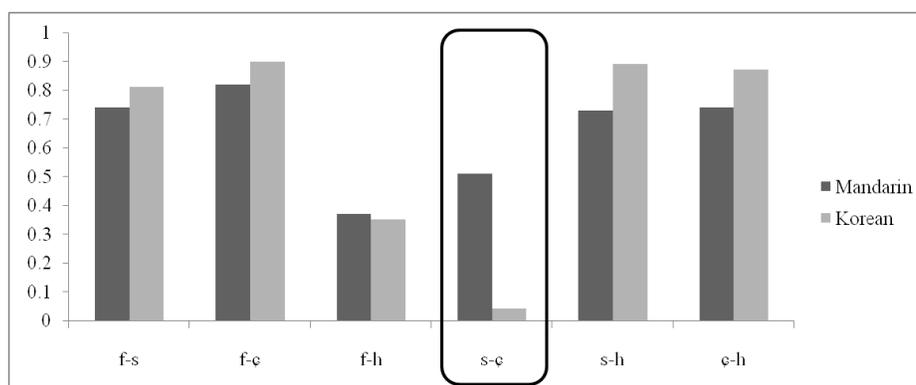
⁵ Materials consisted of two tokens of each of the following VCV sequences: [asa][a]a|[afa][aha], [isi][i]i|[ifi][ihi], or [usu][u]u|[ufu][uhu].

University. They all received up to high school education in South Korea before they came to Stony Brook for undergraduate or graduate education.

3.2. Results

To reduce the variability in analyzing the results of this experiment, the rating scores for each participant were standardized to compensate for differences in using the 5-point scale (Boomershine et al. 2008). The standardized scores were centered around zero, with scores above zero indicating ‘more different’ and scores below zero indicating ‘more similar.’ The results are shown in (4).

(4) SIMILARITY RATING NORMALIZED RESULTS



From the figure in (4), we can see that, except from the target pair [s-ç], the ratings were very similar for the two languages.

A repeated measure analysis (LANGUAGE: Mandarin, Korean; PAIR: [f-s, f-ç, f-h, s-ç, s-h, ç-h]) was performed to interpret the results. The analysis showed that there was a main effect of PAIR ($F(5,38)=73.545, p<.001$). In other words, the pairs were not all rated the same. There was also a significant PAIR by LANGUAGE interaction ($F(5,190)=15.077, p<.001$), meaning that a participant’s response to a given pair was dependent on the language group he/she was in. Simple effect of LANGUAGE in [s-ç] pair was significant ($F(1,38)=36.692, p<.001$), meaning that the ratings of [s-ç] pair from Mandarin and Korean groups were statistically different.

We found that Mandarin listeners rated [s-ç] significantly more different to each other than did Korean listeners. This suggests that Mandarin listeners perceive [s] and [ç] as different categories.

4. Experiment III – Semantic Priming

Previous research has found priming effects between variant pronunciations of a category, but not between sounds belonging to different categories (Ernestus & Baayen 2007; Ranbom & Connine 2007; Sumner & Samuel 2005). In a series of experiments using semantic priming and lexical decision, Sumner & Samuel (2005) found that the target word *music* was primed by the word *flute* articulated with any of the three variants of final [t] (canonical [t], coarticulated [ʔt] and glottalized [ʔ]). However, when the subjects were presented with a contrastive phoneme [flus], no priming effect was shown.

4.1. Methodology

Following up on research showing a priming effect between allophonic variants but not between contrastive sounds, this experiment was designed to investigate the extent to which [s] primed [ç], or vice versa in Mandarin. The results were compared with the priming effects when [s] and [ç] were changed into a contrastive sound [f]. We expect facilitation of lexical decision to a semantically related target to a [s]-prime, or to a [ç]-prime (e.g., [si-jan] ‘breed’ primes [toŋ-wu] ‘animal’). We should find no/less facilitation when [s/ç] are changed into a contrastive sound (e.g., *[fi-jan] should not prime [toŋ-wu]). If we find facilitation when [s] is changed into [ç] or vice versa, then the two fricatives should be variants of the same category (e.g., *[çi-jan] primes [toŋ-wu]). On the contrary, if we do not find facilitation, then [s] and [ç] should belong to different categories (e.g., *[çi-jan] does not prime [toŋ-wu]).

A set of examples is listed in (5), and the complete wordlist can be found in the appendix. Notice that the stimuli for the swapping and contrastive conditions are illegal sequences in Mandarin because of the vowel contexts: [ç] does not occur before non-high-front vowels, and [s] does not occur before high-front vowels.

(5) Stimuli for semantic priming

SAME	[s]-jan] ‘breed’	[toŋ-wu] ‘animal’
	[çi-jan] ‘banquet’	[tçje-hun] ‘wedding’
SWAPPING	*[ç]-jan]	[toŋ-wu] ‘animal’
	*[si-jan]	[tçje-hun] ‘wedding’
CONTRASTIVE	*[f]-jan]	[toŋ-wu] ‘animal’
	*[fi-jan]	[tçje-hun] ‘wedding’

60 Mandarin speaking participants were recruited in Taiwan and were randomly assigned to the three experimental conditions, SAME, SWAPPING, and CONTRASTIVE. In the SAME conditions, [s/ç] were kept unchanged. In the SWAPPING condition, [s] and [ç] were swapped, and in the CONTRASTIVE condition, [s/ç] were changed into a contrastive sound [f]. Among the 72 primes, 36 were connected with related targets, and the other 36 were connected with unrelated targets.⁶ The priming effects were calculated by subtracting the unrelated RTs from the related RTs.

Participants completed the experiment individually or in groups of up to four in a sound-treated booth. Each participant received one experimental list (SAME condition, SWAPPING condition, or CONTRASTIVE condition). Therefore, each participant heard each prime and target only once. All stimuli were presented in a different random order for each participant using E-Prime software. On each trial, participants were presented with an auditory prime, followed by a 500 ms ISI, followed by an auditory target. Participants were instructed to make a lexical decision for the target. Example stimuli were provided and each participant completed a practice session with 8 trials.

⁶ A norming pretest was done to select semantically related word pairs. Three lists of disyllabic [s] and [ç] onset words were put together. These lists were presented to 10 Taiwanese Mandarin speakers. They were instructed to write down a related word for each item. 72 [s] and [ç] onset words were selected from each list as primes (36 [s] onset words and 36 [ç] onset words). The average response of the participants for the related prime-target pairs was 39%.

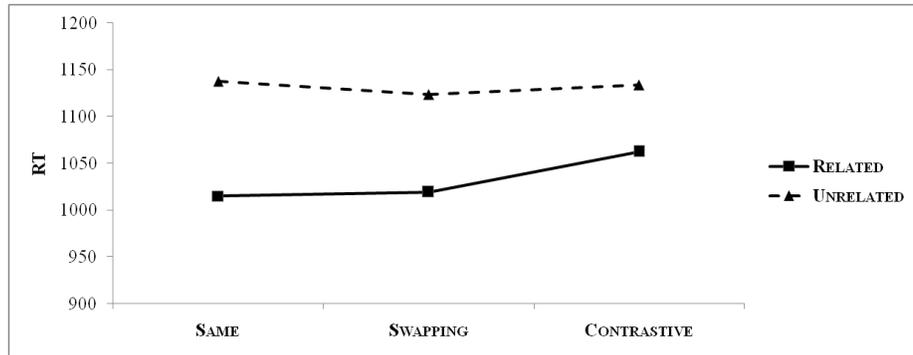
4.2. Results

The mean RTs and the priming effects (the difference between RELATED and UNRELATED) for the three experimental conditions are shown in (6) with standard deviations in parentheses, and are illustrated in (7).

(6) SEMANTIC PRIMING RTs AND STANDARD DEVIATIONS

CONDITION \ RELATION	SAME	SWAPPING	CONTRASTIVE
RELATED	1014.94 (110.34)	1019.19 (106.21)	1062.26 (135.16)
UNRELATED	1137.34 (116.27)	1123.48 (107.22)	1133.46 (139.57)
PRIMING EFFECT	122.4	104.29	71.2

(7) SEMANTIC PRIMING RESULTS



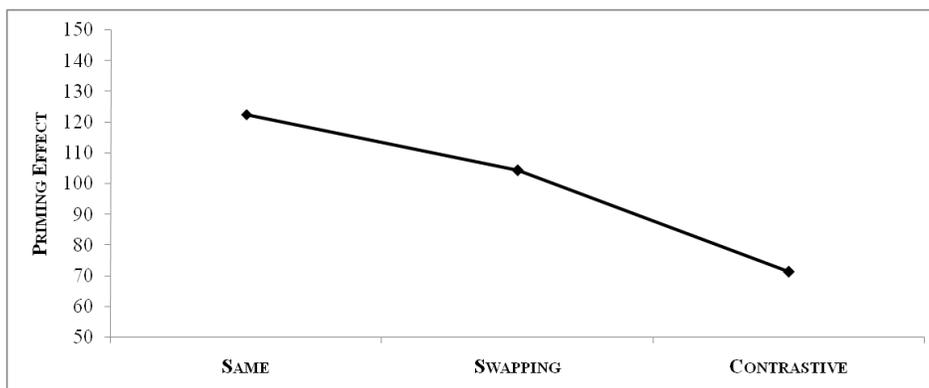
Two-way ANOVAs (CONDITION: SAME, SWAPPING, CONTRASTIVE x RELATION: RELATED or UNRELATED) were performed for subject (F_1) and item (F_2).⁷ Overall, reaction times were significantly faster for related targets than for unrelated targets ($F_1(1, 57) = 171.660, p < .001$; $F_2(1, 210) = 47.836, p < .001$). Planned comparisons showed that targets preceded by related primes were identified more quickly than unrelated primes in all three conditions (SAME $F_1(1, 19) = 85.445, p < .001, F_2(1, 70) = 23.178, p < .001$; SWAPPING $F_1(1, 19) = 44.004, p < .001, F_2(1, 70) = 19.767, p < .001$; CONTRASTIVE $F_1(1, 19) = 53.678, p < .001, F_2(1, 70) = 7.699, p < .01$). Simple effect of condition in RELATED was not significant ($F_1(2, 57) = .985, p = .380, F_2(2, 105) = 2.815, p = .064$) nor does the simple effect of condition in UNRELATED ($F_1(2, 57) = .069, p = .933, F_2(2, 105) = .137, p = .873$).

The factor CONDITION yielded a significant effect in an analysis on priming effects (difference of the RTs in RELATED and UNRELATED in each condition), as illustrated in (8) ($F(2, 117) = 4.356, p < .05$). Pairwise comparisons showed that only the priming effects in

⁷ The results reported here are based on the unscreened data. The screened data (excluding RTs when the lexical decisions were wrong; 253 cases were tossed out), though produced lower standard deviations, yielded the same statistical results.

the SAME & CONTRASTIVE conditions were statistically different ($p < .05$); the other two pairwise comparisons (SAME & SWAPPING, SWAPPING & CONTRASTIVE) were not different (both $p > .1$).

(8) PRIMING EFFECTS



If [s] and [ç] are variants of the same phoneme, we should expect a significant difference between SWAPPING and CONTRASTIVE conditions, but not between SAME and SWAPPING conditions. If [s] and [ç] are contrastive, we should expect a significant difference between SAME and SWAPPING conditions, but not between SWAPPING and CONTRASTIVE. However, the results do not show a three-way difference in RT corresponding to same, allophonic, or contrastive sounds.

4.3. Discussion

The results from semantic priming do not seem to conform to the previous two experiments in which [s] and [ç] are shown to belong to separate categories. A possible explanation for the results is the illegal stimuli in the SWAPPING and CONTRASTIVE conditions. Mandarin listeners might repair the sounds [s], [ç], and [f] in these illegal sequences to a phonetically similar sound that leads to legal sequences. In other words, the priming effects might be due to a repair strategy from the perception.

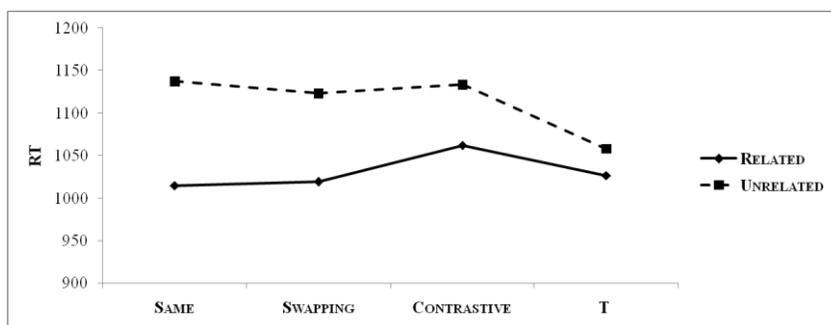
I conducted a follow-up experiment to test this possibility. A fourth condition was added with a contrastive sound that has less phonetic similarity (e.g., [t^h], T condition). Among the stimuli, half of them were illegal sequences, and the other half were legal ones. If the explanation is correct, we should expect an even less priming (or no priming) for the T condition, and we should only find priming effects for the illegal stimuli since Mandarin listeners would repair the illegal sequences, but not the legal ones. The fourth condition, together with the previous three, is listed in (9).

(9) T CONDITION

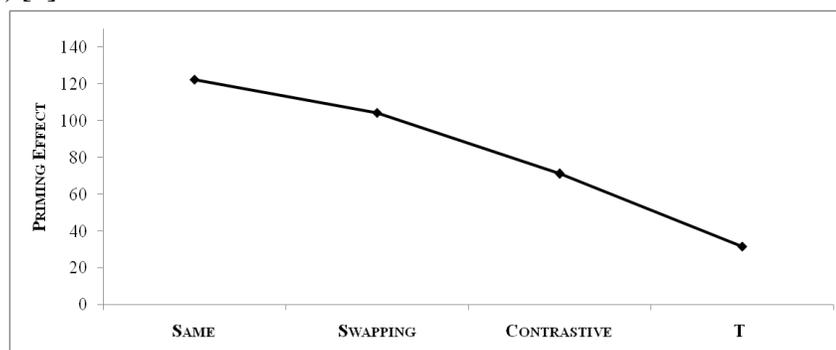
CONDITION	Prime	Target
SAME	[s]-jan] ‘breed’	[toŋ-wu] ‘animal’
	[ɕi-jan] ‘banquet’	[tɕje-hun] ‘wedding’
SWAPPING	*[ç]-jan]	[toŋ-wu] ‘animal’
	*[si-jan]	[tɕje-hun] ‘wedding’
CONTRASTIVE	*[f]-jan]	[toŋ-wu] ‘animal’
	*[fi-jan]	[tɕje-hun] ‘wedding’
T CONDITION	(illegal) * ^h [t]-jan]	[toŋ-wu] ‘animal’
	(legal) [t ^h i-jan]	[tɕje-hun] ‘wedding’

The results are illustrated in (10). Simple effect of relation in T condition was significant ($F(1, 39) = 9.211, p < .005$). In other words, there is still priming in [t] condition. However, the priming effect was significantly less than the other three conditions (all pairwise comparisons $p < .05$), as illustration in 0.

(10) [T] CONDITION RESULTS



(11) [T] PRIMING EFFECTS



The less priming of the fourth condition suggests that Mandarin listeners do perceptually look for a phonetically similar sound to map onto, and that is the reason that

the priming effect of T condition has the least priming effects. If we break the fourth condition into legal and illegal sequences, we see that only the illegal ones have priming (significant simple effect of relation in T-illegal, $F(1,29)=9.173$, $p<.01$) while the legal ones do not ($F(1,19)=2.587$, $p=.124$). This suggests that Mandarin listeners only perceptually repair the illegal sequences.

To summarize this section, at the first sight, the results from semantic priming did not conform to the findings from the previous two experiments. However, after a closer look, the different results reflect a different level of processing in which Mandarin listeners perceptually look for a phonetically similar sound to repair an illegal sequence.

5. Conclusion

This study investigates the psychological reality of the phonological representations of two Mandarin fricatives [s] and [ʃ] by comparing these sounds with a contrastive sound, and with the same sounds in another language. Taken together, the categorical perception on a [s]-[ʃ] continuum, and the phonemic-like judgment on the similarity rating task suggest that the relationship between [s] and [ʃ] are more phonemic than allophonic. The different results from the semantic priming task are due to a different level of processing. The results also shed light on the definition of phonological relationships. Sounds in complementary distribution, like [s] and [ʃ] in Mandarin, need not map onto the same underlying representation, as defined by the traditional structuralist approach. Factors other than distribution need to be taken into account, such as morphological alternation, or phonetic similarity. Future research on how different factors play a role in defining phonological relationships is needed.

REFERENCES

- BABEL, MOLLY & KEITH JOHNSON. 2010. Accessing psycho-acoustic perception with speech sounds. *Laboratory Phonology* 1.179-205.
- BEST, CATHERINE T., GERALD W. MCROBERTS & NOMATHEMBA M. SITHOLE. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu clicks discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14.345-60.
- BOOMERSHINE, AMANDA, KATHLEEN CURRIE HALL, ELIZABETH HUME & KEITH JOHNSON. 2008. The impact of allophony versus contrast on speech perception. *Contrast in Phonology*, ed. by P. Avery, E. Dresher & K. Rice, 143-72. Berlin: de Gruyter.
- CHAO, YUEN-REN. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4.363-97.
- CHENG, CHIN-CHUAN. 1968. *Mandarin phonology*. University of Illinois at Urbana-Champaign.
- CHENG, CHIN-CHUAN. 1973. *A Synchronic Phonology of Mandarin Chinese*. De Gruyter Mouton.
- DUANMU, SAN. 2007. *The Phonology of Standard Chinese*. Oxford University Press.
- ERNESTUS, MIRJAM & HARALD BAAYEN. 2007. Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes* 22.1-24.
- HARTMAN, LAWTON M. 1944. The segmental phonemes of the Peiping dialect. *Language* 20.28-42.

- JOHNSON, KEITH & MOLLY BABEL. 2010. On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* 38.127-36.
- LASKY, ROBERT E., ANN SYRDAL-LASKY & ROBERT E. KLEIN. 1975. VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20.215-25.
- MACKAIN, KRISTINE S., CATHERINE T. BEST & WINIFRED STRANGE. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2.369-90.
- PISONI, DAVID B. & JEFFREY TASH. 1974. Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics* 15.285-90.
- RANBOM, LARISSA J. & CYNTHIA M. CONNINE. 2007. Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57.273-98.
- SUMNER, MEGHAN & ARTHUR G. SAMUEL. 2005. Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52.322-38.
- WAN, I-PING. 2010. Phonological experiments in the study of palatals in Mandarin. *Journal of Chinese Linguistics* 38.157-74.
- WERKER, JANET F. & CHRIS E. LALONDE. 1988. Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24.672-83.

APPENDIX

ç word		s word					
1	xiguan 習慣	19	xiwang 希望	1	sewe 四維	19	sekau 思考
2	xigua 西瓜	20	xiyin 吸引	2	sehu 似乎	20	sewang 死亡
3	xiyou 稀有	21	xicao 洗澡	3	seliao 飼料	21	sefa 司法
4	xixiao 嬉笑	22	xishue 溪水	4	seli 私立	22	seji 司機
5	xiguan1 吸管	23	xire 昔日	5	sepiao 撕票	23	senien 思念
6	xinu 息怒	24	xifong 隙縫	6	seliang 思量	24	segei 賜給
7	xifang 西方	25	xinwen 新聞	7	sefong 祀奉	25	sechu 四處
8	xifu 媳婦	26	xinku 辛苦	8	sechuan 四川	26	sefu 賜福
9	xiyan 喜宴	27	xinren 信任	9	sepuo 撕破	27	sexia 私下
10	xinchu 新竹	28	xingsuo 星座	10	seyang 飼養	28	sejin 絲巾
11	xinshuei 薪水	29	xingxiang 性向	11	semiao 寺廟	29	segue 絲瓜
12	xinguang 星光	30	xinshe 心事	12	sewen 斯文	30	seling 司令
13	xinyang 信仰	31	xintse 刑責	13	sechou 四周	31	sejiao 私交
14	xiyang 夕陽	32	xingyun 幸運	14	seyou 私有	32	seren 私人
15	xihuo 熄火	33	xingmin 姓名	15	sesing 私心	33	sehue 撕毀
16	xinqing 心情	34	xinren1 杏仁	16	senue 肆虐	34	sesin 死心
17	xinshang 欣賞	35	xinxing 新興	17	sezhi 四肢	35	sejiao1 死角
18	xingwei 行為	36	xingwu 省悟	18	semian 四面	36	sesha 廝殺

Chinese as a Foreign Language (CFL) Learners' Representation of Voice in Argumentative Writing

Ying Liu

Qian Du

The Ohio State University

This study investigates how a group of advanced-low level American undergraduate students constructs their voice as they learn to write Chinese argumentative essays in an intensive study-abroad program. We collected students' written essays as well as the sample essays used by the instructor to teach Chinese argumentative writing. We analyzed students' representation of voice in their essays and compared it with the way in which voice is represented in the model essays, focusing on the linguistic devices used in the construction of voice. Our findings suggest that the American learners of Chinese are generally unaware of the importance of constructing their own voice in Chinese argumentative writing, and are in need of explicit teaching of effective linguistic devices that can be used to express their voice.

1. Introduction

Since the early 1980s, when the concept of "voice" was first used by Elbow (1981) to address the question of how writers establish an authorial presence in writing, there has been heated discussion about the application of the concept of "voice" in English L2 writing class. Much of this discussion (Atkinson 1997, Fox 1994, Li 1996, Wu & Robin 2001) argues that such an authorial presence of self in English writing features an individualized world view, which is difficult to achieve for learners from so-called "interdependent cultures" (i.e. East Asian cultures). Along this line, some researchers have proposed that voice exists in all cultures, just as it exists in all the various types of English across the world (e.g., Kachru 1999).

While research of voice has grown tremendously in recent years, such studies in language education have mainly concentrated on the possibility for learners from a collective cultural background to be assimilated into western culture and establish an individualized voice. These studies are illuminating in that they point to the cultural differences of voice construction among writing systems of different languages; however, the comparison is often quite simplistic, as the researchers tend to focus their analysis

on English writing conventions only, and rarely conduct a close examination of the writing systems in so-called collective cultures (e.g. Chinese).

The present study highlights two fundamental problems in current research on voice. First, research related to intercultural differences on voice mainly draws on conclusions from cultural psychology studies, and few studies have examined how voice is presented linguistically in Chinese L1 texts. Second, although many empirical studies have examined English L2 learners' construction of voice in writing, there is a significant lack of empirical research on how CFL learners construct voice in their Chinese writing. To fill in the gap, this study aims to address the following two research questions: (1) What linguistic resources do experienced Chinese writers draw to construct voice in their L1 writing? (2) How do CFL learners establish voice as they learn to write in Chinese?

2. Methodology

2. 1. Participants and data collection

Our study was set in an intensive Chinese summer study-abroad program located in a mid-sized city in East China. The writing course we observed was adopting a genre-based approach, and focusing on argumentative writing at this level (advanced-low). As is stated in the course syllabus, the goal of the writing course was to help students "learn how perspectives, opinions, critiques, etc. are organized and presented in Chinese argumentative writing". Throughout the program, the writing class meets five days a week in the morning, with each meeting lasting for 45 minutes. Students were assigned a specific writing topic and required to write an argumentative essay no less than 800 characters on a weekly basis. In-class activities include: brainstorming on the given composition topics, co-analyzing model texts with the instructor's guidance, and in-class writing exercises. Besides regular class meetings, each student had a 20-minute one-on-one writing tutorial session with the instructor every other afternoon, where they discuss how to revise and improve the drafts.

The participants of our study consist of the course instructor who is a native Chinese speaker with ten years of Chinese teaching experiences, and all of the nine CFL students whose native language is English. Students were recruited from five different American public universities, and thus had different Chinese L2 learning experiences. Only two of them had been to China for a short period for traveling purposes prior to the study-abroad program.

Our data was collected as part of a larger project on CFL learners' learning experience of Chinese argumentative writing. With the permission of participants, we audiotaped the writing class sessions three times a week in order to obtain a general picture of how Chinese argumentation was taught to these foreign language learners. We also collected students' written essays, and interviewed the Chinese-speaking teacher as well as five key participants about their perceptions regarding Chinese argumentative writing. The data we analyzed in this paper is from the second week when students wrote

argumentative essays on the phenomenon of son preference in China. Although our interviews were not directly related to voice, we still obtained much useful information from our interviewees about their perceptions of how they presented themselves in argument construction.

2.2 Data analysis

In order to investigate how linguistic resources are utilized to construct voice by Chinese native speakers, we first analyzed the model essay assigned by the instructor in class. To improve inter-rater reliability, the two researchers, both native speakers of Chinese, first read the model essay separately, and each of us highlighted all the key words, phrases, and sentence structures that we believed were used to express the author's assertiveness and autonomy of his/her arguments. All these tentatively coded data were then cross-checked to ensure that they made sense to both of us. Any inconsistencies between us were discussed until we reached agreement.

We then categorized the codes according to the different types of linguistic devices used by the author of the sample essay to express voice. The focus of our analysis includes lexical items (i.e. first-person pronouns, modal verbs, and content words that express the author's emotions, stances and attitudes), syntactic structures (i.e. rhetorical questions and exclamation sentences), as well as how the author integrates evidence in the argumentative essay to construct voice.

After analyzing the model essay, we then took a similar procedure to code and categorize the nine students' compositions. We highlighted all linguistic devices that express the student author's voice, making notes on the margin of any unit of data that is similar to or different from how voice is linguistically represented in the sample essay. Categories of students' linguistic construction of voice were made by each of us independently and then cross-checked until certain kind of consensus was obtained.

In the following section, we will first present what and how linguistic devices are utilized in the model essay to express the author's voice; then we will compare students' voice construction with the model essay. Relevant pedagogical implications will then be discussed.

3. Analysis of voice in the sample essay

3.1. Linguistic devices used to construct voice

In the sample essay, both lexical items and syntactic structures are used by the author to establish an authorial presence. Lexical items include "I"-phrases that signal the author's forthcoming opinions, modal verbs that express the author's attitudes, emotions, or tones, and various content words that explicitly express the author's emotions and attitudes. Table 1 is a brief summary of the voice-expressing lexical items in the sample essay and their related functions.

Table 1: Lexical items used to construct voice in the sample essay

Lexical Items	Examples	Function
"I"-phrase	我试论, 我想, 我觉得, 我期待, 在我看来 I try to argue, I think, I feel, I hope, in my opinion'	Signal the author's forthcoming opinions
Modal Verbs	可以, 应该, 要 may, should, ought to	Express emotions, attitudes or tones
Content Words: Noun, Verb, Adjective	牺牲品, 痛心疾首, 可笑的 victim, aggrieved, funny	Express emotions, attitudes

Besides the lexical items, syntactical structures are the other type of linguistic resources used to establish the author's voice. Syntactic structures consist of two subcategories: rhetorical questions which are used to convey an emphatic tone, and exclamation sentences that are used to express strong emotions. Examples from the sample essay are:

- (1) 放眼社会, 像我母亲这样的受害女性难道还少吗?
'Are female victims like my mom few in this society?'
- (2) 她们是多么可悲啊! 希望在不久的将来可以看到在家庭中男女平等得到真正的实现!
'How pathetic they are! Hope in the near future we can see sexual equality in families come true!'

The use of rhetorical questions and exclamation sentences, rather than neutral statements, not only expresses the thoughts and ideas of the author, but also conveys strong emotions and attitudes.

3.2. Use of evidence to convey voice

The author's voice is not just embodied in the arguments, but also reflected in the way in which evidence is used and organized. The author's use of evidence allows readers to easily tell the attitudes and tones conveyed in the sample essay, and therefore is by no means neutral or objective.

- (3) 当我妈妈相继生下我姐姐跟我的时候，我奶奶的反应是“怎么是一个女儿”，“怎么又生了一个女儿”，而在我妈妈坐月子时，我奶奶不像别的婆婆一样对媳妇照顾有加，反而是置之不理，甚至是冷嘲热讽。而那些亲戚也只是看看就走，表面寒暄几句而已。

‘When my mom gave birth to my elder sister and me consecutively, my grandma’s reaction was “how come it is a girl,” “how come she had a girl again.” When my mom was in postpartum rest, my grandma was not like other mother-in-laws, who take care of the daughter-in-laws carefully, but ignored my mom, and even humiliated her. Those relatives also just visited and left, having few careless conversations with my mom.’

In example (3), the author uses her mom’s experience as an anecdotal evidence to support her argument about the phenomenon of valuing boys more than girls in Chinese society. Although in this excerpt we cannot see the author’s main thesis, but from the emotional words that she used such as “置之不理” and “冷嘲热讽”, we as readers can still speculate that the author is strongly against the phenomenon of sex discrimination in China. Her mom’s story is not narrated neutrally without any emotions, but is used strategically to express the author’s strong opinions about the injustice of preferring boys to girls.

3.3. Connections between evidence and one’s own opinions

In the sample essay, evidence is not treated as pure isolated “fact”, but is closely weaved into the author’s argument construction. The author always comments on her evidence and makes it clear what her intentions are and why particular evidence is used.

- (4) 而有的男人更可恶，见自己的老婆给自己生的是女儿就对老婆拳打脚踢，要想想，这给女人在心灵上和身体上造成了多大的伤害，女人难道就是生小孩的机器吗？而有的竟还把女婴送给别人抚养，更变态的是还有直接杀死自己亲生女儿的刽子手。可怕的男权制度已经直接伸出了罪恶之手结束还未接触世界的小生命，这种泯灭良知，丧失亲情的做法，何人还能眼睁睁看它存在于这个社会上？

‘But some men are more detestable. They beat and kick their wives when they see their wives bear daughters. Please consider the harm that this kind of behavior brings to women’s mind and body. Are women machines of bearing babies? Some men even have their daughters adopted by others. Some even directly become killers of his own daughter. The terrible patriarchal system has already extended his evil hands to those little lives even before they come to the world. Who can allow such conscienceless behavior to exist in this society?’

Excerpt (4) is an example of how evidence and the author's opinions are interweaved in the sample essay. In this example, the author's descriptions of men's various discriminative behaviors against girls are always followed by her comments and arguments. As a result, readers are clear that the purpose of listing certain evidence here is to express the author's opposition against the patriarchal society that treats girls unfairly.

4. Analysis of voice in students' compositions

4.1. Linguistic devices used by students to construct voice

In this section, we will analyze students' construction of voice in their compositions and compare it with the way in which voice is represented in the model essays. Table 2 summarized the linguistic devices used by students in their compositions.

Table 2: Linguistic devices used by students to construct voice

Linguistic devices	Examples	
"I"-phrase	我觉得, 我发觉, 我认为, 我相信, 在我看来, 对我来说, 从我自己的立场说, 我希望 I feel, I discover, I think, I believe, in my opinion, for me, from my point of view, I hope	(5) 希望在不久的将来可以看到中国家庭中男女平等。 'I hope in the near future we can see sexual equality in Chinese families.'
Modal Verbs	可以, 可能, 会, 不会, 不能 may, probably, will, will not, cannot	(6) 这原因有可能是大城市比较开放, 会受到一些西方的影响。 'The reason probably is that big cities are more open, and more inclined to be influenced by the western countries.'
Content Words: Noun, Verb, Adjective	最基本, 最重要 the most basic, the most important	(7) 家庭被称为社会的最基本及最重要的单位。 'Family is called as the most basic and most important unit of a society.'

Comparing with the model essay, we found that no syntactic structures, such as rhetorical questions or exclamation sentences, were used by students to express their

tones or emotions. Although most students had already learned these structures in their grammar classes, they still had problems applying these structures to serve their voice-constructing purposes.

On the other hand, “I”-phrases were used by most students correctly. One student mentioned in our interview that the instructor’s handout provided them with a list of “I”-phrases which she found very helpful. In example (5), a student directly borrowed a phrase from the model essay “我希望” to express his wishes. As for modal verbs, most students were able to use them correctly on a grammar level. But when we further interview them about their intentions, we found major mismatches between students’ intentions and the native-speaking readers’ interpretations. As in example (6), this sentence is the conclusion of the whole composition which represents the student author’s thesis. However, the use of “可能” greatly weakened the strength of the authors’ argument. The third type, content words, was rarely found in students’ compositions to express their voice, which may be due to their limited vocabulary knowledge as a Chinese foreign language learner.

4.2. Students’ use of evidence

In section 3.2, we discussed that in the model text, evidence is described in such a way that readers can easily interpret the author’s stances, attitudes and perspectives. However, few students were aware of such a skill when they were writing their argumentative essays. Most students tend to adopt a neutral tone as they describe their evidence, which makes it hard for readers to tell their attitudes from the evidence.

- (8) 我祖父母曾经说他们要靠我弟弟来传下梁家的姓。他们的原因是因为梁家所有的女生嫁出去的时候会把姓改成丈夫的，所以只会留我弟弟来传下梁家的姓名。
 ‘My grandparents once said they would rely on my younger brother to pass down our family name of Liang. Their reason is that when girls in our family get married, they would adopt their husbands’ family names. So my younger brother is left to pass down our family name of Liang.’

The evidence in example (8) is quite similar to example (3) from the model essay. Both the student and the author of the model essay used their own experiences as anecdotal evidence. However, while the author of the model essay used emotional words such as “冷嘲热讽” to express her strong emotions against her grandma’s behavior, we could not tell the student author’s attitude and emotions from her use of evidence in example (8). The student’s experience is narrated as pure fact in a neutral tone with no authorial perspective conveyed. The same problem exists in example (9) as well.

- (9) 一位美国律师 Christine M. Bulger 在 2000 年写了一个文章题为《中国职场上斗争男女歧视》，文章讨论职场上的计划生育的问题，讲述 Sun Lili 的描述。Sun

Lili 怀孕的时候没有计划生育的批准，已经流产两次，她跟丈夫决定生孩子。因为她没有批准，所以她的雇主解雇她。她尝试上诉雇主的行为，但全体审判员维护雇主的决定。

'An American lawyer Christine M. Bulger wrote an article, "Gender Discrimination in Chinese Professional Fields" in 2000. This article discussed the effects of one-child policy on professional fields. According to Sun Lili's description, she had two abortions because she had no birth permissions. But she and her husband still decided to have a baby. Her employer fired her because she didn't have the birth permission. She tried to sue her employer, but all the judges affirmed her employer's decision.'

In example (9), the student was using a story from a newspaper article as her evidence to explain how unfair it is for the company to fire the woman employee simply because of her child-bearing needs. However, the student made no effort to adapt the newspaper story to help construct his own argument; thereby the evidence still keeps its original newspaper style. All words in this excerpt are relatively neutral and objective with no emotions of the student author expressed. Without reading the rest of his article, readers could hardly hear this student's voice.

4.3. Lack of connection between students' opinions and use of evidence

When comparing students' compositions with the model text, we also noticed that the connection between evidence and authors' arguments is different. In the model text, the author connects her evidence with his argument closely, and always comments on the evidence and makes it clear why she uses certain evidence in a particular place. However, for our students, the evidence and arguments are often loosely connected. Sometimes, it is quite difficult for readers to understand why certain evidence is used in the essays.

(10) 因此做完这个调查采访我认为中国是重男轻女的国家，但是现代化的年轻人中，这个思想开始慢慢的消失，特别是在比较大的城市。这原因可能是大城市会比较开放，会受到一些西方的影响。

'So after this survey I think China is a country valuing boys more than girls. But for the modern young people, this thought gradually begins to disappear, especially in big cities. The reason probably is that big cities are more open, and more inclined to be influenced by the west.'

Excerpt (10) is the conclusion part of a student's composition. Before writing, the student first conducted a survey to several native Chinese speakers in order to gather some background information about son preference in China. In her composition, she first summarized her survey results in several paragraphs with a neutral tone, and

postponed her own arguments to the last paragraph. The problem of her organization is that the evidence is purely listed without enough comments. And the thesis is left isolated in the end with no support. If we further examine the relationship between her evidence and conclusions, we will see that the evidence she used in the first several paragraphs does not support her core argument in the end about how the phenomenon of boy preference is gradually disappearing among young people in big cities. The lack of connections between students use of evidence and their voice construction is a common problem for most student participants we observed.

5. Pedagogical implications

Based on our preliminary analysis of a sample argumentative essay and students' compositions, it is not unreasonable to conclude that the advanced-low CFL learners in our study do not seem to have an awareness of constructing a clear voice in their writing. Although most students did learn about possible linguistic devices (e.g. "I"-phrases, rhetorical questions, exclamations questions) in grammar classes, they are generally unsure of what purposes can be achieved through the use of these linguistic resources. Consequently, the student authors were not quite capable of expressing their arguments about a controversial issue in writing even though they had strong opinions. Therefore, it is critical for the foreign language instructor to step in and introduce the rhetorical functions of certain phrases and structures to help learners apply them in their own writing to construct voice.

When incorporating the concept of voice in curriculum design, CFL instructors can strategically use model essays to raise students' awareness of voice in argumentative writing. The analysis we did here is for research purposes, but the same type of analysis can also be helpful for instructional purposes. If students are able to notice and understand how experienced Chinese writers utilize certain linguistic resources to construct voice, they are more likely to be successful in imitating what the experienced authors do in their own writing. In our data, it is clear that students are aware of the importance of learning from the model text and borrowing words and phrases, yet they are generally unable to tell the purposes and objectives of the use of certain linguistic devices that they have chosen to borrow. In such a situation, the instructor can play a key role in facilitating CFL learners as they try to identify the rhetorical purposes that can be achieved through the use of particular syntactic and lexical devices.

References

- ATKINSON, DWIGHT. 1997. A critical approach to critical thinking in TESOL. *TESOL Quarterly* 31. 71-94.
- ELBOW, PETER. 1981. *Writing with power*. New York: Oxford University Press.
- FOX, HELEN. 1994. *Listening to the world: Cultural issues in academic writing*. Urbana, IL: National Council of Teachers of English.
- HINKEL, ELI (ed.). 1999. *Culture in second language teaching and learning*. Cambridge: Cambridge University Press.
- KACHRU, YAMUNA. 1999. Culture, context, and writing. In Hinkel, 75–89.
- LI, XIAOMING. 1996. "Good writing" in cross-cultural contexts. Albany, NY: SUNY Press.
- WU, SU-YUEH, & RUBIN, DONALD. 2001. Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English* 35. 148-78.

The Acquisition of Some Properties of the BA Construction by English-speaking Learners of Chinese

Hongying Xu
University of Kansas

There have been numerous studies on the linguistic properties of the BA construction in Mandarin Chinese and the semantic meanings of this construction. However, there have not been as adequate studies on the acquisition of the BA, especially by L2 Chinese learners. This study examines the acquisition of some linguistic properties of the BA construction by the nature of those properties: some are pure syntactic, some involve knowledge from the interface of two domains: syntax and semantics. The results indicate that L2 Chinese learners' performance on properties in the core syntax domain is significantly better than on those at the syntax-semantics interface. However, this advantage is not consistent cross the core syntax domain. L2 Chinese learners' performances within the core syntax domain are significantly correlated but at different accuracy levels.

1. Introduction

The BA construction is a unique yet important construction in the grammar of Mandarin Chinese. It is a very commonly used construction in Chinese people's daily speech. The BA construction was also listed by the well-known Chinese linguist L. Wang as one of the five major changes in Chinese grammar over history (1958). This construction has aroused huge research interest among linguists who are interested in Chinese grammar. There has been numerous research on the linguistic properties of the BA construction, with every possible element of this construction being examined. However, there has always been debate over those issues due to some "difficult-to-characterize, fuzzy" constraints that this construction has (A. Li, 2001, p.1). This construction has also imposed enormous difficulties when it is acquired by L2 Chinese learners. In spite of its notorious difficulties being well-observed by the instructors and learners, there has not been adequate research on the acquisition of the BA construction by L2 learners. Previous studies examined this issue either focusing on a subsection of the BA sentences or some properties/constraints without giving any justification why certain properties were chosen but others were not. This study examines the acquisition of some properties of the BA construction by the nature of those properties. This perspective was inspired by the Interface Hypothesis proposed by Sorace (2005, 2009), who proposes that the final attainment of a property in the target language varies in terms

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

of the nature of the property. The properties that belong to the pure syntax domain are fully acquirable in L2 acquisition, especially in the final attainment. However, the acquisition of the properties which require the integration of knowledge from syntax and other domain is more problematic. It presents “residual optionality in L2” (Sorace, 2005, p.56).

The goal of this study is to compare L2 Chinese learners’ knowledge on some properties of the BA construction which belong to the “core” syntax domain with other properties which belong to the interface of syntax and semantics. By looking at the BA properties in terms of their nature, we are more likely to get a more general picture of L2 learners’ difficulties. A better understanding on this issue can further facilitate classroom instruction by making it more targeted.

1.1 Properties of the BA construction

Many studies have examined the linguistic properties of the BA construction. Observations as well as proposals have been made to generalizing the properties and constraints that this construction has. However, despite of numerous proposals, little consensus has been reached. This study will only focus on several properties that are true to most of the typical and frequently-used BA sentences, since the purpose of this study is to identify the particular difficulties encountered by L2 Chinese learners.

In this study, two properties that belong to the core syntax domain and one that belong to the syntax-semantics domain were chosen. Two pure syntax properties were chosen instead of one is to examine if what Sorace proposes as “fully acquirable” properties are equally well-acquired by L2 learners. Some studies (Yuan, 2010, for example) claim that the acquisition of an interface property is not “domain-wide”, but rather “variable-dependent”.

1.1.1 Core syntax properties of the BA construction

Two core syntax properties were investigated in this study. The first is the non-canonical word order of the BA construction (the word order constraint). A regular Chinese sentence takes the SVO order: the subject goes first, followed by a verb phrase, and the object takes the postverbal position. However, in a BA sentence, the noun phrase introduced by BA (the Ba-NP) always goes before the verb. 1)a is an example of a regular Chinese SVO sentence, in which the object *his car* follows the verb phrase *mai le*. However, in a BA sentence, the object *his car* cannot stay in the postverbal position, as in 1)b, but must move to the preverbal position as in 1)c:

- 1) a. ta mai le ta de che.
 He sell Asp Le he POSS car
 ‘He sold his car.’

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

- b. *ta mai le ba ta de che.
 He sell Asp Le BA he POSS car
- c. ta ba ta de che mai le.
 He BA he POSS car sell Asp Le
 ‘He sold his car.’

It is agreed in the literature that the predicate in a BA sentence cannot be a verb by itself (also referred to as a bare verb), it must take another grammatical unit (the X element from now on) (S. Lu, 1980; Sybesma, 1992; Liu, 1997; A. Li, 2001). This constraint is referred to as the complex verb constraint in this study. The postverbal X element in a BA sentence can be realized in different forms. One of the most typical forms that the X element takes in a BA sentence is a resultative complement, which “refer to the state, degree, accomplishment, achievement, or effect of the action” (Shi, 2002, p. 29). Another form that the X element frequently takes is a propositional phrase to denote destination. A propositional phrase like this usually follows a verb which denotes displacement of objects, as the following example illustrates:

- 2) ta ba shu fang zai zhuo shang.
 He BA book put at desk on
 ‘He put the book on the desk.’

In 2), the proposition *zai* takes *zhuo shang* to for the complement of destination, and goes with the displacement verb of putting. In the tokens that target at the complex verb constraint, the X element only takes the form of a resultative complement or a propositional phrase.

1.1.2 Semantic property of the BA construction

Although most BA sentences have a non-BA counterpart, that is, a regular SVO sentence, they are not interchangeable under certain contexts. The BA construction implies some semantic properties and there are semantic constraints for a BA sentence to be acceptable. Like there is no agreement on the forms that the X element can take, there has also been debate on the semantic properties that the BA construction indicates. Some of the most well-known accounts include the disposal account, the causation account, and more recently, the displacement account. This study will stick with the affectedness account (A. Li 2001), which modifies the original disposal account and regards the BA construction as a “highly transitive” construction, in which the verb extends “a high degree of affectedness” on the object (p. 46), as the following example illustrates:

- 3)a. wo ba guozi zhuang le shui.
 I BA pot fill ASP Le water

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

‘I filled the pot with water.’

- b. wo ba shui zhuang zai guozi li.
 I BA water fill at pot in.
 ‘I filled the water into the pot.’
 (example from A. Li, 2001, p. 46)

In 3)a., the pot, which is the BA NP, is affected in the sense that now it has water; whereas in 3)b., the water, which is the BA NP, is affected in the sense that its location has changed by being in the pot now.

1.2 Previous studies on the acquisition of the BA construction by L2 learners

Compared with the rich literature on the linguistic properties of the BA construction, studies on the acquisition of the BA construction by L2 Chinese learners are scarce.

In her study, Jin (1992) examined L2 Chinese learners’ performance on the BA sentences in terms of the definiteness of the BA NP, the topicality of the BA NP, and the disposability of the VP. Her study indicates that there are clusters of BA sentences in terms of their learnability to L2 Chinese learners. Some BA sentences are acquirable to L2 Chinese learners cross all levels. On the contrary, some BA sentences seem to be beyond L2 Chinese learners at all levels in her study. Other BA sentences are acquired better with learners at higher proficiency levels. This variation in the BA acquisition is insightful.

Zhang (2002) examined the developmental process in the acquisition of the BA construction by L2 Chinese learners by testing their performance on six properties of the BA construction. She claims that there is a U-shaped developmental pattern in the BA acquisition, which is observed across all properties. An acquisition order of those properties was also proposed. However, there is no justification on the choice of these properties. Nor is there any explanation on the role each property plays in the BA construction.

Du (2004) examined the acquisition of two specific constraints on the BA construction by testing L2 learners’ comprehension and production of the BA construction. The results indicate that the L2 Chinese learners performed better in accepting grammatical BA sentences than in rejecting ungrammatical BA sentences. L2 Chinese learners produced fewer BA sentences than native speakers of Chinese, however, their production achieved a high accuracy.

In her cross-sectional study, Wen (2010) examined L2 Chinese learners’ performance on a sub-section of BA sentences (the displacement BA sentences). The results indicate that both frequency and accuracy of the BA construction usage increase as the proficiency level increases.

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

Quite different from what the scholars in the U.S. have done, scholars in Mainland China mainly investigated the acquisition of the BA construction by looking at L2 Chinese learners' performance at different types of BA sentences listed by linguists in Chinese. There are not many experimental or quasi-experimental studies. Many studies examined corpus-based data or data collected from learners' essays, exams and so forth. These error-analysis type of studies provides us with different pieces of the whole picture. However, a more systematic way is called for to gain a more comprehensive understanding of the issue.

The present study examined the acquisition of three specific constraints on the BA construction, namely the word order constraint, the complex verb constraint, and the affectedness on the BA NP constraint. These constraints differ in nature in terms of the knowledge they call for. The first two call for syntactic knowledge, whereas the last one calls for the knowledge from syntax and semantics. The specific research questions that this study examined are: 1) How well have the L2 Chinese learners acquired the core syntax properties and the interface properties at syntax-semantics? Is their performance native-like? Is there any difference between their performances in the two domains? 2) Within the core syntax domain, is L2 Chinese learners' performance consistent?

According to the Interface Hypothesis, the acquisition of the core syntax properties is predicted to be native-like, whereas the acquisition of the syntax-semantics interface is not. It is also predicted there will be a significant difference between these two domains. However, the Interface Hypothesis was proposed on the basis of very advanced level L2 learners or bilinguals. Non-native-like performance in the core syntax domain may be observed since the L2 Chinese learners recruited for this study were mostly at intermediate level.

2. Method

This section introduces the design of this study. First of all, the subjects of this study will be introduced, which is followed by an introduction of the method and procedure used to collect data. The data coding criteria will also be introduced.

2.1 Subjects

There are two groups of participants involved in this study: English-speaking learners of Chinese (the L2 group hereafter), and native speakers of Chinese (the NS group hereafter). The L2 group is the target of this study, and the NS group serves as the control group, which provides a baseline of comparison.

Thirty-two English-speaking learners of Chinese participated in this study. The participants were recruited on a voluntary basis from students who were learning Chinese at a large university in the mid-west at the moment when this study was conducted. Most of them were second-year or third-year Chinese learners. Proficiency is not controlled as a factor, however, a proficiency test was given to all of the L2 group participants. A survey was conducted among the L2 group on their background information as learners

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

of Chinese, such as the age of onset of Chinese learning, length of Chinese learning, length of residence in Chinese-speaking communities, out-of-class exposure to Chinese, and so on. A brief summary of participants' information is presented in Table 1.

Table 1: Background Information of the L2 Group Participants

Demographic features	Mean (range in parenthesis)	S.D.
Gender	F=16; M=16	n/a
Age	22.7 (19-40)	4.2
Age of onset of learning	19.6 (14-27)	3.1
Length of learning	3.1 years (1-8)	4.2
Residence in a Chinese-speaking region	0.5 years (0-7)	1.4

Among all of the L2 participants, 59.4% of them were enrolled in Level II Chinese when this study was conducted; 31.2% of them were enrolled in Level III Chinese; 9.4% were enrolled in Level IV Chinese. Twenty-nine of them (90.6%) started to learn Chinese after 18. Twenty-five of the participants (78.1%) do not have any experience of living in a region where Chinese is spoken as a native language.

Twenty native speakers of Chinese participated in this study. They were undergraduate students at a university in the south-east part of China. The NS group participants were also recruited on a voluntary basis. The proficiency test was also given to the NS group participants so that it could provide a baseline for comparison. The NS group participants were asked to finish a brief survey on their experience of using Mandarin Chinese if they speak another dialect. All the NS data was collected in that university in China.

2.2 Materials

This study employs a grammaticality judgment task to elicit subjects' knowledge on the target properties of the BA construction. This section describes the selection of the tokens and the design of the task.

The grammaticality judgment task was employed for a couple of reasons. First, it makes it possible to collect sufficient data within a short period of time. Secondly, it allows us to tap on subjects' knowledge of a particular property from both sides: to accept grammatical ones and to reject ungrammatical ones.

The grammaticality judgment task used in this study consists of grammatical and ungrammatical sentences which represent the target properties: the word order constraint, the complex verb constraint, and the affectedness on BA NP constraint. There are altogether 80 sentences (tokens) in this task. The distribution of the tokens is shown in Table 2.

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

Table 2. Distribution of Tokens in the Grammaticality Judgment Task

Target properties	Grammatical tokens	Ungrammatical tokens
Word order constraint	8	8
Complex verb constraint	8	8
Affectedness constraint	8	8
Distracters (regular SVO)	16	16

Each constraint is represented by eight minimal pairs of sentences. A minimal pair consists of a grammatical sentence and its ungrammatical counterpart, which is identical with the grammatical one except for the target property. For example, a minimal pair that targeted at the word order constraint is as follows:

- 4)a. ta ba wo de che xiu hao le.
 He BA I POSS car fix good LE
 ‘He fixed my car.’
- b. *ta xiu hao ba wo de che le.
 He fix good BA I POSS car LE

4)a. and 4)b. use exactly the same phrases. The only difference lies in the word order of the BA NP and the verb phrase. If a subject correctly accepts the grammaticality sentence and rejects the ungrammatical one in a minimal pair like 4), it is assumed that the target property has been acquired.

Thirty-two Chinese sentences in the canonical SVO order were used as distracters to keep the target properties from being obviously revealed to the subjects. On the other hand, in order not to “overdistract” the participants, those SVO sentences are limited to those with verb complements of potential. Sixteen of the sentences are grammatical sentences with potential complements, and the rest are their ungrammatical counterparts.

The vocabulary used in this task is largely limited to the words the subjects have learned so far. For those words (only a few) that might be new to some subjects were given a brief definition in English.

2.3 Procedures

The tokens in the grammaticality judgment task were put into two blocks, with each sentence of a minimal pair in one block. As a result, each block has 24 target sentences and 16 distracters, with 20 grammatical sentences and 20 ungrammatical ones. For each participant, the sentences in each block were presented in a random order. In order to eliminate the effect of task order, each subject was randomly assigned to do block 1 or block 2 first.

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

The sentences were presented in the form of an untimed slide show by using *Paradigm*, a software program designed for experimental studies in behavioral science. All of the sentences were presented in Chinese characters, with Pinyin above each character. A simplified characters version and a traditional characters version were provided. The subjects got to choose one of them. In doing so, the possible impact from students' unfamiliarity with the characters has been eliminated.

After reading the task instructions, the subjects were also given two trials before the real task began so that they could get familiar with the task. During the task, the subjects read one sentence on the screen, made judgment by clicking a corresponding box, and then the next sentence was shown on the screen. Subjects were forced to take a rest between two blocks, but they were also allowed to take a rest during the process of one block if they requested.

In the grammaticality judgment task, the subjects got one score for each correct judgment they made and got zero for each wrong judgment. Their total score for each target property was divided by possible total score and resulted in a percentage of the accuracy rate.

3. Results

This section presents the descriptive results of this experiment as well as the results of the statistical analysis conducted to get more information from this experiment.

3.1 L2 learners' acquisition of the core syntax properties vs. syntax-semantics interface properties

The results of the L2 group subjects' performance on the core syntax properties and the syntax-semantics interface property in terms of accuracy rate are presented in Table 3. The accuracy rates on the core syntax properties were obtained by averaging the accuracy rates on the word order constraint and the complex verb constraint. The results from the NS group were also presented for comparison.

Table 3. Accuracy Rates (%) on Core Syntax Properties vs. Syntax-semantics Interface Properties (the L2 group (N=32); the NS group (N=20))

Properties	Mean		Range		S.D.	
	<i>NS</i>	<i>L2</i>	<i>NS</i>	<i>L2</i>	<i>NS</i>	<i>L2</i>
Core syntax	98.9	81.0	93.9-100	59.4-96.9	2.1	10.7
Syntax-semantics	97.2	69.5	87.5-100	50-88	4.3	11.4

As the descriptive statistics in Table 3 shows, the NS group in this study scored a mean accuracy rate of 98.9% (SD=2.1) on the core syntax properties of the BA construction. They scored a mean accuracy rate of 97.2% (SD=4.3) on the syntax-

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

semantics properties of the BA construction. The statistics indicate that there is not much variation among native speakers' response to the properties from both the core syntax domain and the syntax-semantics interface domain, though the variation in the latter is a bit larger than the former.

A paired sample T-test was conducted to test if there was a significant difference between the NS group's accuracy rates between the core syntax properties and the syntax-semantics interface properties. The results show that there was no significant difference between the accuracy rates on the two domains [$t(19)=1.68$, $p=.110$].

As the descriptive statistics in Table 3 shows, the L2 group in this study scored a mean accuracy rate of 81.0% (SD=10.7) on the core syntax properties of the BA construction. They scored a mean accuracy rate of 69.5% (SD=11.4) on the syntax-semantic interface properties. Compared with the NS group, the L2 group shows more variation in their responses to the properties from both domains. A similar trend was found among the L2 subjects that they show larger variation in the syntax-semantics interface domain than in the core syntax domain.

A paired-sample T-test was conducted between the L2 group's performance on the core syntax properties and the syntax-semantics interface properties. The results show that the difference between L2 group's performances on these two domains is significant [$t(31)=4.564$, $p<.001$]. In terms of accuracy, the L2 group performed significantly better with core syntax properties than with syntax-semantics interface properties.

In order to examine if the L2 group has achieved native-like performance on the core syntax properties and on syntax-semantics interface properties, two independent-samples T-tests were conducted to compare the L2 group with the NS group. In those tests, the mean accuracy rate scored by each subject on each domain was used as the dependent variable. Levene's test for equality of variance in each independent samples T-test is significant. Therefore, the t values that do not assume equal variance are reported instead of the standard t values. Table 4 summarizes the results of these T-tests.

Table 4. Independent Samples T-tests of accuracy rates between the L2 group and the NS Group

Properties	Dependent Variables	df	t
Core syntax	Accuracy rate	34.72	9.19*
Syntax-semantics	Accuracy rate	42.98	12.26*

*. The t value is significant at the .05 level

As the results of the T-tests show, the differences in accuracy rates between the L2 group and the NS group in both the core syntax domain and the syntax-semantics interface domain are significant at $\alpha=.05$ level. It means that the L2 group has not achieved native-like performance on the core syntactic properties of the BA construction as well as the properties at the syntax-semantics interface.

3.2 L2 learners' performance within the core syntax domain

This study also attempts to examine if L2 Chinese learners' acquisition of the properties that belong to the core syntax domain is consistent. In order to do this, L2 learners' performances on the two properties within the core syntax domain (the word order constraint and the complex verb constraint) were measured separately in terms of their accurate rates on the two constraints. The results are presented in Table 5. The accuracy rates among the NS group is also presented for comparison.

Table 5. Accuracy Rates (%) on Core Syntax Properties (the L2 group (N=32); the NS group (N=20))

Properties	Mean		Range		S.D.	
	NS	L2	NS	L2	NS	L2
Word order	99.7	93.1	93.9-100	75-100	1.4	7.3
Complex verb	98.1	69.1	87.5-100	37.5-100	4.1	17.3

As the descriptive statistics show in Table5, the native speakers achieved an accuracy rate of 99.8% (SD=1.4) on the word order constraint, with little variation. They scored an accuracy rate of 98.1% (SD=4.1) on the complex verb constraint, with a bit bigger variation. A paired-sample T-test was conducted and the results show that there was no significant difference between the NS group's performance on the word order constraint and the complex verb constraint [$t(19)=1.561$, $p=.135$].

The picture of the L2 group is somewhat different. The L2 group achieved an accuracy rate of 93.1% (SD=7.3) on the word order constraint, with some variation. They scored an accuracy rate of 69.1% (SD=17.3) on the complex verb constraint, with big variation among the subjects. A paired-sample T-test was conducted and the results show that the difference between L2 group's performances on the word order constraint and on the complex verb constraint is significant at $\alpha=.05$ level [$t(31)=8.493$, $p<.001$].

An independent-samples T-test was conducted in order to examine if the L2 group has achieved native-like performance in the word order constraint. The results show that the difference between these two groups with their performance on the word order constraint was significant at $\alpha=.05$ level [$t(34.5)=4.92$, $p<.001$]. It means that the L2 group has not achieved native-like performance on the word order constraint. However, as has been pointed out earlier, this finding could be due to the fact that the L2 participants were not advanced enough. The top-25 per cent sub-group within the L2 group scored an accuracy rate of 96.9%.

The L2 group scored similar accuracy rates on the complex verb constraint (69.5%) and the affectedness on the BA NP constraint (69.1%). However, their performance on the word order constraint and complex verb constraint was significantly different. These findings may make people consider that the complex verb constraint, like the affectedness constraint, belongs to the interface domain rather than the core syntax

domain. Correlation tests were conducted to examine if this is the case. The results show that the L2 group's performance on the word order constraint and the performance on the complex verb constraint was significantly correlated at $\alpha=.05$ level [$r=.384$, $p=.03$]. However, the L2 group's performance on the complex verb constraint and the affectedness constraint was not significantly correlated [$r=.268$, $p=.138$].

4. Discussion

In the previous section, the results of the grammaticality judgment task were summarized. The results show that the L2 group's performance on the word order constraint achieved high accuracy, whereas their performances on the complex verb constraint and the affectedness constraint were not as high.

The first research question that this study attempts to examine is the L2 Chinese learners' performance on the properties that belong to the core syntax domain vs. their performance on the properties at the interface of syntax-semantics. The results indicate that there is a significant difference between their performances on the two domains. The L2 Chinese learners performed significantly better on core syntax properties. Their performances in both domains are significantly different from the native speakers of Chinese.

The findings partially conform to what the Interface Hypothesis would predict concerning the acquisition of linguistic properties of the BA construction. The Interface Hypothesis predicts that the acquisition of the syntax-semantics interface properties will not be as good as the acquisition of the core syntax properties, which is supported by the findings of this study. It also predicts that native-like performance on the core syntax properties, which is not supported by the findings of this study. The reason for this could be the proficiency level of the L2 group. The Interface Hypothesis was proposed on the basis of studies on near-native speakers and bilingual speakers. Obviously, the L2 group subjects in this study have not been there yet. A sub-group ($N=8$) from the L2 group, who has scored at the top 25 per cent of the proficiency test, achieved an mean accuracy rate of 92.8% on the core syntax properties. The difference between this sub-group and the NS group is much smaller. No T-tests were conducted due to the small number of subjects in this sub-group. However, it is still reasonable to expect near-native performance regarding the core syntax properties if more advanced L2 Chinese learners were examined.

Another question that this study attempts to find an answer to is whether the L2 Chinese learners' performance within the core syntax domain is consistent domain-wide. The results show that there is a significant difference between their performance on the word order constraint and the complex verb constraint. They did better in the former than in the latter. The findings are, to some extent, similar to the findings in Yuan's (2010) study, which claims that L2 learners' performance is variable-dependent rather than domain-wide in the syntax-semantics domain. On the other hand, the L2 Chinese learners' performances on the two constraints are significantly correlated. The observed difference

in accuracy rates may be attributed to the driving forces of these two constraints. The non-canonical word order is solely driven by a syntactic constraint in Chinese. In Chinese, the elements that can occur postverbally are limited (A.Li, 2001, p.4). When the postverbal position is occupied by an “internal object” (M. Wang, 1987, p. 61), the real object (the syntactic object) is forced to move preverbally. The complex verb constraint rules out those ungrammatical BA sentences when the predicate is a bare verb. An element, such as a resultative complement, is used to indicate the achievement, state, effect, as a result of the action denoted by the verb. This use of a resultative complement is driven by the semantic meaning of the BA construction. In this sense, this constraint leans towards the interface of syntax and semantics than the word order constraint. However, the L2 Chinese learners took it as a syntactic constraint more than an interface one. Moreover, the L2 Chinese learners’ low accuracy rate on this constraint might also be accounted for by their performance on resultative verb compound (RVC). This is only an assumption and needs further research before any connection can be established.

5. Conclusion

The results have shown that in learning the BA construction, the core syntax properties are better acquired by the L2 Chinese learners than the syntax-semantics interface properties, though neither of their performance has achieved native-like level. These findings, in part, support Sorace’s (2005) Interface Hypothesis that L2 learners have a delayed acquisition of properties at interfaces. However, contradictory to what the Interface Hypothesis would predict, the findings of this study also indicate that the advantage of core syntax properties over interface properties is not consistent. Further research is needed to explore the possible reasons for this inconsistency. This study examines the acquisition of the BA construction properties and constraints by looking at the different performances among L2 learners in terms of the nature of these properties. It helps to pinpoint the difficulties that L2 Chinese learners are experiencing in acquiring the BA construction properties. This knowledge helps teachers make their classroom instruction and practice more targeted and effective. One limitation of this study is that the proficiency levels of the subjects were only loosely controlled by excluding the beginning level L2 Chinese learners. The results would be more informative if more L2 Chinese were recruited and put into different groups. In future research, a more detailed examination could be conducted on the factors that may contribute to the L2 Chinese learners’ deviation in performance from the native speakers’ on the BA construction properties.

XU: ACQUISITION OF BA CONSTRUCTION PROPERTIES

REFERENCES

- Du, H. 2004. The acquisition of the Ba construction by adult second language learners (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (UMI No. 3132215)
- Jin, H. 1992. Pragmaticization and the L2 acquisition of Chinese BA constructions. *Journal of the Chinese Language Teachers Association*, 27(3), 33-52.
- Liu, F. 1997. An Aspectual Analysis of BA. *Journal of East Asian Linguistics*, 6, 51-99.
- Lu, S. 1980. *Xiandai Hanyu Babaici*. [Eight Hundred Words of Modern Chinese.]. Beijing: Shangwu Yinshuguan.
- Shi, Y. 2002. *The establishment of modern Chinese grammar: The formation of the resultative construction and its effects*. Amersterdam, Philadelphia: John Benjamins Publishing.
- Sorace, A. 2005. Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation: reconciling the biological and the social* (pp 55-80). Amsterdam: John Benjamins Publishing Co.
- Sybesma, R. 1992. *Causatives and accomplishments: the case of the Chinese ba*. Doctoral dissertation, Leiden University.
- Wang, L. 1945. *Zhongguo Yufa Lilun*. [Theories of Chinese grammar.]. Beijing: Zhonghua Press.
- Wang, M. (1987). Transitivity and the Ba-construction in Mandarin. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8702139)
- Wen, X. 2010. Acquisition of the displacement BA-construction by English-speaking learners of Chinese. *Journal of the Chinese Language Teachers Association*, 45(2), 73-100.
- Yuan, B. (2010) Domain-wide or variable-dependent vulnerability of the semantics-syntax interface in L2 acquisition? Evidence from wh-words used as existential polarity words in L2 Chinese grammars. *Second Language Research*, 26(2), 219-260.
- Zhang, S-L. 2002. Second language acquisition of the Ba-construction in contemporary mandarin Chinese (Master's thesis). Available from ProQuest Dissertations and Theses. (UMI No. 1414860)
- Li, A. Y-H. (2001). The Ba Construction. Retrieved from <http://www.usc.edu/dept/LAS/ealc/chinling/audreyL.htm>

L2 Acquisition of Chinese Locative Inversion

Jin Zhang
University of Arizona

Locative inversion is a non-canonical structure in Chinese and English. According to the Contrastive Analysis theory (Lado, 1957), a structure in L2 is easier to learn if it has the same meaning and distribution as an “equivalent” in L1. Previous studies suggest that English-speaking Chinese L2 learners would have no difficulty in acquiring the unmarked construction (Jin, 2008). The present study investigates how well CFL (Chinese as a foreign language) learners acquire locative inversion in Chinese and whether or not the positive L1 transfer occurs. The data comes from high-intermediate and advanced CFL learners. It consists of two parts - guided picture description and grammaticality judgments. Results from the two tasks show that the CFL learners did not have a good handle of locative inversion despite the existence of a similar construction in their L1. The CFL learners’ L1 did not help in their acquisition of Chinese locative inversion.

1. Introduction

Locative inversion is a non-canonical structure in Chinese. It is also found in a number of languages, including English. Much research has examined the construction from syntactic and morphological perspectives in Chinese (Pan 1996, Du 1999, Lin 2008, Zhang 2008), but very few studies have investigated the acquisition of the construction both in first language (L1) and in second language (L2).

According to the Contrastive Analysis theory (Lado, 1957), when a L1 structure and a L2 structure have the same form, meaning and distribution, positive transfer would occur in L2 acquisition. The L2 structures that already exist in learners’ L1 with the similar syntactic and semantic properties would not cause learning difficulty.

The similarities shared by Chinese and English locative inversion are frequently mentioned in literature (Levin & Rappoport Havov 1995, Du 1999) as they both follow the syntactic structure of “Locative + V + NP” and are very descriptive (Birner & Ward: 243-244, Lu: 462). Jin (2008) conducted a study on the English-speaking CFL learners’ acquisition of Chinese word order in relation to markedness theories. Locative inversion was listed among the four types of sentences surveyed. Jin (2008) argues Chinese locative inversion is not supposed to cause learning difficulty for English-speaking L2 learners, because the uses of locative inversion are very similar in both Chinese and English in terms of applied conditions and distribution in existential verbs, verbs of

motion, and verbs of positions. The results of the study indicate that the subjects surveyed had less difficulty in learning locative inversion compared with the other two more marked types of Chinese sentence structures. However, as it is not a focused study on locative inversion, the study cannot show us the real picture of L2 acquisition of the structure by comparing to the L2 acquisition of other marked structures in Chinese.

Looking at Chinese locative inversion by itself, the current study investigates how well CFL(Chinese as a foreign language) learners acquire locative inversion in Chinese and whether or not the positive L1 transfer occurs. The results will be analyzed to show whether learners' L1 necessarily has a positive influence on their acquisition of the Chinese locative inversion construction and if tasks cause any difference to learners' performance. The main research questions are:

- 1) How well do the American CFL learners recognize and use Chinese locative inversion?
- 2) Do different tasks influence learners' performance on the construction?
- 3) Does English-speaking CFL learners' L1 has a positive influence on learners' acquisition of Chinese locative inversion?

2. Basic properties of locative inversion

Levin and Rappoport Havov (1995: 218-219) give the three properties of locative inversion: first, as a non-canonical structure (PP V NP), it appears to be “the result of switching the positions of the NP and the PP in the canonical ‘NP V PP’ word order, particularly since the ‘inverted’ and ‘non-inverted’ sentences are near paraphrases of each other” (Levin & Rappoport Havov 1995: 218); second, the construction is so named for the presence of PP, a locative or directional PP, in preverbal position; third, the verbs in locative inversion constructions usually are intransitive. The first two properties are true to the languages, Chinese and English, considered in this study, but the third property is questionable in Chinese which does allow transitive verbs to occur in locative inversion. Another distinct characteristic of Chinese locative inversion, as Du (1999) states, is that an aspect marker is usually obligatory in the construction.

Locative inversion in Chinese is considered by researchers as an existential sentence and generally falls into three types according to semantic references of the verbs (Lu 2006: 460). The three types suggested by Lu (2006: 460) are as follows. Examples are provided in comparison with English and their canonical alternatives.

2.1 Existential state / posture

The verb describes the posture, the manner of motion, or the existential state of the object, either unanimated or animated. It is present in both English and Chinese.

(1) Chinese:

- a. (在) 桌 上 擺 -著/了 一 本 書。(locative inversion)
 (Zai) *zhuo shang bai-zhe/-le yi ben shu*
 (On) table top lay-DUR/PFV a CL book

“On the table lies a book.”

- b. 一 本 書 擺 在 桌 上。 (SVO)

Yi ben shu bai zai shuo shang.

A CL book lay on table top

“A book lies on the table.”

- c. 一 本 書 在 桌 上 擺 - 著。

Yi ben shu zai zhuo shang bai-zhe

A CL book on table top lay-DUR

“A book lies on the table.”

English:

- d. On the table lies a book. (locative inversion)

- e. A book lies on the table. (SVO)

The examples above show that in Chinese the preposition *zai* is optional in locative inversion (as in (1) a) while is required by its two SVO alternatives (as in (1) b, c). Example (1) b uses a postverbal locative phrase to name the place where the object described ends up as a result of the action of the verb. Hence, it does not need an aspect marker. The displacement of the locative phrase in example (1) c mainly shifts the focus of the sentence from the location to the object described. The effect of the change of the word order highly resembles that found in English locative inversion and its SVO counterpart.

2.2 Motions

Verbs in this type of locative inversion describe the motions of the object, unanimated or animated. The object is usually new to the scene, so the verb describes the appearance of the object at a particular location.

(2) Chinese:

- a. (從) 遠 處 來-了 一 輛 車。 (locative inversion)

(From) yuan chu lai-le yi liang che

(From) far place come-PFV a CL car

“From afar comes a car.”

- b. 一 輛 車 從 遠 處 來-了。 (SVO)

Yi liang che cong yuan chu lai-le

A CL car from far place come-PFV

“A car comes from afar.”

English:

- c. From afar comes a bus. (locative inversion)

- d. A bus comes from afar. (SVO)

In English, this case of inversion (as in (2) a) is also optional as it can be expressed in an SVO order (as in (2) b). However, in Chinese, the preposition “*cong* (從)” is optional in locative inversion (as in (2) a) but is required by an SVO order (as in (2) b).

2.3 Disappearance

Verbs in this type of locative inversion describe the disappearance or removal of the unanimated / animated object.

(3) Chinese:

- a. (在) 他 家 死-了 一 條 狗。 (locative inversion)
(Zai)Ta jia si-le yi tiao gou
 (At)his home die-PFV a CL dog
 “A dog has died in his home.”
- b. 一 條 狗 死 在 他 家。 (SVO)
Yi tiao gou die in his home.
 “A dog has died in his home.”
- c. 樓 裡 搬走-了 幾 戶 人。 (locative inversion)
Lou li banzou-le ji hu ren
 Building inside move-PFV several CL people
 “A few families have moved out of the building.”
- d. 幾 戶 人 從 樓 裡 搬走-了。 (SVO)
Ji hu ren cong lou li banzou-le
 Several CL people from building inside move-PFV
 “A few families have moved out of the building.”

English:

- N/A (locative inversion)
 e. A few families have moved out of the building. (SVO)

It is hard to find an equivalent of this type of locative inversion in English as it involves issues of agentive locative subject (Lin, 2008) and transitivity and intransitivity of verbs (Pan, 1996; Du, 1999; Lin, 2008) that are particularly complicated in Chinese.

To have a better control over the variables, the current study only examines the first type of locative inversion suggested by Lu (2006, pp. 460) with a verb plus an aspectual marker *zhe*. The four types of verbs covered in this study include: existence (有 *you*), posture (坐 *zuo*, 站 *zhan*, 躺 *tang*, and 趴 *pa*), existential state (放 *fang*, 擺 *bai*, and 種 *zhong*), and manner of motion (飛 *fei*, and 跑 *pao*). Locatives are provided without a preposition as preposition is usually omitted by native Chinese speakers in locative inversion.

3. Methods

3.1 Participants

Eighteen native English speakers learning Chinese as a foreign language and six native Chinese speakers participated in the experiment. They were all undergraduate / graduate students at the University of Arizona. The English-speaking participants were chosen by the criteria that they had learned Chinese locative inversion before. The eighteen English-speaking participants consisted of three groups, second-year Chinese learners (low-intermediate proficiency level), third-year Chinese learners (intermediate proficiency level), and advanced Chinese learners (actively studying Chinese for more than 4 years), with six participants for each group. The six native Chinese speakers formed the control group to be compared with the English-speaking CFL learner groups. They came from different regions of China and spoke both Mandarin Chinese and their own dialects.

3.2 Materials

A survey comprising two written tasks was used for this study to assess CFL learners' knowledge of Chinese locative inversion. One task was a picture description task, and the other was a grammaticality judgment task. The task directions were given in English. The content was appropriate for the proficiency level of the CFL participants. English translations were provided for some words and phrases in case CFL learners forgot or did not know.

PICTURE DESCRIPTION TASK. The picture description task asked students to describe where the assigned objects were on the picture. The picture presents a scene with animated and unanimated objects on it. In order to get a more controlled and quantifiable result, the task assigned six objects to the participants to describe and provided a matched number of verbs, and locatives that are essential to the locative inversion construction. The task tested whether the learners had the intuition to use the locative inversion construction in a context where the construction is preferred.

GRAMMATICALITY JUDGMENT TASK. The grammaticality judgment task consisted of 42 sentence items with 28 fillers and 14 sentences either correctly or incorrectly using locative inversion. The students were asked to judge the correctness of the sentences as perfect, okay, no intuition, awkward, and horrible. A numerical scale ranged from two to negative two was assigned to the five options to differentiate the degrees of the participants' acceptance of the sentences (2 for perfect, 1 for okay, 0 for no intuition, -1 for awkward, and -2 for horrible). Four types of verbs were used in the locative inversion sentences: existence, posture, existence of state, and motion. All these verbs were learned by the participants before and were commonly used in locative inversion. This task examined participants' recognition and understanding of Chinese locative inversion.

3.3 Procedures

An advertisement was sent to enroll learner participants in Chinese language classes. Those who were interested in participating contacted the investigator and scheduled a half-hour session for the week designated for the survey. The native Chinese participants were obtained through the investigator's personal network. The survey was conducted at the teaching assistant area of the Department of East Asian Studies at the University of Arizona. The investigator administered both tasks with each participant individually. The surveys were finished in two weeks.

3.4 Coding

The participants' responses were numerated for statistical analyses. As students were supposed to write six sentences in the picture description task, their responses were analyzed by the proportion of the number of locative inversion they used out of the six sentences. As for the grammaticality judgment task, to get continuous numbers, for each participant a value averaged over the locative inversion items by the 5-point numerical scale (2, 1, 0, -1, -2) represented the score the participant got. The locative inversion items were categorized into grammatical and ungrammatical sentences. That means the bigger the number value for the correct items, the better the participants understand the construction. On the contrary, the smaller the number value for the wrong items, the better the participants master the construction.

When comparing the two tasks, the participants' scores of the grammaticality judgment task were converted to 0-1 scale to keep them consistent with the scores of the picture description task. That is, the number values 2 and 1 for correct items and the number values -2 and -1 for wrong items were coded as 1, while the number values 1 and 2 for wrong items and the number values -1 and -2 for wrong items were coded as 0. A value averaged over the items was considered as the score the participant got.

4. Results and findings

To address the primary research question, "How well do the American CFL learners recognize and use Chinese locative inversion?" one-factor between-subjects analysis of variance (ANOVA) and a 4×2 mixed ANOVA were performed separately to investigate the CFL learners' overall performance on the picture description task and the grammaticality judgment task. A two-way ANOVA was conducted to explore whether the tasks had significant effect on the learners' performance and whether there were any significant interactions between the proficiency levels and tasks. When significant results were found, a series of one-way ANOVA were performed to further test the simple effects of tasks.

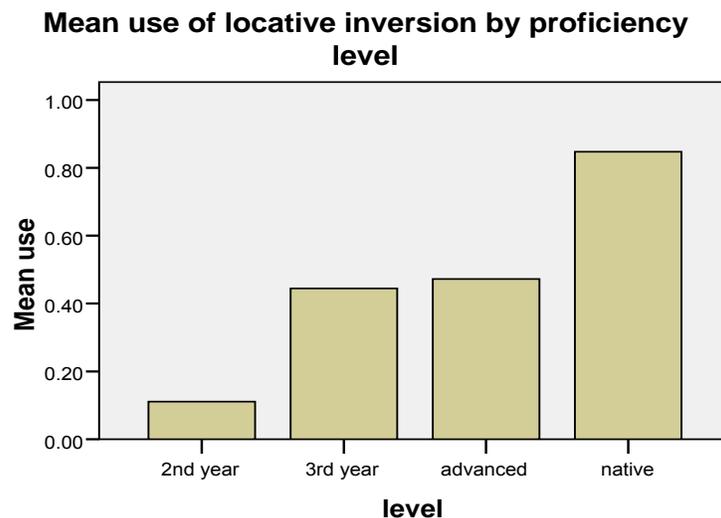
4.1 Overall results

The data on participants' responses to the two tasks was analyzed separately to see how well they use and recognize the locative inversion construction.

The data on locative inversion used in the picture description task was analyzed using a one-factor between-subjects ANOVA with proficiency level (second-year, third-year, advanced, native) as the factor. The subjects' performance was significantly different across different proficiency levels ($F(3, 20) = 18.98, p < .001$). Post hoc pairwise comparisons with Tukey correction were performed among the groups. Significant difference was found between each two groups (2nd year vs. 3rd year: $p < .02$; 2nd year vs. advanced: $p < .01$; 2nd year vs. native: $p < .001$; 3rd year vs. native: $p < .01$) except that between the third-year and the advanced groups ($p = .99$).

PICTURE DESCRIPTION. Figure 1 indicates the overall performance of the CFL learners on this task was poor, although distinct progress was found between the second-year learners and the other two learner groups. The higher the learners' proficiency level was, the more they used locative inversion. The second-year learners barely had the intuition to use locative inversion even with the contextual cues. The third-year learners used locative inversion considerably more than the second-year learners while the advanced learners only used the construction slightly more than the third-year Chinese learners. However, even the advanced learners used locative inversion only half as much as native Chinese speakers did.

Figure 1. Mean use of locative inversion by proficiency level



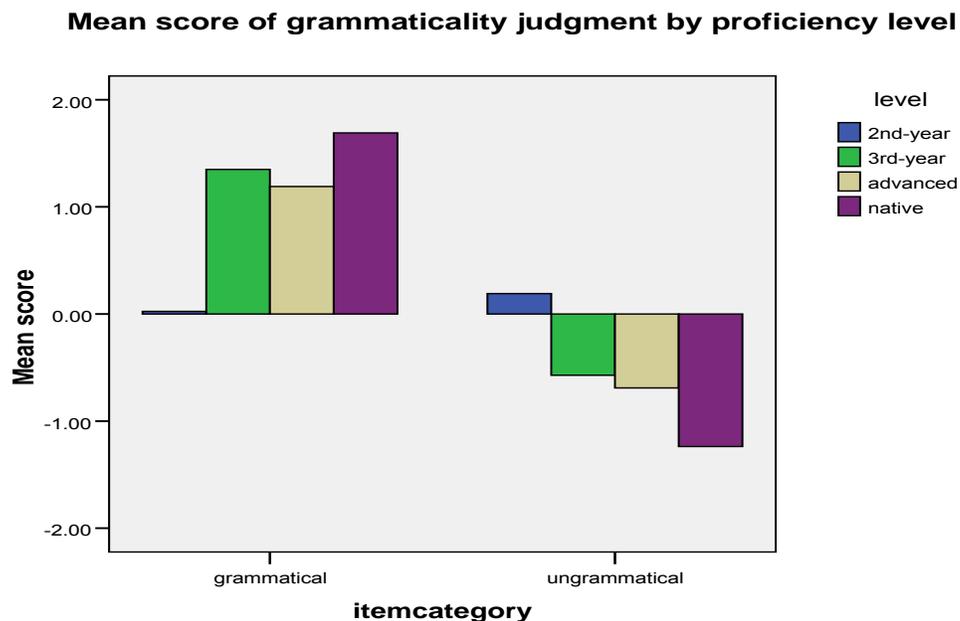
GRAMMATICALITY JUDGMENTS. The data on grammaticality judgments was analyzed using a two-factor mixed ANOVA, proficiency level (second-year, third-year, advanced, native) as a between-subjects factor and item category (grammatical, ungrammatical) as a within-subjects factor. Mauchly's test of sphericity showed there was no violation of the sphericity assumption. The main effect of item category and the

interaction of the proficiency level and item category were significant (itemcategory: $F(1, 20) = 47.22, p < .001$; item category \times proficiency level: $F(3, 20) = 7.40, p < .005$). However, there was no significant main effect of proficiency level ($F(1, 20) = .19, p = .91$). As the interaction of the proficiency level and item category were significant, simple effects of the two factors were tested separately.

Two separate one-way ANOVAs, one for grammatical items and the other for ungrammatical items, indicated that proficiency level had significant effect on grammatical items, $F(3, 20) = 3.36, p < .05$, but not on ungrammatical items, $F(3, 20) = 3.03, p = .53$. The results of post hoc comparisons with Tukey correction showed that the significant effect of proficiency level on grammatical items was caused by the huge gap between the second year CLF learner group and the native speaker group ($p < .03$) but not by any of the other two groups.

A one-factor within-subjects ANOVA, with item category (grammatical, ungrammatical) as the factor, was carried out to examine the simple effects of grammaticality on different proficiency levels. Significant difference between the two types of locative judgment items was found for the third-year CFL learner group ($F(1, 5) = 8.58, p < .05$), the advanced CFL learner group ($F(1,5) = 15.63, p < .02$), and the native speaker group ($F(1, 5) = 25.73, p < .001$), but not for the second-year CFL learner group ($F(1,5) = .19, p = .68$).

Figure 2. Mean score of locative inversion grammaticality judgment by proficiency level



The data reveals that the CFL learners were able to distinguish the grammatical locative inversion sentences from the ungrammatical ones except the second-year learners (see figure 2) who could hardly tell either the right use or the wrong use of the construction. Although proficiency level did not have a significant effect on learners' performance, there was a big leap from the second-year learners' performance to that of the advanced learners. Figure 2 presents that the higher the learners' proficiency level was, the better judgment they tended to have. The third-year learners and the advanced learners had a better judgment on grammatical items than on ungrammatical items. However, the advanced learners did not do significantly better than the third-year learners. In this study, the third-year learners even showed a slightly better performance on judging grammatical locative inversion sentences than the advanced learners (see figure 2), although the difference may be caused by one advanced learner who did particularly poorly on judging the grammatical items. The third-year and the advanced learners did not show a significant difference from the native Chinese speakers in the grammaticality judgment task. Each of the three CFL learner group basically exhibited a consistent performance on judging the two categories of sentences.

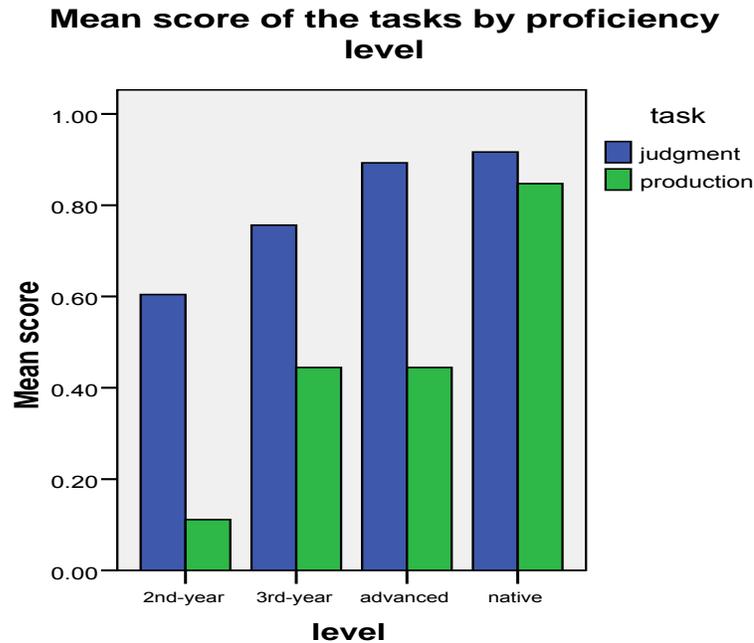
4.2 Task comparison

The data on learners' performance on the two tasks was analyzed using a two-factor mixed ANOVA, with proficiency level (second-year, third-year, advanced, native) and task (picture description, grammaticality judgments) as the two factors. Both the two main effects were significant (proficiency level: $F(1, 15) = 27.33$, $p < .001$; task: $F(1, 15) = 40.48$, $p < .001$), but there was no significant interaction of the two factors ($F(2, 15) = 1.60$, $p = .24$).

A one-factor within-subjects ANOVA was performed to test the simple effects of task at each proficiency level of the CFL learners. The results showed a significant effect of task on the second-year and advanced CFL learner groups (second-year: $F(1, 5) = 18.06$, $p < .01$; advanced: $F(1, 5) = 23.50$, $p < .006$). However, there was no significant difference for the third-year CFL learner group on the two tasks ($F(1, 5) = 5.83$, $p = .06$).

The data suggests that the Chinese learner participants generally did significantly better in the grammaticality judgment task than in the picture description task, while the native Chinese speakers indicated a consistent performance on the two different tasks (see figure 3). Advanced learners' performance on the judgment task was very close to that of the native Chinese speakers. This reveals that the CFL learners had much more trouble using the locative inversion construction than recognizing it, especially at the early stages of learning. The gap became smaller for learners who were more proficient in Chinese. However, while the more proficient learners always did better than the less proficient in the judgment task, the advanced learners did not show a noticeable progress in their ability to use the construction compared with the third-year learners. Compared with native Chinese speakers, the third-year and advanced learners hardly had the intuition to use locative inversion even in a context eliciting the use of the construction.

Figure 3. Mean score of the two tasks by level



5. Discussion

The results of the study essentially answered the research questions and offered some significant findings. First of all, despite the same construction in their L1, the CFL learners did not use locative inversion very well in Chinese. Learners' L1 did not help in CFL students' acquisition of Chinese locative inversion. On most occasions, learners would prefer to use SVO structure rather than the non-canonical locative inversion. This result is consistent with Erbaugh (1982) and Wen (2006)'s finding that Chinese child L1 learners and CFL learners both prefer SVO order in their acquisition of Chinese. In addition, learners' proficiency level played an important role in their using and understanding of the construction. Learners at a higher proficiency level did better on the two tasks than those at a lower level. However, no distinct differences have been found between the third-year learners and the advanced learners in their acquisition of the Chinese locative inversion construction. The findings indicate that the locative inversion construction is not as easy to acquire as it is assumed as an unmarked construction but requires certain amount of exposure to construction and the language as well. It is hard for the CFL learners to learn at the early stage of acquisition, as was demonstrated in their poor performance on judging both the grammatical and ungrammatical sentences of local inversion. Learners at a higher proficiency level showed a better judgment on the correctness of the use of locative inversion. Moreover, the CFL learners performed

significantly better on grammaticality judgment task than on the picture description task. Even the advanced learners still could not successfully use the construction in a context eliciting the use of it. This suggests that after reaching certain proficiency level, learners may have persistent problem which keeps them from doing better with the locative inversion construction.

Some factors may contribute to the learning difficulty. It is a rarely-used and an alternative construction for its canonical counterpart both in the learners' L1 and in Chinese. The use of the construction to a great extent is determined by discourse; although all languages share a similar way to produce "non-canonical" sentence structure, they differ in the realization of the construction. Also, Chinese locative inversion is not exclusive to intransitive verbs as in English. Transitive verbs can also be applied to the construction. The CFL learners may not be aware of the difference, which is indicated by the errors they made in the picture description task. The pragmatic, syntactic, and morphological differences could be the obstacles in the CFL learners' acquisition of the construction.

Besides the major findings mentioned above, the study also indicates that the verb types may influence the learners' acquisition of the construction or learners' acquisition of the construction may follow a certain sequence which is characterized by the verb types. The results of the picture description task show that learners had less difficulty in using locative inversion with the existential verb *you* which in most cases is bound to the construction and does not allow a canonical alternative. Therefore, learners seemed to be aware of the constraint set by *you* on locative inversion either in an explicit or an implicit way. Also, textbooks may influence learners' acquisition of the construction, too. The learners felt more comfortable with locative inversion wherever the verbs with which they were taught the construction were used, such as *bai* and *fang*. Although verb types were considered, it is not the focus of this study. A more detailed classification of verbs is needed in future study to investigate the acquisition pattern of locative inversion in terms of verb types.

6. Conclusion and implications

To conclude, my study provides evidence that Chinese locative inversion is not easy to acquire for English-speaking CFL learners. They showed a notable learning difficulty at the early stage of learning either in recognizing or using the construction. The CFL learners could hardly use locative inversion very well in Chinese even after reaching a high proficiency level. Although locative inversion has the same grammatical structure in English and Chinese, L1 did not help much in the English-speaking CFL learners' acquisition of Chinese locative inversion.

The results of the study also shed light on classroom instruction. First, instructors should draw students' attention to the similarities between Chinese and English locative inversion. They should also address and treat consistent errors made by students with locative inversion. Second, contextualized practice should be provided for students to get

them familiar with the discourse constraints set by locative inversion. Third, as mentioned previously, teachers should be aware of the acquisition pattern determined by verb types in their instruction.

REFERENCES

- Birner, Betty J. & Ward, Gregory. (1998). *Information Status and Noncanonical Word Order in English*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Erbaugh, Mary S. (1982). *Coming to Order: Natural Selection and the Origin of Syntax in the Mandarin Speaking Child*. Berkley: University of California.
- Du, JulietWai-hong. 1999. Locative inversion and temporal aspect in Mandarin Chinese. *Chicago Linguistic Society Regional Meeting*, 35 (1), 339-353.
- Jin, Lingxia. 2008. Markedness and second language acquisition of word order in Mandarin Chinese. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, 1, 297-308.
- Lado, Robert. (1957). *Linguistics Across Cultures*. Ann Arbor, MI: University of Michigan Press.
- Levin, Beth & Rappaport Havov, Malka. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- Lin, T.-H. Jonah. 2008. Locative subject in Mandarin Chinese. *Nanzan Linguistics*, 4, 69-88.
- Lu, Qinghe. 2006. *實用對外漢語教學語法 Shiyong duiwai hanyu jiaoxue yufa*. Beijing: Peking University Press.
- Pan, Haihua. 1996. Imperfective aspect ZHE, agent deletion, and locative inversion in Mandarin Chinese. *Natural Language and Linguistic Theory*, 14, 409-432.
- Wen, Xiaohong. 2006. Acquisition sequence of three constructions: an analysis of the interlanguage of learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association*, 41 (3), 89-113.
- Zhang, Lan. 2008. Locative inversion and aspect markers of *le* and *zhe* in Mandarin Chinese. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, 1, 893-902.