

Buckeye East Asian Linguistics Forum 4

Friday, 5 March 2021



<http://u.osu.edu/beal/beal-forum/>



THE OHIO STATE UNIVERSITY

Buckeye East Asian Linguistics Forum 4

Friday, 5 March 2021

The Ohio State University
Columbus, Ohio

A virtual event via Zoom

Program Book



THE OHIO STATE UNIVERSITY

Organizers and Sponsors

BEAL Forum 4 Organizing Committee

Faculty Co-Chairs

Mineharu Nakayama, Marjorie K.M. Chan, and Zhiguo Xie

Student Co-Chairs

Junyu Ruan (President, GACL) & Shunichi Maruyama (President, GREALL)

Committee Members

John Bundschuh, Jingyi Chen, Paul Cockrum, Hannah Dahlberg-Dodd, Skylor Gomes, Yuki Hattori, Saori Wakita, Seojin Yang, Jinwei Ye, Xuan Ye, Ying Zhang, Wei William Zhou, and Yuhong Zhu

Administrative Assistants

Dani Cooke (East Asian Studies Center)
David Liu (Dept. of E. Asian Languages & Literatures)

Commentators

Dana Scott Bourgerie (Brigham Young University)
Seth Goss (Emory University)
Hana Kang (University of Notre Dame)
Sue-mei Wu (Carnegie Mellon University)

Ohio State University Sponsors

Institute for Korean Studies
East Asian Studies Center

Graduate Association of Chinese Linguistics (GACL)
Graduate Students of East Asian Languages and Literatures (GREALL)

Department of Linguistics
Department of East Asian Languages and Literatures

This event is sponsored in part by a U.S. Department of Education Title VI grant for The Ohio State University East Asian Studies Center, programming fund for GACL from the Council on Student Affairs, and by the [James H-Y. Tai Buckeye East Asian Linguistics Fund](#).

Table of Contents

	Page
Organizers and Sponsors	ii
Table of Contents	iii
General Information	iv
Program	1
Plenary Sessions Abstracts	4
Professor Professor Kang-kwong Luke	5
Professor Jiwon Yun	6
Professor Haruko M. Cook	7
Session I-A Abstracts	8
Session I-B Abstracts	22
Session I-C Abstracts	33
Session II-A Abstracts	46
Session II-B Abstracts	56

General Information

1. Online Registration

<https://u.osu.edu/beal/beal-forum/2021-2/registration/>

- Register by 3 March 2021.

2. Buckeye East Asian Linguistics Forum 4 (BEAL Forum 4)

Date: Friday, 5 March 2021 (Eastern Standard Time)

Place: Virtual event via Zoom, The Ohio State University, Columbus, Ohio

Buckeye East Asian Linguistics Forum 4 (BEAL Forum 4) provides a platform primarily for graduate students to articulate and exchange ideas on their research findings with forum participants. It is a one-day event with keynote lectures and students' presentation sessions. Previously, students' presentation sessions were all posters. Due to the COVID-19 pandemic, these sessions will, instead, be oral, 15-minute presentations via Zoom. The BEAL Forum is an excellent opportunity to present one's research in a public arena and to receive comments individually. The Forum is intended to showcase regional research activities in East Asian linguistics.

As in the case of previous BEAL Forums, as part of this event, a Proceedings volume will be published through Ohio State University's Knowledge Bank, a service of The Ohio State University Libraries.

This year, we also have a pre-BEAL Forum 4 event announcement, namely, a guest lecture hosted by the Institute for Korean Studies at the East Asian Studies Center, to be held via Zoom on 4 March 2021 at 2:30-4:00 p.m. EST. The invited speaker is **Professor Ross King** (University of British Columbia), who will speak on "Cosmopolitan and Vernacular in the Sinographic Cosmopolis and Beyond: Traditional East Asian Literary Cultures in Global Perspective." (For more details, including registration, see: <https://easc.osu.edu/events/iks/rking>).

Commentators:

For this year's event, in addition to keynote speakers, we have invited several OSU alumni to serve as commentators, to provide feedback to the presenters. Invited to participate as commentators are:

- Professor Dana Bourgerie (Brigham Young University)
- Professor Seth Goss (Emory University)
- Professor Hana Kang (University of Notre Dame)
- Professor Sue-mei Wu (Carnegie Mellon University)

BEAL Forum 4 is free and open to the public, although registration is required in order to obtain the Zoom links.



THE OHIO STATE UNIVERSITY



**The 4th Buckeye East Asian Linguistics Forum
(BEAL Forum 4)**

Friday, 5 March 2021

Venue: Virtual event via 3 Zoom Meeting Rooms – Zoom A, B & C
Time Zone: Eastern Standard Time (EST)
Registration: Free and open to the public, but online registration required

8:45-9:00 am

Welcoming Remarks (Zoom A)

*Professor Mitchell B. Lerner, Director, East Asian Studies Center
Professor Mark Bender, Chair, Dept. of E. Asian Languages and Literatures*

9:00-10:20 am

Plenary Session 1: Keynote Speaker (Zoom A)

Chair: Professor Marjorie K.M. Chan

Professor Kang-kwong Luke
Nanyang Technological University, Singapore

“Languages and Cultures in Action: Snippets of Interactions from Singapore”

Session I. 10:25-11:45 a.m.			
	Session I-A (Zoom A)	Session I-B (Zoom B)	Session I-C (Zoom C)
	Chair: Mineharu Nakayama	Chair: Björn Köhnlein	Chair: Etsuyo Yuasa
10:25 am	Jonathan Him Nok Lee (Chinese U. of Hong Kong), Regine Yee King Lai (Chinese U. of Hong Kong), Stephen Matthews (U. of Hong Kong), and Virginia Yip (Chinese U. of Hong Kong) Bilingual Intonation in Cantonese-English Bilingual Children's Sentence-final Particles	Ian Joo and Yu-Yin Hsu (Hong Kong Polytechnic U.) A Preliminary Survey on Linguistic Areas in East Asia Based on Phonological Features	Ka-Fai Yip (Yale U.), Tommy Tsz-Ming Lee (U. of Southern California), and Sheila Shu-Laam Chan (Chinese U. of Hong Kong) Deriving Separable Verbs in Cantonese
10:45 am	Jun Lyu (U. of Southern California) "Tone Sandhi in Mono/polysyllabic Single Words in Shanghai Chinese"	Wei William Zhou (Ohio State U.) On the Status of a Tone Merger in Dalian Mandarin	Ka-Fai Yip (Yale U.) Two Types of Temporal Adverbial Clauses in Cantonese
11:05 am	Ke Wang (Ohio State U.) Dialect and Watershed Distributions in Shaanxi Province and the Jiang-Huai Area in China	Alexandra Konovalova and Alena Tsvetkova (National Research University Higher School of Economics) Comparative Analysis of Grapheme-to-Phoneme Models for the Russian-Chinese Parallel Corpus	Jinwei Ye (Ohio State U.) Character Writing, Lexicon and Syntax in 19th Century Cantonese: Some Observations from Bridgman (1841)
11:25 am	Paul D. Cockrum (Ohio State U.) Reflecting Identity through Song Lyrics: Usage of Variant Graphs in Popular Taiwanese Southern Min Songs	Yuhong Zhu (Ohio State U.) Phrasal Tonology in Suzhou Chinese: Some Preliminaries	Xuan Ye (Ohio State U.) and Yao Chen (Fujian Normal U.) Doubling of Perfective Aspect Markers in Hui Chinese
11:45 am - 1:00 pm	~~ LUNCH BREAK ~~		

1:00-2:20 pm	Plenary Session 2: Keynote Speaker (Zoom A) Chair: Professor Zhiguo Xie
	Professor Jiwon Yun Stony Brook University “ <i>Wh</i> -words in East Asian Languages: Intonation and Meaning”

Session II. 2:25-3:25 p.m.		
	Session II-A (Zoom A)	Session II-B (Zoom B)
	Chair: Hannah Dahlberg-Dodd	Chair: Danielle O. Pyun
2:25 pm	Yuning Cao (Stanford U.) Gender in Japanese Youth Language	Eunhye Kim Hess (Oklahoma State U.) A Usage-based Construction Approach to Korean DO-Causatives
2:45 pm	John Bundschuh (Ohio State U.) The Narrative Functions of Perfective Auxiliaries in Early Heian <i>Kundokubun</i> Texts	Seojin Yang (Ohio State U.) Korean Writing System: Adaptation and Creation
3:05 pm	Yuki Hattori (Ohio State U.) Loanword Adaptation in Japanese Kansai Dialect	

3:30-4:50 pm	Plenary Session 3: Keynote Speaker (Zoom A) Chair: Professor Mineharu Nakayama
	Professor Haruko M. Cook University of Hawaii “Referential and Non-referential (Im)politeness: The Use of Honorifics in Face-attacking Acts in a Japanese Company’s Orientation Meeting”
4:50-5:00 pm	Closing Remarks (Zoom A) BEAL Forum 4 Organizing Committee

Plenary Session
Abstracts

Languages and cultures in action: Snippets of interactions from Singapore

K.K. Luke

Nanyang Technological University, Singapore

The city-state of Singapore is known for its linguistic and cultural diversity, with a community made up of people from a variety of cultural traditions and an education system that promotes plurilingualism in English, Chinese, Malay, and Tamil. Within each of these languages, there are further variations and complications. Speakers of ‘Chinese’, for example, find themselves using, in addition to Mandarin, a mix of ‘Chinese dialects’, with Hokkien (Min) being the most popular, but there are equally large numbers speaking Toechew (or Chaozhou), Hakka (or Kejia) and Cantonese (Yue).

In this paper we move from a bird’s eye view of the community to an engagement with languages and cultures at the ground level by zooming in on talk-in-interaction in Singapore as people go about their everyday businesses. Using snippets of social interactions in the form of video recordings and an Ethnomethodological and Conversation Analytic approach, I will show how a host of interactional goals are achieved via the skilled use of a pool of interlingual and intercultural resources that have a distinctly ‘Singaporean flavour’. By approaching these interactions as ‘culture in action’ (Hester & Eglin 1997), I will unpack the localisms and distinctive forms of expression in an attempt to document and celebrate a rich and colorful kaleidoscope of creative practices that we can identify as ‘speaking, the Singapore way.’

BIO: K.K. Luke is President’s Chair Professor of Linguistics and Chair of the School of Social Sciences at Nanyang Technological University, Singapore. K.K.’s research is on talk and social interaction using an Ethnomethodological and Conversation Analytic approach. This research explores the ways in which joint actions are achieved through talk (and ‘body language’), and is driven by the fundamental question of what makes communication possible. Among K.K.’s publications are *Utterance Particles in Cantonese Conversation*, *Telephone Calls: Unity and Diversity in the Structure of Telephone Conversations across Languages and Cultures*.

Kang-kwong Luke: KKLUKE@ntu.edu.sg

Wh-words in East Asian languages: Intonation and meaning

Jiwon Yun

Stony Brook University

Wh-words in East Asian languages are ambiguous between interrogative and indefinite readings (e.g. *shei* in Chinese, *dare* in Japanese, *nwukwu* in Korean can all appear in phrases that indicate either ‘who’ or ‘someone’). This talk concerns how the two meanings of *wh*-words can be distinguished by intonation. While several intonational factors such as phonetic prominence, *phonological phrasing*, and sentence-final intonation have been mentioned in the literature to affect the interpretation, I argue that the most decisive factor is phonological phrasing that affect the global intonation pattern, presenting evidence from production/perception experiments (Jun & Oh 1996; Maekawa 1991; Yang et al. 2020; Yun 2019). In particular, the experimental results disconfirm the previous impressionistic observations that *wh*-interrogatives are intonationally distinguished from their indefinite counterparts by their prominence; the prosodic prominence of the *wh*-word *per se* does not induce an interrogative reading unless the change of the global intonation pattern is accompanied.

BIO: Dr. Jiwon Yun is Associate Professor of the Department of Linguistics at Stony Brook University. Her areas of research include semantics, prosody, computational linguistics, and cognitive science. Her research has paid attention to the East Asian languages including Chinese, Japanese, and Korean. She received a Bachelor’s degree in Computer Science and Engineering at Seoul National University, Korea, and a Ph.D. in Linguistics at Cornell University.

Jiwon Yun: jiwon.yun@stonybrook.edu

**Referential and non-referential (im)politeness:
The use of honorifics in face-attacking acts in a Japanese company's orientation meeting**

Haruko M. Cook

University of Hawai'i at Manoa

Speakers of honorific languages can linguistically express politeness/impoliteness referentially through the propositional content and/or nonreferentially through the presence or absence of honorifics. What is problematic is when co-occurring referential and nonreferential expressions are in conflict as in the co-occurrence of honorifics and referentially face-attacking expressions. Why do speakers use honorifics when they express a face-attacking referential message?

Brown and Levinson's politeness theory (1987) explains that the use of honorifics is a negative politeness strategy that mitigates an FTA (face-threatening act). Except for Brown and Levinson (1987), this problem has not been discussed much in the literature of linguistic (im)politeness in the West due to the lack of honorifics in English. However, it is an important issue when studying linguistic (im)politeness in languages with elaborate systems of honorifics such as Japanese and Korean, because speakers of these languages constantly need to make choices between honorific and non-honorific forms.

By examining a Japanese company's new employee orientation discourse, this talk explores why honorifics are used when the propositional content of an utterance is face-attacking. A new employee orientation in a Japanese company is a context in which impoliteness is ideologically legitimized and often deployed. At the same time, the goal of the orientation is to train new employees to behave in an extremely polite manner. By qualitatively analyzing the speech of the trainer of an orientation, this paper concludes that the trainer's use of honorifics while attacking the positive face of the new employees is a way of resolving the conflicting demands of a Japanese company.

This paper contributes to (im)politeness research in that it points to the importance of distinguishing referential and non-referential (im)politeness, and it shows why the co-occurrence of honorifics with face-attacks involves more than mitigating an FTA, as claimed by Brown and Levinson (1987).

BIO: Haruko Minegishi Cook is a professor in the Department of East Asian Languages and Literatures at the University of Hawai'i at Manoa. Her research interests include Japanese linguistics, language socialization, discourse analysis, and pragmatics. Her research explores how participants of social interactions utilize linguistic forms as a resource to construct a social world. She published a book on style shift between JFL learners and their host families and a co-edited volume on Japanese workplace discourse. She also published a number of articles in major journals and edited volumes. A co-edited volume on language socialization and *Journal of East Asian Pragmatics'* Special Issue on impoliteness are forthcoming.

Haruko M. Cook: hkcook@hawaii.edu

Session I-A
Abstracts

Bilingual intonation in Cantonese-English bilingual children's sentence-final particles

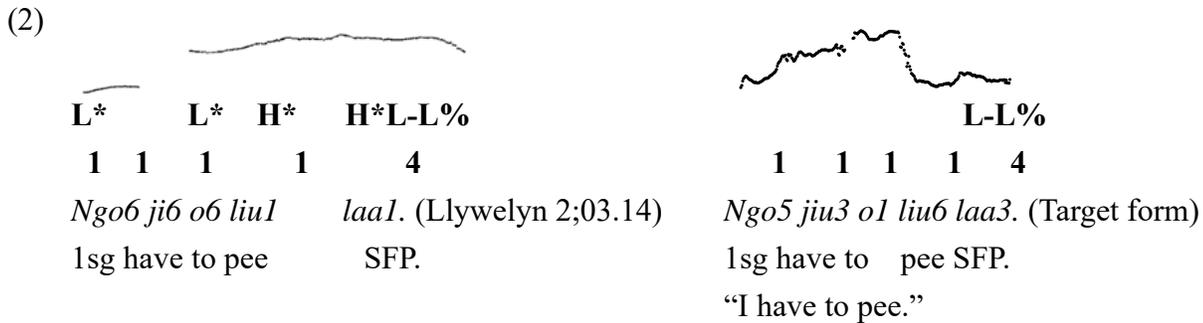
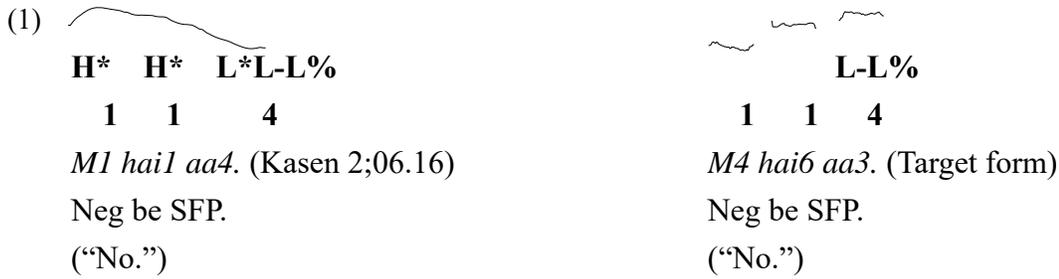
Jonathan Him Nok Lee¹, Regine Yee King Lai¹, Stephen Matthews², & Virginia Yip¹

¹*The Chinese University of Hong Kong, and* ²*The University of Hong Kong*

This corpus-based study investigates the intonation patterns of Cantonese sentence-final particles (SFPs) produced by Cantonese-English bilingual children. In recent bilingual first language acquisition research, the interaction between two or more languages in children's development (e.g., cross-linguistic influence) has been at the center of discussion. However, few studies in bilingual phonological acquisition have focused on prosodic aspects of bilingual children's production. This study will discuss the interaction of a tonal language (Cantonese) and an intonational language (English), and how the use of pitch for intonation interacts with the pitch patterns of lexical tones in their Cantonese and code-mixed utterances. In addition, many of the semantic/pragmatic functions encoded by intonation patterns in English are expressed with sentence-final particles (SFPs) rather than prosodic cues in Cantonese (Wakefield, 2010). This difference between English and Cantonese can potentially give rise to a transfer effect in the speech of Cantonese-English bilingual children.

The current study investigates the production of 8 Cantonese-English bilingual children (5 Cantonese-dominant; 1 English-dominant; 2 balanced) from ages 2;0 to 3;0 in the longitudinal Hong Kong Bilingual Child Language Corpus (N=8, Yip & Matthews, 2007). The corpus data were analyzed by sampling children's speech at 3-month intervals at ages 2;0, 2;3, 2;6, 2;9 and 3;0.

We have observed two bilingual intonation patterns: (i) "high pitch followed by a fall" and (ii) "low pitch followed by a rise", regardless of language dominance of the bilingual children and sentence types. For instance, the intonation pattern (i) in example (1) can be attributed to the superimposition of an English intonation with high pitch followed by a low boundary tone at the utterance final position (predominantly SFPs). The intonation pattern (ii) in example (2) can be attributed to the superimposition of an English intonation with low pitch followed by a high boundary tone at the utterance final position (predominantly SFPs). The superimposition of the two bilingual intonation patterns seems to dominate in the bilingual prosody produced by the English-dominant child Charlotte (ranging from 1.6% to 28.0% of all utterances in the sample recordings from ages 2;0 to 3;0). In contrast, Cantonese-dominant children produced bilingual prosody in much lower frequencies (< 10% of all utterances in the sample recordings from ages 2;0 to 3;0). They usually produced target-like lexical tones and intonation in their Cantonese and code-mixed utterances.



We argue that bilingual prosody is the result of cross-linguistic influence from English intonation. We hypothesize the following relationship between language dominance and the production of bilingual prosody: if the Cantonese-English bilingual child is more English-dominant, we predict greater cross-linguistic influence from English, and thus the more frequent production of bilingual prosody and vice versa. The results show that the two variables are found to be positively correlated, $r(25) = .431$, $p = .025$. The more dominant the Cantonese-English bilingual child is in English, the more bilingual prosody is produced in Cantonese and code-mixed utterances.

Several possible natures and sources of the bilingual intonation patterns will be discussed. First, the two bilingual intonation patterns can be viewed as emergent prosodic templates which are applied without conscious “strategies” to sentences that exceed the children’s cognitive capacities and physiological capabilities. Such templates can be taken to mediate between output- and input-based learning. The templates can be seen as reflecting distributional learning over the database constituted by repeatedly used child output forms (Vihman & Wauquier, 2018), which is supported by the prominent application of the bilingual intonation pattern (i) to some frequently used phrases (e.g., *m4 hai6 aa3* “no” (1)) and indirectly evidenced by the predominant use of high-level tone (tone 1) in L2 native English learners of Cantonese. It is also observed that the bilingual children were often emotional and would raise the pitch of speech, resembling high-level tone in Cantonese, when producing these frequently used phrases. Such a rise in pitch may alter the tones of syllables, and may in turn reinforce the distributional learning of the intonation pattern (i). At the same time, the bilingual intonation patterns reflect the typological differences of the ambient languages as perceived and filtered by the children. For instance, the low boundary tone in

intonation pattern (i) and the high boundary tone in intonation pattern (ii) are strictly licensed at the utterance-final position (predominantly SFPs). The fact that SFPs are used as the domain for intonation realization is consistent with the location of high boundary tone in Cantonese questions (Xu & Mok, 2011). Additionally, the two bilingual intonation patterns are a “mirror image” of each other. This echoes with the “Octave Shift Hypothesis” proposed by Hirst (2013) who argues that declarative rises in English are essentially the same as declarative falls, despite the seemingly opposite patterns, owing to physiological constraints.

Second, certain inputs of the ambient languages resemble the bilingual intonation patterns. Intonation pattern (i) resembles a transformation of English declarative intonation. Possible sources of intonation pattern (ii) include the English intonation of the Filipino caretakers (3), infant-directed speech (4), and listing intonation in English.

(3) *HOU: I don't know. You see in the box. (Alicia 2;09.15)

(4) *FAT: Okay, put those away. In a safe place. (Alicia 3;00.10)

The above phenomena demonstrate how cross-linguistic influence is manifested in the prosody of Cantonese-English bilingual children when they produce SFPs.

References

- Hirst, D. (2013). UK declarative rises and the frequency code. In S.Hancil & D.Hirst (Eds.), *Prosody and Iconicity* (pp. 149–160). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Vihman, M., & Wauquier, S. (2018). Templates in child language. In M.Hickmann, E.Veneziano, & H.Jisa (Eds.), *Sources of variation in first language acquisition: Languages, contexts, and learners* (pp. 27–44). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Wakefield, J. C. (2010). *The English Equivalents of Cantonese Sentence-final Particles: A Contrastive Analysis*. Ph.D dissertation, The Hong Kong Polytechnic University.
- Xu, B. R., & Mok, P. K. P. (2011). Final rising and global raising in Cantonese intonation. *Proceedings of 17th International Congress of Phonetic Sciences*, (January), 2173–2176. Hong Kong.
- Yip, V., & Matthews, S. (2007). *The Bilingual Child: Early Development and Language Contact*. Cambridge: Cambridge University Press.

Jonathan Him Nok Lee: jonjonleehn@link.cuhk.edu.hk
Regine Yee King Lai: ryklai@cuhk.edu.hk
Stephen Matthews: matthews@hku.hk
Virginia Yip: vcymatthews@cuhk.edu.hk

Tone sandhi in mono/polysyllabic single words in Shanghai Chinese

Jun Lyu

University of Southern California

Introduction

This study examines tone sandhi for mono/polysyllabic single words in Shanghai Chinese. Five contour citation tones exist in this dialect ([6], see **Table 1**), although underlyingly there has been proposed to be only /LH/ and /HL/ (e.g., [1], [3], [5]). When several syllables combine to form a single word (compounds not considered here), tone sandhi occurs.

Table 1. Citation tones in Shanghai Chinese.

	T1 (<i>Yinping</i>)	T2 (<i>Yinqu</i>)	T3 (<i>Yangqu</i>)	T4 (<i>Yinru</i>)	T5 (<i>Yangru</i>)
Xu et al. (1988)	53	34	23	55	12
Duanmu (1997)	HL	MH	LH	MH	LH

Tone sandhi in Shanghai Chinese can be characterized by tonal redistribution or a “spread-cum-delink” process. When two syllables form a word, the tone of the second syllable is delinked from its tone bearing unit (TBU). Then the contour tone of the first syllable spreads to the entire word (see (1)). The **scope of tonal redistribution** is called a “redistribution domain” in this work. In a polysyllabic word, if only one TBU gets to keep and redistribute its citation tone, we say there is one redistribution domain; but if two TBUs redistribute their tones, there are two domains.

(1) Tonal redistribution in Shanghai Chinese: example with disyllabic words

HL LH **HL** LH **H L** *Step 1: The 2nd syllable delinks from its citation tone.*
 pa li (‘Paris’) → pa li → pa li *Step 2: Each “segment” tone occupies TBU, left to right.*

It has been proposed that a tonal redistribution domain is also a foot domain ([2], [3], [7]) as only unstressed syllables lose their citation tones (like vowel reduction for unstressed syllables in English). However, there is a lack of consensus on the number of redistribution domains for quadrisyllabic and quintesyllabic words. Specifically, Xu et al. (1988) suggest that there is *only one redistribution domain* for quadrisyllabic words (although variation exists for T5-initial words), but Duanmu (1997) suggests that there are *two redistribution domains* not only for quadrisyllabic (e.g., (2a)) but also for quintesyllabic words (e.g., (2b)).

(2) a. **HL** LH **HL** LH b. **HL** HL **LH** LH LH
 (**H L**) (**H L**) (**H L**) (**L H** L) (last TBU receiving default L tone)
 ja lu sa l. (‘Jarusalem’) piŋ ɛi faʔ ni ja (‘Pennsylvania’)

The goals of this study are (i) to empirically examine the tonal patterns of polysyllabic words given the lack of consensus and (ii) to see whether the tonal patterns of polysyllabic words show variation. The results will allow us to assess the empirical foundation of the foot structure analysis of Duanmu (1997) and shed light on the constraint interaction in the OT framework ([4]).

Methods

Pronunciations of a male native speaker based on the **word list below** were recorded. Words with the same syllabic length are near minimal pairs. For words with more than 3 syllables, most are nonce words due to the lack of sufficient real word minimal pairs (but for quadrisyllabic and quintsyllabic words, extra real words were also collected, see more below). The pitch values of the words were extracted with Praat. A pitch value was sampled every 10ms for each audio file.

Monosyllabic		Disyllabic		Trisyllabic		Quadrisyllabic		Quintsyllabic	
Initial Tone	Word in IPA Gloss	Initial Tone	Word in IPA Gloss	Initial Tone	Word in IPA Gloss	Initial Tone	Word in IPA Gloss	Initial Tone	Word in IPA Gloss
T1 (52)	ε "grief"	T1 (52)	ε.tein "sorrow"	T1 (52)	ε.li.ja nonce-word	T1 (52)	ε.sa.ni.ja nonce-word	T1 (52)	ε.se.ɲu.pi.ja "Ethiopia"
T2 (34)	ε "love"	T2 (34)	ε.tein "love"	T2 (34)	ε.li.ja nonce-word	T2 (34)	ε.sa.ni.ja "Estonia"	T2 (34)	ε.se.ɲu.pi.ja nonce-word
T3 (23)	ε "salty"	T3 (23)	ε.te "salted egg"	T3 (23)	ε.li.ja nonce-word	T3 (23)	ε.sa.ni.ja nonce-word	T3 (23)	ε.se.ɲu.pi.ja nonce-word
T4 (55)	aʔ "duck"	T4 (55)	aʔ.tsʒ "duck"	T4 (55)	aʔ.li.ja nonce-word	T4 (55)	aʔ.sa.ni.ja nonce-word	T4 (55)	ε.se.ɲu.pi.ja nonce-word
T5 (12)	aʔ "box"	T5 (12)	aʔ.tsʒ "box"	T5 (12)	aʔ.li.ja nonce-word	T5 (12)	aʔ.sa.ni.ja nonce-word	T5 (12)	ε.se.ɲu.pi.ja nonce-word

Results

Fig. 1 shows the (preferred) tonal patterns for mono/polysyllabic words in Shanghai Chinese. Beside each word is its tonal contour (e.g., “HL”). The pitch contours of all words are summarized in **Table 2**. There are three main findings.

Fig. 1 Preferred tonal patterns in Shanghai dialect

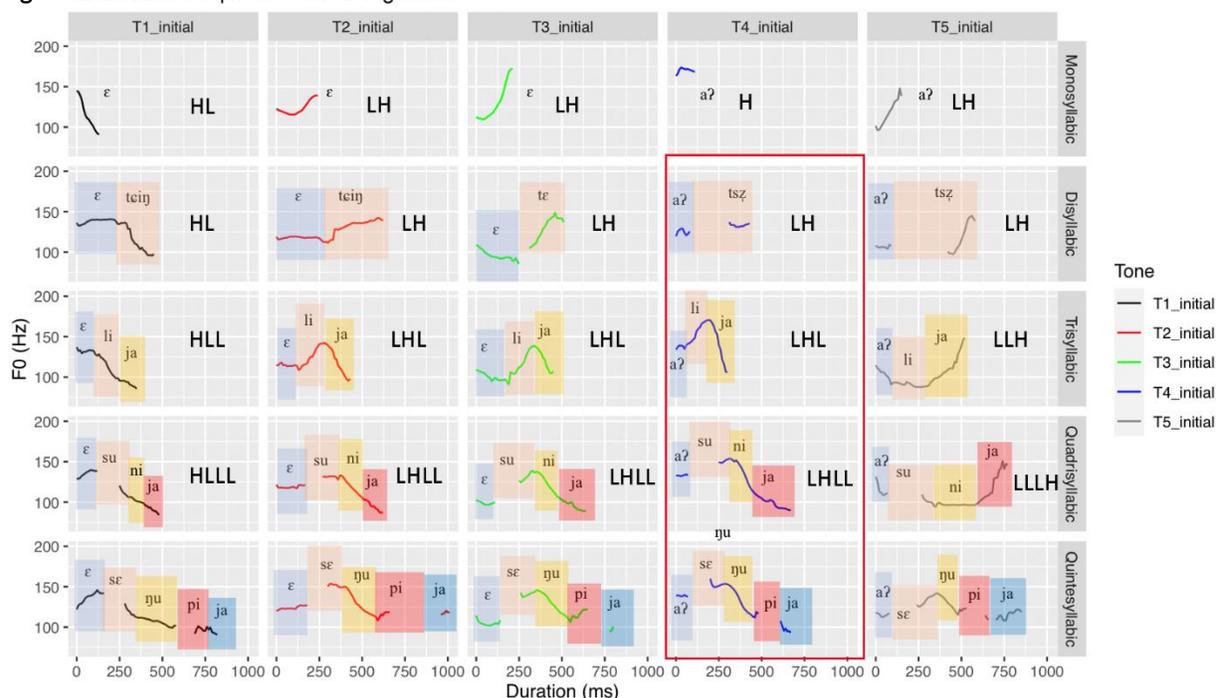


Table 2. Pitch contours for mono/polysyllabic words with different initial tones (preferred pattern).

Initial Tone	Monosyllabic	Disyllabic	Trisyllabic	Quadrisyllabic	Quintsyllabic
T1	HL	H-L	H-L-L	H-L-L-L	H-L-L-H-L
T2	LH	L-H	L-H-L	L-H-L-L	L-H-L-H-L
T3	LH	L-H	L-H-L	L-H-L-L	L-H-L-H-L
T4	H	L-H	L-H-L	L-H-L-L	L-H-L-H-L
T5	LH	L-H	L-L-H	L-L-L-H	L-H-L-L-L

First, for monosyllabic words, T4 is actually a high-level tone, not a LH or MH tone. However, in polysyllabic words, the underlying contour of T4 emerges (see column 4 in Fig. 1) as the first two syllables form a LH contour. This suggests that *T4 phonologically behaves like a /LH/ tone*, although its surface realization is a H tone on monosyllabic words (presumably due to its short duration with a glottal stop /ʔ/).

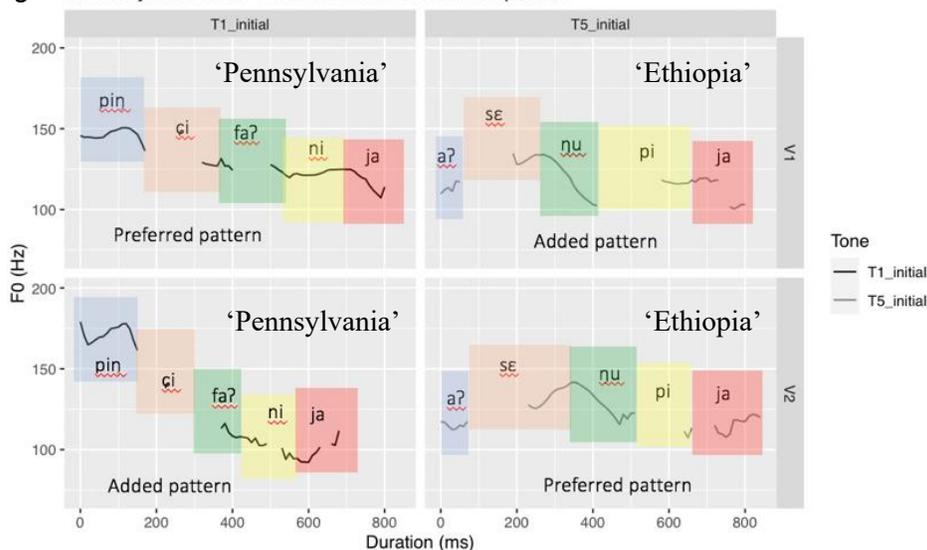
Second, the native speaker prefers to treat quadrisyllabic words as *one tonal redistribution domain* rather than two (see **Table 3**): for T1-initial words, only the first syllable has the H tone, with the rest getting default L tones; for T2/T3/T4-initial words, the first tone spreads its H tone to the second syllable with the rest syllables getting default L tones; for T5-initial words, the tone of the first syllable (i.e. LH) spans across the whole word: the initial portion of the tone drops to the lowest point on the first syllable and stays low until the last syllable for a rise. Despite the different tonal redistribution patterns linked to different initial tones – well documented in the literature (e.g., [3], [6]) – *all quadrisyllabic words were pronounced as one redistribution domain*, thus supporting Xu et al. (1988)’s description. This raises a potential challenge to Duanmu’s foot structure analysis, as the attested pattern (i.e. one redistribution domain) is harmonically bounded (i.e. can never win) in his OT analysis and thus should not occur, contrary to observation.

Table 3. Tonal domains for mono/polysyllabic words with different initial tones (preferred pattern): “S” refers to a stressed syllable (keeps its citation tone for redistribution) and “s” refers to an unstressed syllable (loses its citation tone). Parentheses demarcate the redistribution domain.

Initial Tone	Monosyllabic	Disyllabic	Trisyllabic	Quadrisyllabic	Quintesyllabic
T1	S	(Ss)	(Sss)	(Ssss)	(Ss)(Sss)
T2	S	(Ss)	(Sss)	(Ssss)	(Ss)(Sss)
T3	S	(Ss)	(Sss)	(Ssss)	(Ss)(Sss)
T4	S	(Ss)	(Sss)	(Ssss)	(Ss)(Sss)
T5	S	(Ss)	(Sss)	(Ssss)	(Sssss)

Third, variations within T5-initial quintesyllabic words were found, see **Fig. 2** (‘Pennsylvania’ is one of the “extra” real words not shown in list Fig. 1). The V1 variety represents the two-domain pattern while V2 variety represents one-domain tonal pattern. The “added” pattern was less preferred but still accepted and added by the native informant himself. Suffice it to say that unlike quadrisyllabic words, quintesyllabic words tend to be dissected into two redistribution domains (intuitively, redistribution domains cannot be stretched too long). Overall, the results in this study call for a careful re-examination of the previous metrical analysis of tones in Shanghai Chinese.

Fig. 2 Quintesyllabic tonal variation within the same speaker



References

- [1] Chen, M. Y. 2000. *Tone Sandhi*. Cambridge University Press, Cambridge.
- [2] Duanmu, S. 1995. Metrical and tonal phonology of compounds in two Chinese dialects. *Language* 71(2): 225-259.
- [3] Duanmu, S. 1997. Recursive constraint evaluation in optimality theory: Evidence from cyclic compounds in Shanghai. *Natural Language & Linguistic Theory* 15(3): 465-507.
- [4] Prince, A. and Smolensky, P. 1993. Optimality theory: Constraint interaction in generative grammar. Manuscript.
- [5] Selkirk, Elisabeth & Shen, T. 1990. Prosodic domains in Shanghai Chinese. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection*, University of Chicago Press, Chicago, pp. 313-337.
- [6] Xu, B., Tang, Z., You, R., Qian, N., Shi, R. & Y. Shen. 1988. *Urban Shanghai Dialect*. Shanghai Education Press, Shanghai.
- [7] Zhang, J. 2014. 汉语方言变调的能产性及其理论分析 (Productivity of tone sandhi in Chinese dialects and its analysis). *Contemporary Linguistics* 16(3): 273-287.

Jun Lyu: junlyu@usc.edu

Dialect and watershed distributions in Shaanxi Province and the Jiang-Huai area in China

Ke Wang

The Ohio State University

Inspired by several dialect geography studies (Coblin, 2002; Chamberlain, 2005; Lu, 2019), the present study mainly examined whether geography plays a deterministic role in the formation of the current dialectal conditions of certain localities in China. Through a map-to-map comparison and a literature review, similar overlapping patterns were found between the watershed and dialect boundaries in two areas with geographical and dialectal diversity: the Jiang-Huai area and Shaanxi province. Three criteria were used to determine whether the linkage between dialect and geography holds at the regional level: (1) dialectal diversity, (2) geographical diversity, and (3) the availability of geographic and linguistic data. Preliminary findings indicate a correlation between watershed distributions and dialectal boundary formation in these localities.

The literal meaning of the term dialect, 方言 (fāng yán), is regional speech, suggesting an internal, natural, and close bond between dialect and geography; this type of symbiotic relationship has been noticed and discussed by various scholars. For example, Chamberlain (2005) studied the watershed distribution in the Tibetan region and concluded that the watersheds are closely correlated with known linguistic groupings, suggesting the possible delineation of local language variations despite the natural, formidable physical obstacles of the Tibetan region. Simmons et al. (2006) investigated the boundary between the Wu and Jiang-Huai Mandarin dialects and outlined the common southern Jiang-Huai phonological system. Zhang and Zhuan (2008) demonstrated that natural and historical geography both play important roles in the distribution of the Hakka, Yue, and Min dialects in Guangdong province.

The Jiang-Huai area comprises the plain between the Yangtze River, the longest river in Asia, and the Huai River, a major river in east-central China with numerous tributary streams. The Jiang-Huai area is 243,300 square kilometers, has a population of 164.4 million, and is known as a dialect aggregation for its abundance of early Chinese language varieties (Coblin, 2002) and local dialects. Shaanxi province is a landlocked province in northern-central China. Bordering eight other provinces, it is about 195,800 square kilometers, has a population of more than 30 million, and is one of the renowned original sites of Chinese civilization. Part of the famous line from the Qin Ling Mountains, 秦岭 (qín líng), to the Huai River, 淮河 (huái hé), is in Shaanxi province; this line is the most important geographical marker dividing China into north and south. Many dialects are spoken in Shaanxi province.

Through direct comparisons of the *Language Atlas of China* and Chinese river basin distribution maps via contour matching, boundary comparisons, and opaqueness adjustments, an overlapping pattern of geographical markers and dialectal group boundaries was observed. For example, Figure 1 illustrates a relatively high level of overlap between the borders of the Yangtze River Basin (pink color) and the Hong Chao dialect subgroup region (green color) in the Jiang-Huai area. A geographic demarcation line largely connects the cities of Huai An, 淮安 (huái ān); Beng Bo, 蚌埠 (bèng bù); and Lu An, 六安 (lù ān) (red circles), and delineates the watershed between the Huai River and the Yangtze River Basin. It also largely coincides with the boundaries between the Central Plains Mandarin group and the Jiang-Huai Mandarin group (Berman, 2011; ILCASS et al., 2012, p. B1–B9).

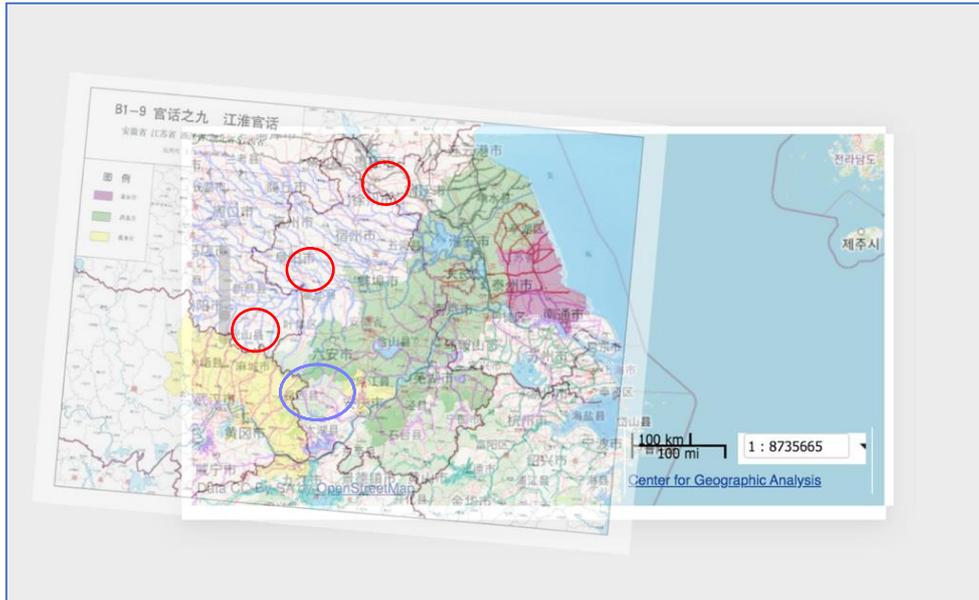


Figure 1. Overlaid maps of the dialect boundaries and watersheds of the Jiang-Huai area (Berman, 2011; ILCASS et al., 2012, p. B1–B9).

An overlapping pattern was also observed between the geographical markers and dialectal group boundaries in Shaanxi province. For example, as shown in Figure 2, in the northern area of the province, the contour of the Jin dialect group overlaps with the contour of the water drainage basin of the Luo River, 洛河(luò hé). In the southern area, the contour between the Central Plains Mandarin group and the Southwestern Mandarin group generally fits with the west-east direction of the Qin Ling Mountains as the geographical demarcation of north and south China. The contour between the Qinlong dialect subgroup and the Chuanqian dialect subgroup is aligned with the watershed borderline between the Jialing River, 嘉陵江(jiālíng jiāng), and the Han River, 汉江(hàn jiāng) (OCLGIMGCSP, 2011, p. 53; ILCASS et al., 2012, B1–B6, B1–B11, B1–B13).

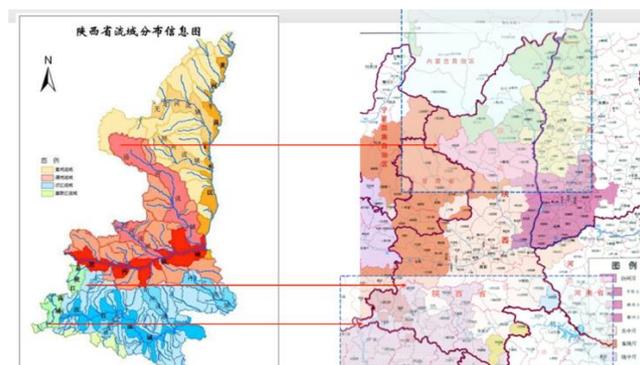


Figure 2. A comparison of the major water drainage basins and distributions of the major dialect groups of Shaanxi province (OCLGIMGCSP, 2011, p. 53; ILCASS et al., 2012, B1–B6, B1–B11, B1–B13).

Preliminary study findings validate Li and Xu (2019) and Lu’s (2019) arguments that mountains can act as natural barriers to dialect distribution. For example, the yellow area in Figure. 1. around An Qing, 安庆(ān qìng), (the blue circle in Figure 1) coincides with the

distribution of the Dabie Mountains, 大别山 (dà bié shān), on the boundary of the Hubei, Huanan, and Anhui provinces. This area also largely overlaps the Huang Xiao dialect subgroup.

The findings also partly substantiate Lu's (2019, p. 24, p. 56) argument that rivers function more to facilitate dialect diffusion than to isolate dialects because, along the rivers, the contours of dialect groups and subgroups do not strictly end with a river's drainage basin border. For example, due to the presence of many ferries, communication across the Yellow River is not entirely prevented by nature (Lu, 2019, p. 59). Therefore, the same dialects are often spoken on both sides of the river. For example, cities in northern Shaanxi, such as Fugu, 府谷 (fǔ gǔ); Shenmu, 神木 (shén mù); and Suide, 绥德 (suí dé), and across the river in Shanxi province, such as Xinshou, 忻州 (xīn zhōu), and Hequ, 河曲 (hé qǔ), all belong to the Wutai subgroup of the Jin dialect group.

The present findings generally support Chamberlain's (2005) conclusion that linguistic groups fall along geographic lines and the view that watersheds are a major factor in delineating micro- and macro-level language variations. Future studies in this area could yield valuable results that deepen our understanding of dialect typology.

References

- Berman, L. (Ed.). (2011). *China river basins*. Worldmap.harvard.edu. http://worldmap.harvard.edu/data/geonode:River_basin_num2
- Chamberlain, B. (2015). Linguistic watersheds: A model for understanding variation among the Tibetic languages. *Journal of the Southeast Asian Linguistics Society*, 8, 71–96.
- Coblin, W. S. (2002). Migration history and dialect development in the lower Yangtze watershed. *Bulletin of the School of Oriental and African Studies*, 65(3), 529–543.
- Institute of Linguistics, Chinese Academy of Social Sciences (ILCASS), 中国社会科学院语言研究所; The Institute of Ethnology and Anthropology (CASS), 中国社会科学院民族学与人类学研究所; and Language Information Sciences Research Centre of City University of Hong Kong (LISRCCUH), 香港城市大学语言资讯科学研究中心. (2012). *Zhongguo yu yan di tu ji (di er ban) hanyu fangyan juan* 中国语言地图集 (第2版) 汉语方言卷 [Language atlas of China, 2nd edition, Chinese dialects]. Beijing: Commercial Press.
- Li, F.-Y., & Xu, Y.-S. (1996). Yingxiang fangyan de dili yinsu chuyi 影响方言的地理因素刍议 [The influence of the geographical factors on the idiom]. *Human geography*, 11(2), 63–64.
- Lu, J.-P. (2019). *Han yu fang yan fen bu ge ju yu zi ran di li, ren wen di li de guan xi* 汉语方言分布格局与自然地理、人文地理的关系 [The relationship between the distribution pattern of Chinese dialects and physical geography and human geography]. Beijing: China Publishing House.
- Office of the Central Leading Group for Information Monitoring of the Geographical Conditions of Shaanxi Province (OCLGIMGCSP). (2011). *Shaanxi ji ben di li sheng qing bai pi shu* 陕西基本地理省情白皮书 (2011) [White paper on the basic geography of Shaanxi province (2011)]. Shaanxi: Shaanxi Provincial Bureau of Surveying, Mapping and Geographic Information.
- Simmons, R., Shi, R.-J., & Gu, Q. (2006). *Jianghuai guan hua yu wu yu bian jie de fang yan di li xue yan jiu* 江淮官话与吴语边界的方言地理学研究 [Chinese dialect geography: Distinguishing Mandarin and Wu in their boundary region]. Shanghai: Shanghai Education Press.
- Zhang, S.-Q., & Zhuang, C.-S. (2008). Guangdong fangyan de dili geju yu ziran dili jishi dili de guanxi 广东方言的地理格局与自然地理及历史地理的关系 [Geographical distribution of Guangdong dialects: Their linkage with natural and historical geography]. *中国文化研究所学报* [Journal of Chinese Studies], 48, 407–423.

Ke Wang: wang.11436@osu.edu

**Reflecting identity through Song lyrics:
Usage of variant graphs in popular Taiwanese Southern Min songs**

Paul D. Cockrum

The Ohio State University

Language serves as a distinctive marker of identity for groups of people, including written language (Milroy 1982). Individuals are able to use language in order to construct their desired identity. However, how exactly language reflects identity is not always agreed upon. From the Hong Kong Protests to the literary debates in Taiwan during the 1960s, the writing system of the Chinese language itself has often been at the center of these ideological debates (Gomes and Chan, forthcoming; Chang, 1999: 412). In Taiwan, the younger generation relies upon both Mandarin and Taiwanese to construct their linguistic identity (Liu et al. 2016:688). The younger generations attended school after the end of the Mandarin Language Policy in 1987, which allowed students to use Taiwanese or other indigenous languages in the classroom without punishment (Kloter 2005:132, 188-189.) In a multi-generational analysis, Todd Sandel (2003: 541) found that in individuals receiving their schooling since 1987, there is unanimous support for the acquisition of Taiwanese, although a rural-urban effect was found in various aspects of their language attitudes. This discussion is increasingly relevant in light of the 2019 National Languages Act, which aims to promote native Taiwanese languages, including Taiwanese, the Taiwanese variant of Hakka, and the Austro-Polynesian Taiwanese Aboriginal languages.

The interlaced questions of the written language and how to portray one's native identity, have remained central questions in the vitality and transmission of minority languages. The debate has led to multiple characters being used to represent the same word (Heylen 2001:146). The *Taiwanese Southern Min Commonly Used Words Dictionary* created by the Taiwanese Ministry of Education refers to these different variants as either *běnzì* 本字, *xùndúzi* 訓讀字, or *súzi* 俗字 (2008, *Bianji Jian Li* 編輯簡例). Graphs that contain a long history of usage with its referent are called *benzi*, translatable as 'original character,' and are often the same graphs that would be found in Modern Standard Mandarin. However, due to the complicated history of dialectal forms in Chinese, there has not often been an official standard of how to write characters. As a result, dialect speakers created locally acceptable graphs known as *suzi* [common characters] that are often unique to an individual dialect.

The current presentation takes three popular songs performed in Taiwanese as a case study to examine what strategies are used to assert their *běnsǎngrén* 本省人 or 'native born' identity. The native-born identity is operationalized through the usage of more *suzi* variants, which represent a refutation of the legitimatization of Mandarin (Kloter 2019: 203). The songs are *Waves Wandering* (*lōng liú lián* 浪流連) and *Back Here Again* (*lōng tsú huê thâu* 浪子回頭) by EggplantEgg 茄子蛋, as well as *Far Away Home* (*tshut guâ ê gín á* 出外的囡仔) by Nine Chen 陳零九. These songs were selected due to their popularity and modernity, thus being prime candidates for a case study on identity construction. The oldest of these songs, *Back Here Again*, was released in 2017. Additionally, all of the performers were born between 1987 and 1990, after the removal of the Mandarin only language policy, and as such, are part of the same age group and would have experienced similar schooling. However, the two are also opposed in a north and south divide, as the members of EggplantEgg all graduated high school in Taipei, whereas Nine Chen graduated high school in Kaohsiung before moving to Taipei.

In these songs, there are several examples of either distancing the performer from a Taiwanese native identity, such as in *Back Here Again* with the official music video choosing to use *benzi* instead of *suzi*. However, in EggplantEgg's other song, *Waves Wandering*, there is a preference for *suzi*, instead of *benzi*. One example of such are the graphs [薰] and [菸] *hun* which both mean 'to smoke', yet only the *benzi* [菸] would be readily accessible to a Mandarin speaker, and the *suzi* [薰] would be either an archaic or semantically odd choice. In *Far Away Home*, the choice of the characters *kiánn á* 囡仔 "child" instead of the *benzi* variant 兒子 conveys a strong connection with the local Taiwanese identity, due to a long history with the term *kiánn á* used similarly in many Taiwanese songs. As such, the choice to use the graphs [囡仔] *kiánn á* would recall the history of the past songs which also bore the same title (such as the song by Qi Lang 七郎 in 2008, who is also from Kaohsiung).

These different life experiences have created a different experience and attitude towards Taiwanese amongst the performers, which is then reflected in their music, as Kaohsiung-born Nine Chen will primarily perform in Mandarin—only referencing his *benshengren* identity in specific contexts such as *Far Away Home*. On the other hand, EggplantEgg continually asserts their Taiwanese identity through performing in Taiwanese, despite Mandarin being a much more popular language in cosmopolitan Taipei (Ding 2016:98). Their music, in turn, is a strong indicator of their artist's personal identity in the context of their environment—reflecting and expanding upon the rural-urban divide mentioned in Sandel (2003). In Kaohsiung, which is associated far more with a 'rural' context despite being Taiwan's second largest city, there are many speakers of Taiwanese and a stronger identification with the language—Kaohsiung Taiwanese served as the primary base of phonetic information for the Ministry of Education's *Commonly Used Words Dictionary* (2008, *Bianji Jian Li* 編輯簡例). Yet, Nine Chen moved from Kaohsiung to Taipei to pursue a singing and acting career, as such, he is under pressure to actively conform to the linguistic landscape of Taipei and attempt to market himself to a larger audience, thereby going with the linguistic pressure. Conversely, the members of EggplantEgg all grew up in Mandarin-dominated Taipei, and thus reject the dominance through performing in Taiwanese.

For both of the artists in the case study, we can see similar, yet distinct, identity formation strategies along a rural-urban divide. For the artist who came from a rural, Taiwanese-dominated context—Nine Chen—there is more pressure to conform to the dominance of Mandarin in the entertainment industry, which is reflected in his music predominantly being performed in Mandarin. Yet for his one song performed in Taiwanese, Nine Chen displays a deep knowledge of the local language variant, through his usage of a familiar title and *suzi* characters. On the urban side, EggplantEgg is almost the mirror image, with only a select few of their songs being released in Mandarin. In the context of Taipei, the usage of Taiwanese is a subversion of the language pressure and an overt assertion of the identity-building ability of language, as noted in Milroy (1982). Yet, EggplantEgg is able to switch between both *benzi* and *suzi* in the lyrics to their songs, depending on their imagined audience.

As illustrated by the case study of these two artists, the question of the graphic representation of Taiwanese remains important in modern times. As Taiwanese grasps with a bifurcated role as both a local, personal language as well as a legislated national language, graphic variants allow for an insight to personal language attitudes and goals (Liu et al. 2016). With the passage of the 2019 National Language Act, the status and utility of Taiwanese as an (inter-)national language remains in flux, yet hopeful.

References

- Chang, Sung-sheng Yvonne. 1999. Literature in Post-1949 Taiwan, 1950 to 1980s. *Taiwan: A New History* edited by Murray A Rubinstein. 403–418.
- Ding, Picus Sizhi. 2015. Taiwan: The Haven for Southern Min. *Southern Min as a Migrating Language*. Springer 55-75.
- Gomes, Skylor E. and Marjorie K.M. Chan. 2021. The 2019-2020 Hong Kong Protests: Dualling messages of the Authorities and the Protesters. *DEALL 50th Anniversary Celebration Volume: Current Issues in East Asian Languages and Literatures*, to appear.
- Heylen, Ann. 2001. Missionary Linguistics on Taiwan Romanizing Taiwanese: Codification and Standardization of Dictionaries in Southern Min (1837-1923). *Authentic Chinese Christianity: Preludes to its development (nineteenth and twentieth centuries)* 135-2001.
- Kloter, Henning. 2005. *Written Taiwanese (studia formosiana)*. Harrassowitz Verlag: Wiesbaden.
- Kloter, Henning. 2019. Writing Taiwanese: Then and now, how, why and who? In *Taiwan Studies Revisited* edited by Dafydd Fell and Hsin-Huang Michael Hsiao. 197-214.
- Liu, Yu-chang, Johan Gijssen and Chung-ying Tsai. 2016. Lower domain language shift in Taiwan: The case of Southern Min. *Folia Linguistica* 50.677–718
- Milroy, Lesley. 1982. Language and group identity. *Journal of Multilingual & Multicultural Development* 3(3).207-216.
- Jiaoyubu Guoyu tuixing weiyuanhui 教育部國語推行委員會 [Taiwanese Ministry of Education Committee on the Promotion of National Languages]. 2008. *Taiwan Minnanyu Changyongci cidian 台灣閩南語常用詞辭典* [Taiwanese Southern Min Commonly Used Words Dictionary]. Ministry of Education: Taipei.
- Sandel, Todd L. 2003. Linguistic capital in Taiwan: The KMT's Mandarin language policy and its perceived impact on language practices of bilingual Mandarin and Tai-gi speakers. *Language in Society* 32. 523-551.
- Wei, Jennifer M. 2013. Perspectives on Marked Language Choices and Uses in Taiwan. *Taiwan Journal of Linguistics* 11.67–82.

Paul D. Cockrum: cockrum.5@osu.edu

Session I-B
Abstracts

A preliminary survey on linguistic areas in East Asia based on phonological features

Ian Joo and Yu-Yin Hsu

The Hong Kong Polytechnic University

In this presentation, we will demonstrate the preliminary findings of our project aimed at identifying linguistic areas in East Asia. A linguistic area is here defined as geographically close languages sharing a high proportion of linguistic features not due to genealogical relatedness but due to historical contact. East Asia is defined as the area consisting of China, Japan, Korea, and Mongolia.

Based on 19 phonological features as binary parameters in 52 East Asian languages that we sampled, we calculated the Simple Matching Coefficient of the features between geographically close languages. The 19 phonological features, seven of them from the World Atlas of Linguistic Structures (WALS, Dryer and Haspelmath 2013), are listed in Table 1, with a column indicating whether each feature is borrowed from WALS. We drew a line between two languages if their geographical coordinates are within 1,500km distance and their 19 binary features show a Simple Matching Coefficient higher than 0.7.

Table 1: List of phonological features

Feature	WALS
Consonant Clusters	No
Consonant Inventories	Yes
Coronal Sonorants	Yes
Falling Diphthongs	No
Front Rounded vowels	Yes
Glottal Stop	No
Labiodental Fricatives	No
Long Vowels	No
Palatal Nasal	No
Plosive Codas	No
Retroflex Consonants	No
Tone	Yes
Uvular Consonants	Yes
Velar Fricatives	No
Velar Nasal Onset	No
Voiced Plosives	No
Voiceless Glottal Fricative	No
Vowel Nasalization	Yes
Vowel Quality Inventories	Yes

Figure 1 shows the preliminary results, suggesting some interesting patterns:

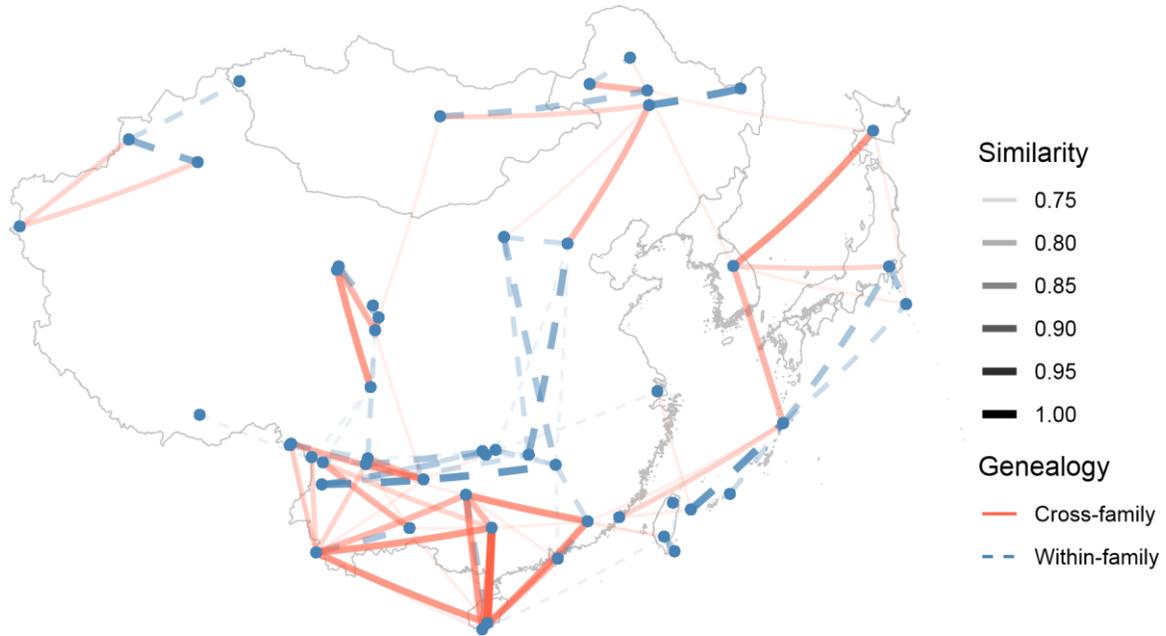


Figure 1: Connections representing phonological similarity between geographically close languages, across or within families

- The languages spoken in the Chinese provinces of Qinghai and Gansu share strong cross-family connections, as predicted by previous studies on the Qinghai-Gansu linguistic area (Xu 2017, Ch. 1, cf.).
- Languages in southwestern China are generally densely connected to each other, supporting the previous theories of the Mainland Southeast Asian linguistic area (Enfield 2018, cf.).
- Formosan languages show no similarity to Ryukyuan languages, despite their geographical proximity, in line with a genetic study demonstrating no genetic similarity between Taiwanese aboriginals and Ryukyuan islanders (Matsukusa et al. 2010).
- Manchu is connected to Mongolian and northern Sinitic languages (Mandarin and Hohhot Jin), in line with the historical contact between Manchu and these languages.
- Korean is most strongly similar to Ainu, and less so to Japonic languages and Dagur (Mongolic).
- Sarikoli, an Indo-European language spoken in northwestern China, show some connection with Turkic languages (Kazakh and Uyghur) spoken nearby.

Even though these observable patterns must be approached with caution given the preliminary stage of the data, they offer a promising outlook to our ongoing project and lead us to believe that with more features (other than phonological) examined, we will have a clearer view on linguistic areas within East Asia.

References

- Dryer, Matthew S. and Martin Haspelmath, eds. (2013). *WALS Online*. URL: <https://wals.info/>.
- Enfield, N. J. (2018). *Mainland Southeast Asian Languages: A Concise Typological Introduction*. Cambridge, UK ; New York, NY: Cambridge University Press.
- Matsukusa, Hirotaka et al. (2010). “A genetic analysis of the Sakishima islanders reveals no relationship with Taiwan aborigines but shared ancestry with Ainu and main-island Japanese”. In: *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 142.2, pp. 211–223.
- Xu, Dan (2017). *The Tangwang language: An interdisciplinary case study in Northwest China*. Cham: Springer.

Ian Joo: ian.joo@connect.polyu.hk

Yu-Yin Hsu: yu-yin.hsu@polyu.edu.hk

On the status of a tone merger in Dalian Mandarin

Wei William Zhou

The Ohio State University

This present study investigates the status of a tone merger in Dalian Mandarin,¹ a northern Mandarin variety spoken by about seven million native speakers. Specifically, it attempts to examine the merger of tones *yinping* and *qu* from Middle Chinese and determine the extent to which the two tones have merged.

In the 1960s, it was reported that Dalian Mandarin had four lexical tones (Song, 1963). This is presented in Table 1.

Table 1 Lexical Tones in Dalian Mandarin in the 1960s

	Ping	Shang	Qu	Ru
Yin	312	213	53	
Yang	34			

In more recent years, Liu (2012) and Qiu (2014), who respectively relied on the speech from one native speaker informant, reported only three citation tones in Dalian Mandarin. Since determining lexical tone inventory should be the prerequisite of more complex tone analyses, this study also begins by identifying citation tones in Dalian Mandarin, based on the speech data from a 26-year-old, female speaker reading 96 characters in carrier sentences. The speech of our informant is very representative of Dalian Mandarin, as verified by several native speakers. Moreover, she is from the same area as the informants surveyed in Liu's and Qiu's studies. Three citation tones are identified in this study, as shown in Table 2.

Table 2 Lexical Tones in Dalian Mandarin in this Study

	Ping	Shang	Qu	Ru
Yin	51	412	51	
Yang	24			

From Table 2, it can be seen that the tones for *yinping* and *qu*, which were distinctive in Song's (1963) study, have merged into one tone, namely, /51/. An examination of the pitch and register shows that the two tones are indeed identical. In addition, the length for the two tones (with controlled initials, voicing, and aspiration) also does not differ, contra Liu (2012), who reported that the tones derived from *yinping* were longer in duration than those from *qu*.

Although acoustically undistinguishable in isolation, the tones for *yinping* and *qu* exhibit different tone sandhi behaviors, suggesting that the merger is not complete (see also Liu, 2012). The data for the tone sandhi analyses was obtained from another reading task done by the same speaker reading 64 disyllabic words in carrier sentences. Essentially, when *yinping* and *qu* (51) occur as the first syllable of a disyllabic word and precede another *yinping* tone (51), they undergo a change to become a high-level tone (55). However, when *yinping* and *qu* precede another *qu* tone (51),

¹ For a detailed discussion of tone merger, refer to Lien (1986) and Wang (1987).

they do not undergo any tone sandhi change. Although both this study and Liu's (2012) study have dealt with the tone sandhi phenomenon, the specific configurations of the tone changes appear to be very different, and this will be elaborated upon in the presentation.

In conclusion, it appears that the merging of *yinping* and *qu* is nearly complete, as the two tones do not differ in their pitch, register, and length, but only in their tone sandhi patterning. The merger is likely to have progressed further than a decade ago, as exhibited in Liu's (2012) experiment, since the two tones no longer exhibit a length difference that was present in Liu's study. However, more research is still needed. First-hand data on Dalian Mandarin in the literature is scarce. This study is thus a small contribution in advancing our knowledge of this variety of Mandarin. At the same time, it contributes to enriching our growing understanding of the complexity of sound change and mergers that take place among the Chinese dialects.

References

- Lien, Chinfa. (1986). Tone Merger in the Dialects of Northern Chinese. *Journal of Chinese Linguistics* 14.2, 243-291.
- Liu, Te-hsin. (2012). The Phonology of Incomplete Tone Merger in Dalian. *Journal of Chinese Linguistics* 40.2, 362-396.
- Qiu, Mingyu (秋鳴宇). (2014). Dalian Fangyan Danzidiao Yanjiu (大連方言單字調研究) [A Study on Citation Tones in Dalian]. *Xiandai Yuwen* (現代語文) [Modern Chinese] 10, 15-18.
- Song, Xue (宋學). (1963). Liaoning Yuyin Shuolue (遼寧語音說略) [A Sketch of Liaoning Phonology]. *Zhongguo Yuwen* (中國語文) [China's Chinese] 2, 104-114.
- Wang, William S.-Y. (1987). A Note on Tone Development. In The Chinese Language Society of Hong Kong (ed.), *Wang Li Memorial Volumes* (English Volume). 435-443. Hong Kong: Joint Publishing Company.

Wei William Zhou: zhou.2874@osu.edu

Comparative analysis of grapheme-to-phoneme models for the Russian-Chinese parallel corpus¹

Alexandra Konovalova and Alena Tsvetkova

National Research University Higher School of Economics, Moscow

Introduction²

The Russian-Chinese parallel corpus of Ruscorpora (henceforth the Corpus) is an online corpus of texts provided with linguistic markup and meta-information. The Corpus contains 1070 text samples of literary works, news, and others. So far, it lacks a proper annotation of pinyin (Chinese transcription). Now the Grapheme-to-Phoneme (G2P) model applied to the Corpus is based on the dictionary CEDICT (Luo, Xu, Zhang, Ren and Sun, 2019). Thus, all the possible transcriptions are ascribed to each character without disambiguation. Therefore, the purpose of this study is to compare six G2P models to improve the quality of pinyin markup of the Corpus. Chinese G2P conversion seems to be challenging because of homophones and polyphones, which means that a character may have multiple pronunciations. Besides, the Chinese texts in the Corpus contain many phonetic borrowings from Russian which commonly are not included in dictionaries.

Related Work

Approaches to Chinese G2P conversion can be divided into rule-based (Wang, Chen and Yeung, 2004) and data-driven. Generally, rule-based methods suggest searching transcriptions of the words in a dictionary and mapping them to the text in accordance with the context. Although a set of rules is efficient for processing the majority of data, it faces problems with ambiguous characters. Data-driven approaches use statistical methods. Park and Lee (2020) proposed a developed dataset for polyphone disambiguation and trained a Bi-LSTM model on it. Chen, Zhao and Wang, (2015) argued that converting neural network language models into back-off n-gram language models helps to reduce computational cost. Following this approach, Cai, Yang, Zhang, Qin, and Li (2019) explored a bidirectional recurrent neural network. Recently, in Huang, Li, Zhang and Zhao (2018) and Zhang, Huang and Zhao (2019), an attention-based model which translates from pinyin sequence to Chinese character sequence was implemented.

Method

For this work, we explored the following models: A Context-aware Grapheme-to-Phoneme for Chinese (G2pC³) (2019), A Neural Grapheme-to-Phoneme Conversion Package for Mandarin Chinese Based on a New Open Benchmark Dataset (G2pM⁴) (2020), xpinyin⁵ (2002) and Chinese Pinyin Conversion Tool for Python (pypinyin⁶). All the methods tested are data-driven: xpinyin is based on stochastic decision lists and pypinyin uses n-gram statistics. G2pM is entirely built on sequence translating. Among those, the only model to use grammar information is G2pC which supports POS-tagging. One of the main challenges for disambiguating polyphones is correct word segmentation and POS-tagging. Therefore, we fine-

¹ The project was supported by the Commission of the Support of Educational Initiatives of the Faculty of Humanities within the framework of the Competition of Project Groups for Students (the name of the project is «Linguistic Markup of Chinese Texts in the Russian-Chinese Parallel Corpus of Ruscorpora»)

² Code is available on https://github.com/vydra-v-getrax/pinyin_annotation

³ <https://github.com/Kyubyong/g2pC>

⁴ <https://github.com/kakaobrain/g2pM>

⁵ <https://github.com/lxneng/xpinyin>

⁶ <https://github.com/mozillazg/python-pinyin>

tuned G2pC model with different state-of-the-art tools for parsing Chinese texts: FastHan⁷ (Geng, Yan, Qiu, and Huang, 2020), UDPipe⁸ (Straka and Straková, 2017, August) and pkuseg⁹ (Luo, Xu, Zhang, Ren, and Sun, 2019). For the test, we used a small dataset of 20 human-annotated sentences which were randomly selected from the Corpus and contained polyphones and proper names.

Results

Table 1 presents accuracy scores on the test dataset for each model.

Table 1. Metrics on evaluation dataset

Model	Accuracy
G2pC-pkuseg	0.903
G2pC-FastHan	0.899
G2pC-Udpipe	0.880
Xpinyin	0.861
Pypinyin	0.831
G2pM	0.824

(suffix *pkuseg/FastHan/UDPipe* refers to the tool for POS-tagging in G2pC annotator)

One of the typical mistakes is provoked by character 了 *le* or *liǎo*. For example, in the phrase 巴扎罗夫瞅了他一眼 ‘Bazarov looked at him’ it is pronounced as *le*, but algorithms annotated it as *liǎo*. As for the Russian-Chinese corpus, specific mistakes in the output of G2P models were not found. We plan to test this hypothesis on a larger dataset. Names and loan words are transcribed correctly because phonetic borrowings from Russian are translated using a limited set of Chinese characters. Nevertheless, linguistic issues common for written Chinese G2P should be addressed.

Based on our analysis, we revealed that the best algorithm for our Corpus is G2pC annotator on texts preprocessed with pkuseg package as it seems that for interpreting Chinese characters correct word segmentation and POS-tagging are crucial. Unlike UDPipe and FastHan, pkuseg includes multiple domain-specific segmentation CRF models. Therefore, superiority of this model is mainly due to pre-training on a large-scale, multi-domain dataset. In the future, we plan to proceed with experiments, by applying other tools for word segmentation, fine-tuning the algorithms to our data, and providing an extended evaluation dataset.

References

- Cai, Z., Yang, Y., Zhang, C., Qin, X., & Li, M. (2019). Polyphone Disambiguation for Mandarin Chinese Using Conditional Neural Network with Multi-level Embedding Features.
- Chen, S., Zhao, H., & Wang, R. (2015). Neural Network Language Model for Chinese Pinyin Input Method Engine. Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation, 455–461. <https://www.aclweb.org/anthology/Y15-1052>.
- Geng, Z., Yan, H., Qiu, X., & Huang, X. (2020). fastHan: A BERT-based Joint Many-Task Toolkit for Chinese NLP. arXiv preprint arXiv:2009.08633.
- Huang, Y., Li, Z., Zhang, Z., & Zhao, H. (2018). Moon IME: Neural-based Chinese Pinyin Aided Input Method with Customizable Association. Proceedings of ACL 2018, System Demonstrations, 140–145. <https://doi.org/10.18653/v1/P18-4024>.
- Luo, R., Xu, J., Zhang, Y., Ren, X., & Sun, X. (2019). Pkuseg: A toolkit for multi-domain chinese word segmentation. arXiv preprint arXiv:1906.11455.

⁷ <https://github.com/fastnlp/fastHan>

⁸ <http://ufal.mff.cuni.cz/udpipe>

⁹ <https://github.com/lancopku/pkuseg-python>

- Park, K., & Lee, S. (2020). g2pM: A Neural Grapheme-to-Phoneme Conversion Package for Mandarin Chinese Based on a New Open Benchmark Dataset. ArXiv:2004.03136 [Cs]. <http://arxiv.org/abs/2004.03136>.
- Semenov, K., Durneva, S., & Kuznetsova, Y. (2020). The Russian-Chinese Parallel Corpus of Ruscorpora: Achievements and Challenges. DHN2020: Parallel Corpora as Digital Resources and Their Applications. Retrieved from https://parallellcorporadhn2020.github.io/talks/Durneva_Kuznetsova_Semenov.html.
- Straka, M., & Straková, J. (2017, August). Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipes. In Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies (pp. 88-99).
- Wang, X., Chen, Q., & Yeung, D. S. (2004). Mining Pinyin-to-Character Conversion Rules From Large-Scale Corpus: A Rough Set Approach. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), 34(2), 834–844. <https://doi.org/10.1109/TSMCB.2003.817101>.
- Zhang, Z. R., Chu, M., & Chang, E. (2002). An efficient way to learn rules for grapheme-to-phoneme conversion in Chinese. In International Symposium on Chinese Spoken Language Processing.
- Zhang, Z., Huang, Y., & Zhao, H. (2019). Open Vocabulary Learning for Neural Chinese Pinyin IME. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 1584–1594. <https://doi.org/10.18653/v1/P19-1154>.

Alexandra Konovalova: askonovalova@edu.hse.ru

Alena Tsvetkova: a.tsvetkova@hse.ru

Phrasal Tonology in Suzhou Chinese: Some preliminaries

Yuhong Zhu

The Ohio State University

1. Introduction

This study is a continuation of previous research on tonal behaviors in Suzhou Chinese (Northern Wu; Zhu 2020; 2021). Cited work focused mostly on prosodic words that are disyllabic, and pointed at intriguing alternations among light-initial lexical items. Although a foot inventory following the tenets of Kager (1993) offers a systematic account of all disyllabic sandhi patterns, processes that apply to longer prosodic phrases are yet unexplained, and exactly what roles phrase structure may play (cf. Duanmu 1995, 1999) still remains unknown. In this study, I aim to investigate two phrasal tonal phenomena in Suzhou: pre-boundary L tone insertion, and morphology-sensitive footing.

2. Pre-boundary L tone insertion

Previous descriptive grammars and phonological accounts differ considerably in the treatment of a few sandhi patterns. One pair of examples is given in (1). /T/ stands for any lexical tone.

(1a). /HH/ + /T/ → [H.H] (Shi & Jiang 2013; my fieldwork)

(1b). /HH/ + /T/ → [H.L] (Ye 1979, 1993; Xie 1982; Wang 1996, 2011)

This apparent difference is potentially due to a L tone insertion rule at the end of Prosodic (or Intonational) Phrases: H → L /_]PhP/IP. This sort of edge-sensitive phonological processes is discussed in Selkirk (1984) for Xiamen Chinese (Southern Min), but is not yet examined in Suzhou. Moreover, the L tone insertion rule fails to apply to certain sandhi forms. Consider the pair in (2). In the current study, I aim to both determine the domain of application and blocking conditions of this L tone insertion rule.

(2a). /HH/ + /T/]PhP/IP → [H.L]

(2b). /LH/ + /T/]PhP/IP → [L.H]/*[L.L]

3. Morphology-sensitive footing

I use the cover term "morphology" to refer to patterns conditioned by phrase structures. Duanmu (1995) compares trisyllabic words in Shanghai (Northern Wu) and Taiwan (Southern Min) with the structure σσ-σ and σ-σσ (the dash stands for a word/morpheme boundary), attributing the differences across the two types to a prohibition of adjacent foot heads (related to stress clash avoidance effects). An example of Suzhou is shown in (3) (σ: metrical head; Ø: toneless).

(3a). ba.həu ti:
grocery store
H L HL/Ø
(σ.σ) (σ)
(σ.σ) σ

(3b). dəu vɛ:.ti:
big restaurant
L H Ø
(σ. σ) σ
* (σ) (σ.σ)

In (3a), it is possible for the monosyllabic morpheme *store* to form its own foot and preserve its underlying tones (unfooted syllables lose all phonological tones in Suzhou); in contrast, it is impossible for the monosyllabic morpheme *big* to form its own foot in (3b), since doing so will trigger a stress clash: *(σ)(σ.σ). Crucially, stress clash can neither be avoided by positing a second iamb, as all feet in Suzhou must be trochaic: *(σ)(σ.σ). The current project explores the role phrase structure can play in determining surface tonal patterns in Suzhou.

4. Methodology

All examples above come from my previous fieldwork, but they are preliminary in nature. Since both aspects are rarely, if ever, discussed in the relevant literature, the current project is also documentary work of an understudied language variety. I will analyze conversational speech data in the form of sociolinguistic interviews, and try to locate phrases and utterances that directly address my two research questions:

- (a) What is the domain and condition of the pre-boundary L tone insertion rule in Suzhou?
- (b) What role do different phrase structures play in determining the tonal output in Suzhou?

References

- Duanmu, San. 1995. "Metrical and Tonal Phonology of Compounds in Two Chinese Dialects." *Language* 71 (2): 225–59.
- Duanmu, San. 1999. "Metrical Structure and Tone : Evidence from Mandarin and Shanghai." *Journal of East Asian Linguistics* 8 (1): 1–38.
- Kager, René. 1993. "Alternatives to the Iambic-Trochaic Law." *NLLT* 11 (3): 381–432.
- Selkirk, Elisabeth. 1986. "On Derived Domains in Sentence Phonology." *Phonology Yearbook* 3 (May): 371–405. <https://doi.org/10.1017/S0952675700000695>.
- Shi, Xinyuan, and Ping Jiang. 2013. "A Prosodic Account of Tone Sandhi in Suzhou Chinese." In *Proceedings of the 25th North American Conference on Chinese Linguistics*. Education Press.
- Wang, Ping. 1996. *Suzhou Fangyan Yuyin Yanjiu [Phonetic Studies of Suzhou Dialect]*. Wuchang: Huazhong Ligong University.
- Wang, Ping. 2011. *Suzhou Fangyan Yanjiu [Suzhou Dialect Studies]*. Beijing: Zhonghua Shuju.
- Xie, Zili. 1982. "Suzhou Fangyan Liangzizu de Lianxu Biandiao [Tone Sandhi of Bi-Characters in Suzhou]." *Fangyan* 3: 117.
- Ye, Xiangling. 1979. "Zailun Suzhou Fangyan Shangsheng He Yinqu de Lianxu Biandiao [Revisiting the Shangsheng and Yinqu Tone Sandhi in Suzhou]." *Fangyan*, no. 4: 306–8.
- Ye, Xiangling. 1993. *Suzhou Fangyan Cidian [Suzhou Dialect Dictionary]*. Nanjing: Jiangsu
- Zhu, Yuhong. 2020. "Extending the Autosegmental Input Strictly Local Framework : Metrical Dominance and Floating Tones." In *Proceedings of the Society for Computation in Linguistics*, 3:393–401. <https://scholarworks.umass.edu/scil/vol3/iss1/38>.
- Zhu, Yuhong. 2021. "Moraic Footing in Suzhou Chinese: Evidence from Toneless Moras." In Ryan Bennett, Richard Bibbs, Mykel Loren Brinkerhoff, Max J. Kaplan, Stephanie Rich, Nicholas Van Handel & Maya Wax Cavallaro (eds.), *Supplemental Proceedings of the 2020 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.

Yuhong Zhu: zhu.1447@osu.edu

Session I-C
Abstracts

Deriving separable verbs in Cantonese

Ka-Fai Yip¹, Tommy Tsz-Ming Lee², and Sheila Shu-Laam Chan³

¹Yale University, ²University of Southern California, and ³The Chinese University of Hong Kong

Separable verbs.

In Cantonese, some disyllabic verbs allow verbal suffixes to occupy an infix position, e.g., (1)-(2) (Chan & Cheung forthcoming).

- (1) *feilou-zo sap-gei ci* fail-PFV ten-several time “failed a dozen times” (AB-x)
 (2) *fei<zo>lou sap-gei ci* fail<PFV> ten-several time =(1) (A-x-B)

The two syllables (referred to as A and B) of the disyllabic *monomorphemic* verb *feilou* ‘fail’ (an English loanword) is separated by perfective suffix *zo* (referred to as x). The same is true of other monomorphemic verbs like *pisen* ‘present’ and *sowi* ‘sorry’. They are known as *separable verbs* and seem to host infixation, but the separation pattern is not exclusive to affixal elements. That phrasal elements like frequency phrases may appear after *zo* in (3) suggests that genuine infixation is not involved.

- (3) *fei<zo><sap-gei ci>lou* fail<PFV><ten-several time> =(1) (A-x-XP-B)

Against a reanalysis approach.

A (syntactic) reanalysis approach suggests that the disyllabic verbs like *feilou* are indeed reanalyzed as a Verb-Object (VO) phrase, a prevailing approach for separable VO compounds in Mandarin (Chao 1968, Huang 1984, Packard 2000, *i.a.*). However, the second syllable in disyllabic verbs barely displays nominal/object properties. First, it cannot be preceded by the nominal modifier marker *ge* (= *sen* in (4)), which is otherwise allowed on a genuine object (= *hei* ‘movie’ in (5)).

- (4) *pi-zo sam ci (*ge) sen* present-PFV three time MOD present “presented three times.”
 (5) *tai-zo sam ci (ge) hei* look-PFV three time MOD movie “watched movies three times”

Also, the second syllable does not saturate the thematic requirement of the transitive verb, since the verb can still take a (preposed) thematic object (= (6)), which is surprising if the second syllable is reanalyzed as an object (*cf.* the true object in (7)).

- (6) *zeong ni-fan-je pi<zo>sen* DISP this-CL-thing present<PFV> “presented on this document”
 (7) *(*zeong ni-coet-hei) tai-zo hei* DISP this-CL-movie look-PFV movie Int.: “watched this movie”

The proposal.

We motivate a hybrid (syntactic + phonological) approach to separable verbs. Assuming that verbal suffixes head a projection above the verb, we suggest that verbal suffixation generally involve syntactic verb movement to the suffix (Tang 2003, Tsai 2001). Crucially, we propose an optional PF deletion rule triggered by affixes in (8).

(8) Affix-induced Syllable Deletion

Affixes optionally trigger deletion on an *adjacent* syllable of their hosts.

Implementation.

Under the copy theory of movement (Chomsky 1995 *et seq.*), verb movement creates two copies, i.e., (9)a-b. If (8) does not apply, copy deletion will erase the lower copy, giving rise to (9)ci as suffixation. If (8) applies, the second syllable (i.e., B adjacent to the suffix) is deleted as in (9)cii(I). Then when copy deletion applies to the lower copy, it only *partially* deletes the complement syllable (i.e., A) for recoverability. A suffix is now sandwiched between A and B (=9)cii(II)). It immediately explains why the following patterns in (10) are unattested.

(9) Derivation steps for suffixation and “infixation”

- a. [-x [AB]] (base structure)
- b. [<AB>-x [<AB>]] (verb movement)
- ci. [<AB>-x [<AB>]] = AB-x (copy deletion targeting the lower copy, deriving (1))
- cii. (I) [<AB>-x [<AB>]] (affix-induced syllable deletion in (8))
- (II) [<A>-x [<AB>]] = A-x-B (partial copy deletion, deriving (2))

	adjacent syll. deletion	recoverability	copy deletion
(10)a. * <i>lou</i> < <i>zo</i> > <i>fei</i> (*B-x-A)	✗ non-adjacent deletion		
b. * <i>fei</i> < <i>zo</i> > <i>fei</i> (*A-x-A)		✗ unrecoverable	
c. * <i>lou</i> < <i>zo</i> > <i>feilou</i> (*B-x-AB)	✗ non-adjacent deletion		✗ fail to apply
d. * <i>fei</i> < <i>zo</i> > <i>feilou</i> (*A-x-AB)			✗ fail to apply

Deriving (3).

We assume that the frequency phrase is (left-)adjoined to the vP, below the projection headed by the suffix, as in the base structure in (11)a. Crucially, (11)b indicates that the verb moves across the frequency phrase to head-adjoin to the suffix. Then (8) applies and deletes *lou*, followed by the partial copy deletion on *fei* only.

- (11)a. [_{AspP} -*zo* [_{vP} *sapgei ci* [_{vP} ... [_{VP} *feilou*]]]] (base structure)
- b. [_{AspP} <*feilou*>-*zo* [_{vP} *sapgei ci* [_{vP} ... [_{VP} <*feilou*>]]]] (verb movement)
- c. [_{AspP} <*feilou*>-*zo* [_{vP} *sapgei ci* [_{vP} ... [_{VP} <*feilou*>]]]] (by (8) and partial copy deletion)

An extension to prefix.

It is argued that the *lin...dou*-focus construction can target verbs, leading to verb doubling (Shyu 1995, Cheng & Vicente 2013). In cases of disyllabic monomorphemic verbs, it is possible to double the whole verb A+B (as in (12)). Crucially, it is possible for B to occupy the higher position (stranding A), but not vice versa, as contrasted in (12) and (13). The pattern in (13) mirrors (2) and immediately follows from the proposal if we assume *lin* is a *prefix*: by (8), *lin* triggers syllable deletion on A (adjacent to *lin*), and copy deletion partially applies to the lower copy, deleting B. (14) is disallowed since syllable deletion is not adjacent.

- (12) *Lin pisen keoi dou mou pisen* (lin-AB ... AB)
- even present 3SG also NEG.PFV pre(sent) “He even didn’t do the presentation.”

- (13) *Lin -sen keoi dou mou pi-* (lin-AB ... AB)
 even (pre)sent 3SG also NEG.PFV pre(sent) =(12)
- (14) **Lin pi- keoi dou mou -sen* (lin- AB ... AB)
 even (pre)sent 3SG also NEG.PFV pre(sent)

Implications.

(a) A monosyllabic preference in Cantonese for verbs has been noted (Tang 2002, 2003, Li et al. 2016) and (8) may potentially extend to other cases in the phenomenon: (i) syllable deletion may be obligatory in some environments (e.g., *-dak* and *-ngaang* suffixation, Tang 2002, 2003); (ii) syllable deletion may apply recursively (e.g., in A-not-A formation); (iii) idiosyncratic properties of verbs may show varying resistance to deletion (e.g., different registers, frequency, morpho-phonological structures).

(b) We offer novel evidence for a non-lexicalist view that verbal suffixes are *syntactic heads* (*contra*. Gu 1993, Huang et al 2009).

(c) Affixation, at least in Cantonese, is achieved not by lowering but *head raising* (*contra*. Cheng, Yi & Xiong 2016), which creates copies for partial deletion.

(d) Copy deletion interacts with PF operations like (8) and can be disturbed (*cf.* Lee 2020), leading to word-level partial deletion (*cf.* Fanselow & Cavar 2002).

Selected References.

- Chan**, S. S.-L., & L. Y.-L. **Cheung**. Forthcoming. “Morpho-syntax of non-VO separable compound verbs in Cantonese”. *Studies in Chinese Linguistics* 41(2).
- Tang** S.-W. 2003. “Properties of Ngaang and the Syntax of Verbal Particles in Cantonese.” *Journal of Chinese Linguistics*. 31(2), 245-269.
- Li**, D. C. S., C. S. P. **Wong**, W. M. **Leung**, and S. T. S. **Wong**. 2016. “Facilitation of transference: The case of monosyllabic salience in Hong Kong Cantonese.” *Linguistics*. 54(1), 1–58.

Ka-Fai Yip: kafai.yip@yale.edu

Tommy Tsz-Ming Lee: tszming1@usc.edu

Sheila Shu-Laam Chan: sheilachan@cuhk.edu.hk

Two types of temporal adverbial clauses in Cantonese

Ka-Fai Yip
Yale University

The typology of adverbial clauses

Adverbial clauses, according to their internal and external syntax, may be dichotomized into two groups: central and peripheral adverbial clauses (Haegeman 2003, 2010 *i.a.*). Temporal adverbial clauses (TACs) belong to central adverbial clauses and are assumed to have operator movement (e.g. *when*) and a lower attachment site to the main clause (Geis 1970, Larson 1987, Haegeman 2009). In Cantonese, TACs may be headed by either *hai* ‘at’ in (1) or *dong* ‘at, while’ in (2):

- (1) *Hai Aafan fan-gan gaau gozan, Aaming lai wankeoi* (hai-TACs)
 at Fan sleep-PROG nap that.time Ming come find 3SG
 ‘M. came find F. when she was sleeping.’
- (2) *Dong Aafan fan-gan gaau gozan, Aaming lai wan keoi* (dong-TACs)
 while Fan sleep-PROG nap that.time Ming come find 3SG
 ‘M. came find F. while she was sleeping.’

However, it is observed that *hai*-TACs and *dong*-TACs behave differently in both internal and external syntax. Specifically, while *hai*-TACs act like central adverbial clauses, *dong*-TACs do not. This suggests that temporal adverbial clauses, classified by a *semantic* label, may not be a homogeneous *syntactic* class.

Proposal

I propose that both types of TACs contain a null temporal operator comparable to *when* (hereafter OP_{temp}) but they differ in the position of the operator. In *hai*-TACs, OP_{temp} is base-generated in the TP domain and undergoes operator movement to CP to form an adverbial clause, schematized in (3). In *dong*-TACs, in contrast, OP_{temp} is base-generated higher at CP and does not require further movement as in (4).

- (3) Hai-TACs with operator movement: (4) Dong-TACs with operator base-generated at CP
hai [_{CP} OP_{temp} [_C ... [_{TP} ... *t* ...]]] *dong* [_{CP} OP_{temp} [_C ... [_{TP} ...]]]

I show how this proposal may capture the internal syntax of the two types of TACs by locality restrictions and minimality effects associated with the operator movement. Further evidence for the existence of OP_{temp} comes from an agreement marker *-haa*. I also discuss how the proposal may relate the internal and external syntax of TACs in terms of the height of operator generation sites and attachment sites to main clauses.

High-low reading and locality

First, *hai*-TACs are ambiguous between a ‘high’ reading and a ‘low’ reading when they further embed a CP (*cf.* English as Larson 1987 & Mandarin as Liou 2003). (5) contains an odd ‘high’ reading that I am alive for thousands of years. Yet, it also allows for a felicitous ‘low’ reading that I am alive in 2012, the foretold year of apocalypse.

- (5) *Hai*<sub>[CP Maangaa jan jyujin [CP saigaai wui waimit]] gozan, ngo zung saangngaungau.
 at Maya people foretell world will destroy that.time 1SG still alive
 i. #‘I’m still alive at the time when Maya people made the apocalyptic prophecy.’ (high read.)
 ii. ‘I’m still alive at the time when Maya people predicted to be the end of the world.’ (low read.)</sub>

Notably, the ‘low’ reading disappears when the lower CP is contained in an NP:

- (6) #*Hai* [Maangaajan gong]<sub>[NP [CP saigaai wui waimit] ge jyujin]] gozan, ngo zung saangngaungau.
 at Maya ppl. say world will destroy MOD prophecy that.time 1SG still alive
 #‘I’m still alive at the time of Maya people making the apocalyptic prophecy.’ (only high r.)</sub>

Under (3), the (non-)availability of the ‘low’ reading can be explained by long distance movement of OP_{temp} from the lower CP to the higher CP which is barred from a complex NP island. Contrarily, *dong*-TACs can only be interpreted ‘high’ and they lack a ‘low’ reading as in (7):

- (7) #*Dong*<sub>[CP Maangaajan jyujin [CP saigaai wui waimit]] gozan, ngo zung saangngaungau.
 while Maya ppl. foretell world will destroy that.time 1SG still alive
 #‘I’m still alive at the time when Maya people made the apocalyptic prophecy.’ (only high r.)</sub>

(7)’s reading follows directly from (4) where OP_{temp} is merged directly to the (higher) CP without movement.

Quantificational elements and minimality effects

Second, occurrence of identificational focus marked by *hai* ‘be’ (bears a different tone with *hai* ‘at’) is disallowed in *hai*-TACs yet is allowed in *dong*-TACs, contrasted in (8). The contrast carries over to ‘only’ focus *dak* and epistemic modal *jinggoi* ‘should’ etc.

- (8) **Hai*^{OK}*dong* [*hai-Aaming* fangaau] gozan, lousi zau faatnau.
 at while be-Ming sleep that.time teacher then become.mad
 ‘The teacher became mad *when/^{OK}while it was MING (but not someone else) that fell asleep.’

The ban on focus (and modals) could be explained by Rizzi’s feature-based Relativized Minimality (2001, 2004), where a Qu(antificational) element may disrupt a syntactic dependency like *wh*-movement. Both focus and modals are classified as Qu elements and they also block ‘why’ and A-not-A dependencies in Chinese (Law 2001, Soh 2005, Hagstrom 2006, Tsai & Yang 2015). Hence, the contrast in (8) can be accounted for if *hai*-TACs involve operator movement while *dong*-TACs do not. Qu elements induce minimality effects and disrupt OP_{temp} movement in *hai*-TACs but not in *dong*-TACs, diagrammed below.

- (9) Minimality effects in *hai*-TACs: (10) Lack of minimality effects in *dong*-TACs
 **hai* [CP OP_{temp} [C ... Z_[+Qu] [TP ... t ...]]] *dong* [CP OP_{temp} [C ... Z_[+Qu] [TP ...]]]

Note that only Qu elements *higher* than TP would block the movement in (9). Deontic modals, for example, are lower than TP (Tsai 2015) and may occur in *hai*-TACs. These low Qu elements can only induce minimality effects if OP_{temp} originates *lower* than them, e.g. in an embedded CP. Hence, a *hai*-TAC with a deontic modal is predicted to lack a ‘low reading’. The prediction is borne out:

- (11) #*Hai*_{[CP Maangaa jan [hoji] jyuujin [CP saigai wui waimit]] gozan, ngo zung saangngaungau.}
 at Maya people can foretell world will destroy that.time 1SG still alive
 #‘I’m still alive at the time when Maya people could make the apocalyptic prophecy.’ (only high r.)
- (12) *hai* [CP **OP**_{temp} [C ... [TP *f*_{high} ... Modal^{Deontic}_[+Qu] [CP ... *t*_{low}]]]] (minimality blocks long-d. mvt.)
-

Temporal agreement marker -*haa*

The existence of null *OP*_{temp} in TACs receives further support from *-haa*, a progressive suffix on reduplicated verbs (Matthews & Yip 1994). Clauses with *-haa* lose clausal independency and must attach to a main clause (= (13)). While it may be reminiscent of English participle *-ing*, the subordinate clause in (13) differ from *-ing* in bearing only temporal reading (i.e. they are TACs), resisting of conditional, causal or concessive interpretation even if the respective subordinators are added.

- (13) *Aafan fanfan-haa gaau, *(Aaminglai wankeoi)*
 Fan RED.sleep-HAA nap Ming come find 3SG
 ‘Ming came find Fan when she was sleeping.’

The obligatory TAC formation by *-haa* can be explained if we assume that it establishes an agreement relation with the temporal operator. One argument comes from the ban on *low* Qu elements in TACs with *-haa*. Negation is disallowed in (14), which can be made grammatical by replacing *haa* with another progressive suffix *-gan*. The same is true for *lin* ‘even’-focus, deontic modals and quantificational adverb ‘often’, etc. Crucially, they are also banned in *hai*-TACs and *dong*-TACs once *-haa* is present.

- (14) * (*Hai/dong*) [*Aafan* *m-hai* *fanfan-haa gaau*]gozan, *Aaming lai wankeoi*
 at while Fan NEG-be RED.sleep-HAA nap that.timeMing come find 3SG
 Int.: ‘Ming came find Fan when she wasn’t sleeping.’

The surprising constraint on low Qu elements in *hai/dong*-TACs with *-haa* supports an agreement analysis, where the syntactic dependency between *-haa* and *OP*_{temp} may be disrupted by minimality effects:

- (15) ‘Extra’ minimality effects in *hai*-TACs and *dong*-TACs with *-haa*:
 hai/dong* [CP **OP_{temp} [C ... [TP (*f*) ... Z_[+Qu] ... [AspP *-haa*_{temp} ... [vP ...]]]]] (agree blocked)
-

Internal syntax correlates with external syntax

Apart from internal syntax, *hai*-TACs and *dong*-TACs are also different in external syntax. Only *hai*-TACs, but not *dong*-TACs, may follow a main clause subject (also attested in Mandarin, see Wang & Feng 2015):

- (16) *Aaming* [^{OK}*hai*/**dong* [*Aafan fan-gan gaau*]gozan] *lai wan keoi*
 Ming at while Fan sleep-PROG nap that.timecome find 3SG = (1)/(2)

(16) is indicative of their attachment sites to main clauses: while *hai*-TACs may occur *within* a TP and follow a subject (presumably at Spec,TP), *dong*-TACs must occur *outside* a TP, i.e. it attaches

to a CP. Together with their generation sites of OP_{temp} , a correlation between internal and external syntax can be drawn: If the operator originates from TP, the TACs (i.e. *hai*-TACs) may attach to a TP; whereas if the operator base-generates higher at CP, the TACs (i.e. *dong*-TACs) will attach to a bigger clause CP rather than TP. The height of operator sites seems to determine the attachment sites (*cf.* Endo & Haegeman 2019).

Implications

(a) Temporal adverbial clauses, at least in Cantonese, are not uniformly central adverbial clauses and differ in their base-generation sites of temporal operators and attachment sites to main clauses. The distinction may extend to Mandarin, or even other languages like English where *while*-clauses are reported to lack a ‘low’ reading, unlike *when*-clauses (Larson 1990).

(b) Cantonese TACs lend potential support to Endo & Haegeman (2019)’s claim that the internal syntax of adverbial clauses determines their external syntax.

Selected references

- Endo, Yoshido, and Liliane Haegeman. 2019. Adverbial clauses and adverbial concord. *Glossa* 4(1): 48.1-32.
- Haegeman, Liliane. 2010a. The internal syntax of adverbial clauses. *Lingua* 120:628-648.
- Liou, Yu-fen. 2003. On Chinese temporal clauses. Doctoral Dissertation, Michigan State University.
- Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti, ed., *Structures and Beyond*, 223-251. Oxford: Oxford University Press.

Ka-Fai Yip: kafai.yip@yale.edu

Character Writing, Lexicon and Syntax in 19th Century Cantonese: Some Observations from Bridgman (1841)

Jinwei Ye, The Ohio State University

Bridgman's *A Chinese Chrestomathy in the Canton Dialect* (1841) is the first and most comprehensive record of colloquial Cantonese, which contains thematic collections of Cantonese phrases and conversations about a wide range of topics: human body, kindred relations, classes of men, commercial affairs, mechanical affairs, etc. The book is a precious documentation of the 19th century local vernacular, given that the dominant written variety of Chinese was still *wen2yan2wen2* 文言文 'literary Chinese' at the time.

Chapter 5 *jat6jung6pin1* 日用篇 'Domestic Affairs' from the book is chosen for characterizing 19th century Cantonese. As the name of the chapter suggests, this chapter covers topics which are also discussed today. Character writing, lexicon and syntactic structures are examined to find out (i) differences in Chinese character writing, lexicon and syntactic structures between the 19th century variety and the variety today and (ii) the implication of these differences.

The study has the following findings: (i) In terms of character writing, graphic variants are very common. *Faan4tai2* 繁體 'traditional variants' are dominant, and slight differences between graphic variants of the same character are observable, e.g. some additional strokes (1a) or difference in the shape of a stroke (1b), or an additional radical for one variant over the other (1c). (ii) As for lexicon, some common nouns, verbs, idiomatic expressions and loan words have undergone semantic shifts, while others retained their core meaning. On the other hand, Bauer's (2018: 127) findings of Hong Kong Cantonese on borrowing Chinese characters for transcription apply to loan words in the 19th century variety. (iii) Syntactically, A-not-A questions show diversified structures in the text: VP-NEG-V (2a), VP-NEG (2b) and V-NEG-V (2c). Double object constructions contain a particle right before the indirect object (3). These two syntactic findings are also observed in Tai (2009). The study also discovers an undiscussed construction: the *gu2* 估 "suppose" + sentence construction today appears as sentence + *gu2* in 19th century Cantonese (4).

The above findings are indicative in three ways: (i) from character writing, it is evident that the lack of standardized writing practice in Cantonese can be traced back to as early as the 19th century; (ii) the special lexicon and syntactic structures may be the result of interaction between Cantonese and literary Chinese; (iii) the uncertain grammatical status of *gu2* 估 'suppose' in the "suppose" construction remains to be verified, although in this study it is assumed to be a verb.

To further the study, additional language data from literary Chinese are necessary for comparison with Cantonese language materials. It is hoped that additional linguistic evidence will provide insights into understanding the "suppose" construction in 19th century Cantonese.

Examples

- (1) a. (19th c. variants) 番, 番 *faan1* ‘again’
 b. (19th c. variant)  *sam1* ‘heart’
 (21th c. variant) 心
 c. (19th c. variant)  *wui3* ‘(modal) can’
 (21th c. variant) 會
- (2) a. 重 要 的 乜野 添 唔 要 呀 (VP-NEG-V)
zung6 jiu3 di1 mat1je5 tim1 m4 jiu3 aa3
 still need bit what thing TIM NEG need SFP
 ‘Do you need anything else?’
 b. 佢哋 有 賞 冇 呢 (VP-NEG)
keoi5dei6 jau5 soeng2 mou5 ne1
 3PL have reward NEG SFP
 ‘Have they any rewards?’
 c. 你 試 吓 個 件 帽 睇 中意 唔 中意 (V-NEG-V)
nei5 si3 haa5 go2 gin6 mou2 tai2 zung1ji3 m4 zung1ji3
 2SG try HAA5 that piece hat see like NEG like
 ‘Just try on this bonnet and see if it will suit you.’
- (3) 遞 隻 過 我 (V DO Particle IO)
dai6 zek3 gwo3 ngo5
 send CL EXP 1SG
 ‘Send me one of them.’
- (4) 我 睇 得 你 杯 茶 濃 估 呀 (19th c. variety)
ngo5 tai2 dak1 nei5 bui1 caa4 nung4 gu2 aa3
 1SG see DAK 2SG CL tea strong suppose SFP
 ‘I fear, sir, your tea is too strong.’

Select References

- Bauer, Robert. (2018). Cantonese as written language in Hong Kong. *Global Chinese*, 4(1), 103-142.
- Bridgman, Elijah Coleman. (1841). *A Chinese Chrestomathy in the Canton Dialect*. S. Wells Williams: Macao.
- Tai, Chung Pui 戴忠沛. (2009). Characteristics of mid-19th century Cantonese as reflected in Su Hua Qing Tang 《俗話傾談》反映的 19 世紀中粵方言特徵. In Chi-on Andy Chin 錢志安, Bit-Chee Kwok 郭必之, Po Lun Peppina Lee 李寶倫, & Ka Yin Benjamin T'sou 鄒嘉彥 (eds.), *Multi-disciplinary Approaches to Cantonese Studies: Papers from the 13th International Cantonese Studies 粵語跨學科研究: 第十三屆國際粵方言研討會論文集*, pp. 245-260. Hong Kong: City University of Hong Kong.

Doubling of perfective aspect markers in Hui Chinese

Xuan Ye¹ and Yao Chen²

¹The Ohio State University and ²Fujian Normal University

Research on the doubling of synonymous function words in standard Mandarin Chinese mainly involves system-internal grammaticalization (Liu 2001, Jiang 2008, Zhang 2012). That does not appear to be the case in the doubling of perfective markers in the Hui (徽) dialect of Chinese,¹ spoken in Anhui and Jiangxi provinces. It is proposed here that this phenomenon in Hui Chinese can be attributed to language contact, and more specifically, to contact with standard Mandarin Chinese. This paper provides evidence for such a proposal, based on historical evidence and on data collected in Hui-speaking areas in 2019.

Both *zhe* 着 and *le* 了 are used as perfective aspect markers in the Hui dialect. In contrast, in modern Standard Mandarin, only *le* is used as a perfective aspect marker, while *zhe* is used as a durative aspect marker (Li & Thompson 1981). According to Mei (1988), in the case of northern Mandarin, this process of the division of labor of *le* and *zhe* took place during the Tang dynasty. Nevertheless, ever since the Song dynasty, dialects along the Yangtze River use *zhe* both as a perfective aspect marker and as a durative aspect marker. Long and Sun (2013), however, hold a different view, namely that in the Song dynasty, *zhe* was not used as a perfective aspect marker, but only as an aktionsart, with the semantic meaning of “completion”. In this presentation we will apply the concept of perfectivity broadly to particles that suggest the meaning of “completion” and “realization.”

In the Hui dialect, not only can both *zhe* and *le* be used as perfective markers, but they can co-occur in the same syntactic position as well (Hirata 1998). The different usages of *zhe* and *le* as perfective markers in Hui is demonstrated in Table 1². An example of *le-zhe* co-occurrence in the Qimen Hui dialect 祁门徽语 is shown in (1).³ As *le* was not originally employed as a perfective aspect marker in Hui, the most likely source of this usage in modern Hui is standard Mandarin. This phenomenon can thus be regarded as the result of language contact between Hui and standard Mandarin.

(1) 我跟渠都讲了着。(祁门徽语 Qimen Hui)

a42 kuẽ11 tɛi55 tu11 kɔ̃42 lia42-tʂo0.

Wo gen qu dou jiang le-zhe.

1st-SG with 3rd-SG already tell le-zhe

‘I already told him that.’

1 Hui Chinese, or the Hui dialect, is one of the southern dialects of Chinese. It is mainly spoken in Xin’anjiang and its surrounding area which is to the south of Yangtze River. The specific counties and cities speaking different subvarieties of the Hui dialect include Jixi, Jingde, Shexian, Tunxi, Xiuning, Yixian, Qimen and Shitai in Anhui Province, as well as Wuyuan in Jiangxi Province (Meng 1997).

2 Table 1 models Hirata’s sample sentences (1998: 277), and the results are collected from Chen’s fieldwork in Anhui and Jiangxi in 2019.

3 The example is from Chen’s fieldwork.

Such an analysis is supported by both historical evidence and subdialectal evidence. Historically speaking, migrants from Northern China occupied the Hui-speaking area from the Northern and Southern dynasties to the Song dynasty, followed by migrants who speak Jianghuai Mandarin during the Ming and Qing dynasties (Campbell 2004). As for the subdialectal evidence, geographical distribution of the aspect markers among various subvarieties of Hui supports the proposed language contact between Hui and standard Mandarin.

Table 1 presents the usage of *le* and *zhe* in different subdialects of Hui, based on geographic distribution from north to south. As Table 1 shows, there are significant differences between northern Hui subdialects on the one hand, and central and southern Hui subdialects on the other, in the choice of *le* and/or *zhe*. The basic distributional pattern is as follows. In northern Hui subdialects, *le* is now used as the perfective aspect marker, replacing the original marker, *zhe*. A different pattern emerges in central and southern Hui subdialects. Although *zhe* is retained, it has a competing form involving aspect-doubling of *le* and *zhe*, with *le-zhe* used in central Hui dialects, and *zhe-le* in southern Hui dialect.

Table 1. The Use of Perfective Aspect Markers in Hui Dialects

Mandarin	我买了一个碗。	他吃了饭了。
Hui dialect subvarieties	Wo mai <i>le</i> yi ge wan. 1 st -SG buy PFV one QTF bowl I have bought a bowl.	Ta chi <i>le</i> fan <i>le</i> . 3 rd -SG eat PFV rice SFP He has eaten.
Meixi (northernmost)	我买了一只碗。	渠饭吃了。
Jixi (northern Hui)	我买了一只碗。	渠饭吃了。
Shexian (northern Hui)	我买了一只碗。	渠饭吃了。
Tunxi (central Hui)	我买着一只碗。	渠吃饭了着。
Xiuning (central Hui)	我买着一只碗。	渠吃饭了着。
Qimen (central Hui)	我买着一个碗。	渠吃饭了着。
Wuyuan (southern Hui)	我买着一只碗。	渠饭吃着了。

The results suggest that the northern Hui subdialects are the most greatly influenced by standard Mandarin Chinese, since the perfective aspect marker *le* has supplanted the original element *zhe* in Hui dialects. In the case of the central and southern Hui subdialects, however, *le* has not completely supplanted, the native form, *zhe*, resulting in the retention of *zhe* occurring side-by-side with the overlaying of *le* and *zhe*, either as *le-zhe*, in central Hui dialects, or *zhe-le*, in southern Hui dialects.

Thus, it is proposed here that the co-occurring perfective markers in the Hui dialect is the result of language contact with standard Mandarin Chinese. Hui Chinese originally used *zhe* as the perfective marker. However, due to influence from the standard language's use of *le* as the perfective marker, the two aspect markers—the local and the standard—competed within the same syntactic position. The consequence is the mixed distribution pattern that we see today, with the inconsistent usages of *le* and *zhe* in the different subvarieties of Hui Chinese. The native aspect marker *zhe* is supplanted by *le* in northern Hui dialects. In central and southern Hui dialects, although *zhe* is retained there, its territory is, nonetheless, encroached upon by a competing form involving the aspect-doubling of *le-zhe* in central Hui dialects, and *zhe-le* in southern Hui dialects.

What does the future hold for the Hui dialects? Will *zhe* survive as the stand-alone perfective aspect marker? Or will the *le-zhe* and *zhe-le* eventually replace it? Will the aspect-doubling survive, or will those forms be replaced by *le* in central and southern Hui dialects, as it has already in northern Hui dialects? Only time will tell.

References

- Campbell, Hilary. (2004). *Chinese Grammar—Synchronic and Diachronic Perspectives*. Oxford University Press.
- Hirata, Shōji 平田昌司. (1998). *Huizhou Fangyan Yanjiu* 徽州方言研究 [Research on the Hui dialects]. Tokyo: Kobun Shuppen.
- Li, Charles N. & Thompson, Sandra A. (1989). *Mandarin Chinese: A functional reference grammar* (Vol. 3). Univ of California Press.
- Liu, Danqing 刘丹青. (2001). 语法化中的更新, 强化与叠加. *Yuyan yanjiu*, (2), 71-81.
- Long, Guofu, & Sun, Chaofen. (2013). Aspectual Marker: Issues of *zhe* of grammaticalization in the History of Chinese and Wu dialect. *Journal of Chinese Linguistics*, 41(2), 392-417.
- Mei, Zulin 梅祖麟. (1988). 汉语方言里虚词“著”字三种用法的来源. *Zhongguo yuyan xuebao*, 3.
- Meng, Qinghui 孟庆惠. (1997). *Anhui Shengzhi: Fangyan zhi* 安徽省志: 方言志 [Anhui Province Journal: Dialect Journal]. Fangzhi chubanshe.
- Jiang, Lansheng 江蓝生. (2008). 概念叠加与构式整合——肯定否定不对称的解释. *Zhongguo yuwen*, (6), 483-497.
- Zhang, Yisheng 张谊生. (2012). 试论叠加, 强化的方式, 类型与后果. *Zhongguo yuwen*, (2), 128-138.

Xuan Ye: ye.914@osu.edu

Yao Chen: chenyaofujianshida@163.com

Session II-A
Abstracts

Gender in Japanese Youth Language

Yuning Cao

Stanford University

The current study discusses the gender aspect of Japanese Youth Language (abbreviated as “YL”) and focuses on real-life data. Previous literature has investigated gender in YL from various aspects including the differences of personality and characteristic between genders (Tanimitsu, 2006), the vicissitudes of gender differences in YL (Yonegawa, 1998), particular words that exhibited gender differences (Yonegawa, 1998), and how male and female university students used YL differently as well as the reasons accounting for such differences (e.g., Akiyama, Uesugi, & Suzuki, 1991; Kim & So, 2014). At the meantime, YL usage can be considered as using a variant to construct certain persona or identities, the issue of which was extensively discussed in terms of phonological variants (e.g., Gratton, 2016; Podesva, 2011a), contours (Podesva, 2011b), and phonation types (Podesva, 2007), to name a few. The influence of phonological variants on listeners’ perceptions particularly shed light on the current study, including the examination of *-ing/-in*’ variants (Campbell-Kibler, 2007, 2008), released /t/ used by U.S. politicians (Podesva, Reynolds, Callier, & Baptiste, 2015), and fronted /s/ among Copenhagen youth (Pharao, Maegaard, Møller, & Kristiansen, 2014). The current study is inspired by these studies on perceptions and explores how YL usage is perceived from a hearer’s point of view, especially how speakers’ gender influences hearers’ perception on YL usage.

Deriving from a larger project, this presentation reports quantitative data collected from an online survey and omits discussion on other data sources. In total, 113 online surveys (male = 54, female = 59) asking a native speaker of Japanese to evaluate a speaker after listening to a short conversation were collected. Sampling service provided by Qualtrics was requested to ensure a population as unbiased as possible. The distribution of age is detailed in Table 1. There were 16 conversations (4 YL words × 2 YL/non-YL options × 2 genders), and each participant randomly listened to only one conversation. All conversations were recorded by a female undergraduate student and a male graduate student. After the conversation, participants were asked to rate their perception on the speaker’s age, femininity/masculinity, education level, and twenty more items on personalities (see list attached) using a 1-6 Likert scale. The experimental design of controlling speakers’ gender and usage/non-usage of YL words aims at eliminating other factors that might influence hearers’ perception. Sample conversations of one YL word can be found in Table 2.

Regarding whether listeners’ perception would differ based on speaker’s gender providing a YL utterance, two items appeared statistically significant—interestingness and intelligence. That is, a female user is perceived more intelligent (female = 3.50, male = 2.71, $p = .04$) and more interesting (female = 3.36, male = 2.52, $p = .01$) when she uses a YL. Intriguingly, analysis on non-YL utterances revealed a difference in leadership (female = 2.60, male = 3.31, $p = .01$), with no other significant differences. In other words, when both genders are conducting the same conversation using unmarked expressions, people tend to rate male speakers as possessing more leadership capability. Considering that a short conversation discussing casual topics can barely be related to leadership capability, we can interpret this result as a rating under the influence of gender stereotypes irrelevant of contexts provided by the conversation. Nevertheless, it is worth mentioning that leadership was the only category that instantiated such view, and that no other categories showed a difference.

Regarding listeners’ perception on speakers depending on usage/non-usage of YL, for female speakers the only difference is observed in competitiveness (YL = 3.57, non-YL = 2.84, $p = .04$).

This can be interpreted as female speakers using YL were considered more competitive. Equally interestingly, no other categories demonstrated a statistical difference, suggesting that in general, people’s perception on female speakers does not differ based on YL usage, echoing the point that women are considered more active in creating and using new expressions, and thus people are used to their YL usage. In contrast, for male speakers, YL usage tends to be more marked and is perceived more negatively. As summarized in Table 3, male YL users were considered childish (perceived age), less masculine, less handsome, less serious, less interesting, less intelligent, and with less leadership capability. Reversely, being rated negatively may attribute to why males use YL less frequently in real-life conversations, in return reinforcing the stereotype that YL words are reserved for female speakers.

The current research provides an updated view on how YL/non-YL users of different genders are perceived by the general public. As demographic information on rater’s gender and age were also collected, further analysis may explore how rater’s gender influence his/her perception on speakers of each gender, as well as on their YL/non-YL usage. As only two native speakers in their twenties recorded all conversations for this study, possible research direction also includes how the age of YL users affects people’s perceptions.

List of personalities: “popularity” “likelihood to be welcomed” “sociability” “capability of communication” “fondness of fashion” “degree of handsomeness/beauty” “seriousness” “interestingness” “intelligence” “loveliness” “determination” “self-disclosure” “leadership” “competitiveness” “activeness” “empathy” “positivity” “preemptive-ness” “outgoingness” “self-centeredness”

Table 1 Age distribution of questionnaire respondents

Age	18-22	23-25	26-29	30-34	35-39	40-44	45-49	50-54	55-60	>60
Count	5	10	12	25	17	21	7	6	8	2

Table 2 Sample conversations

Word	YL?	Gender	Japanese	English translation
草不可避 (kusafukahi)	YL	F	F:今日の番組、 <u>草不可避</u> ね。 M:そうそう！	F: Today’s TV show is so <u>interesting!</u> M: That’s right!
	non-YL	F	F:今日の番組、 <u>面白くて笑っちゃう</u> よね！ M:そうそう！	F: Today’s TV show is so <u>interesting!</u> M: That’s right!
	YL	M	M:今日の番組、 <u>草不可避</u> ね。 F:そうそう！	M: Today’s TV show is so <u>interesting!</u> F: That’s right!
	non-YL	M	M:今日の番組、 <u>面白くて笑っちゃう</u> よね！ F:そうそう！	M: Today’s TV show is so <u>interesting!</u> F: That’s right!

Table 3 Perception differences - Male

Item	Mean (using YL)	Mean (using non-YL)	P Value
perceived age	10.76	15.03	.040
perceived masculinity	19.76	27.00	.031
handsomeness	2.59	3.31	.036
seriousness	2.71	3.55	.042
interestingness	2.53	3.17	.033
intelligence	2.71	3.48	.013
leadership	2.41	3.31	.009

Selected References:

- Akiyama, Y., Uesugi, T., & Suzuki, M. (1991). Research on the phrase in fashion (1st Report) — On the student at Bunkyo University, Faculty of Human Science —. *Language and Culture*(4), 67-92. Retrieved from <https://ci.nii.ac.jp/naid/120006418503/>
- Campbell-Kibler, K. (2008). I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, 37(5), 637-659.
- Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, 82(1), 32-64.
- Gratton, C. (2016). Resisting the gender binary: The use of (ING) in the construction of non-binary transgender identities. *University of Pennsylvania Working Papers in Linguistics*, 22(2), 7.
- Kim, Y.K., & So, K.W. (2014). The features of the new generation through their unique words. *Journal of Japanese Language and Culture*, 27, 183-210. Retrieved from <https://www.earticle.net/Article/A281292>
- Kuwamoto, Y. (2003). Production and stability of *wakamono kotoba*. *Research Reports of Akita National College of Technology*(38), 113-120. Retrieved from <https://ci.nii.ac.jp/naid/110000477734/en/>
- Pharao, N., Maegaard, M., Møller, J. S., & Kristiansen, T. (2014). Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society*, 43(1), 1-31.
- Podesva, R. J., Reynolds, J., Callier, P., & Baptiste, J. (2015). Constraints on the social meaning of released /t/: A production and perception study of US politicians. *Language Variation and Change*, 27(1), 59-87.
- Podesva, R. J. (2011a). The California vowel shift and gay identity. *American Speech*, 86(1), 32-51.
- Podesva, R. J. (2011b). Salience and the Social Meaning of Declarative Contours: Three Case Studies of Gay Professionals. *Journal of English Linguistics*(3), 233. Retrieved from <https://stanford.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=edsfra&AN=edsfra.24469748>
- Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, 11(4), 478-504.
- Tanimitsu, T. (2006). Youth language and male language/female language (translated by the author). *The Bulletin of Musashino Gakuin University Japanese Research Institute*, 3, 50-57. Retrieved from <https://ci.nii.ac.jp/naid/110006161489/>
- Yonegawa, A. (1998). *Science of Youth Language*: Meiji Shoin.

Yuning Cao: yuning2@stanford.edu

The narrative functions of perfective auxiliaries in Early Heian *Kundokubun* texts

John Bundschuh

The Ohio State University

Kundokubun is a linguistic style of early spoken Japanese adapted for reading texts written in Sinitic script aloud as Japanese or quoting Sinitic texts in Japanese. The earliest extant texts clearly rendered in *kundokubun* via a system of morphosyntactic and phonetic glosses are found at the onset of the Heian period (794–1185CE). The focus of this study is two earlier Japanese perfective auxiliaries, or inflecting suffixes, the exoactive perfective *-tu* and endoactive perfective *-nu* (Quinn 1987). Scholars such as Kasuga Masaji (1942b), Tsukishima Hiroshi (1967), . tsubo Heiji (1992), and Kobayashi Yoshinori (2012) have investigated the semantics of these perfectives in Heian period *kundokubun* texts. They primarily find *-nu* governing intransitive and passive predicates and *-tu* governing transitive and causative predicates. While the narrative functions of these auxiliaries have been studied in Heian period vernacular tales (Suzuki 2009, Watase 2013), there has yet to be a study of their relation to narrative structure in *kundokubun*. The purpose of this study is to determine their narrative functions in early-Heian *kundokubun* texts.

Adhering primarily to the methodology of Nara (2011), this study codes finite predicates governed by *-nu* and *-tu* for transitivity (Hopper & Thompson 1980), lexical aspect (Kenny 1963, Vendler 1967), and discourse function (Labov 1972) to determine the auxiliaries' respective narrative functions. Data were drawn from the following Buddhist texts rendered into *kundokubun* via gloss during the first half of the Heian period: an early-9th-century CE complete rendering of the *Golden Light Sutra* (*Konkōmyōsaishōkyō*) (Kasuga 1942a), a mid-9th-century CE rendering of the 5th and 9th scrolls of the *Golden Light Sutra* (Tabuchi 1987), a late-9th-century CE rendering of eight scrolls of the *Kṣitigarbha Ten Wheels Sutra* (*Jizōjūrinkyō*) (Nakada 1958, Nakada 1980), and an early-10th-century CE complete rendering of the *Innumerable Meanings Sutra* (*Muryōgikyō*) (Kabutogi & Nakada 1979). The cited researchers have deciphered the original glossed texts and rendered them into *yomikudashibun*, a more accessible written form of *kundokubun*.

Examples (1)–(4) below display prototypical uses of *-nu* and *-tu* in these early *kundokubun* texts.

- (1) 右に三帛遶りて、退（き）て一面に**坐しヌ**。(Kasuga 1942a: 1)
Migi ni sansapu mapari-te, sirizoki-te itimen ni **za.si-nu**.
'Having circumambulated to the right three times, they withdrew and **sat down** in one corner.'
- (2) 是（の）如く法を護ル人は、諸法の一昧を**知りヌ**。(Nakada 1980: 187)
Ko no gotoku papu wo mamoru pito pa, syopapu no itimi wo **siri-nu**.
'One who thus protects the dharma **learns** the nature of all things.'
- (3) 時に諸（の）大衆、佛の此（の）甚深（の）空性（を）説（き）たまふ（を）**聞（き）ツ**。(Tabuchi 1987: 58)
Toki ni moro no daisyu, potoke no ko no zintan no kūsyau wo toki-tamapu **wo kiki-tu**.
'Then the entire gathering **heard** the Buddha explaining this deep emptiness.'
- (4) 一切の法に於て勇健の想を**得ツ**。(Kabutogi & Nakada 1979: 110)
Issai no papu no oite yūgon no sau wo **e-tu**.
'They gain robust thoughts regarding all dharma.'

In example (1) *-nu* governs an intransitive predicate of motion. Its use in example (2), while transitive to the extent the verb *siru* ('learn') takes an object marked by the accusative particle *wo*, is low on Hopper & Thompson's (1980) transitivity scale due to the lack of change in the object. Auxiliary *-nu* often governs predicates indicating changes of mental state, such as learning and thinking. Beyond verbs of motion and thought, auxiliary *-nu* also often governs *naru* ('become'). Auxiliary *-tu*, on the other hand, in addition to being found in more prototypical transitive predicates such as those including *nasu* ('make'), *usinapu* ('lose'), and the causative auxiliary *-simu*. Example (3) displays its governing *kiku* ('hear'), which parallels another common use governing *miru* ('see'). Finally, *-tu* is the auxiliary that most often governs *u* ('gain'), seen in example (4) in its participial form *e*. These examples both are representative of the transitivity differences between the two auxiliaries and demonstrate their prevalence in governing telic predicates (see Kenny 1963), reflecting their perfectivity. Among these four examples, only (3) depicts an accomplishment, a predicate that is both telic and durative, rather than an achievement, one that is telic and instantaneous. Both *-tu* and *-nu* predominantly govern achievements, which is because in *kundokubun* predicate morphology specific context is required to have a verb represent the latter. In other words, in *kundokubun* all accomplishments are predicates that, by themselves outside of context, index activities (durative and atelic) that gain an end point through context.

Regarding the discourse functions of *-tu* and *-nu*, in narrative sequences both are only found sentence-finally in complications, or events that progress the story, and resolutions, events that conclude the story. Example (5) below depicts the first action in the *Innumerable Meanings Sutra* after its orientation is a motion governed by *-nu*.

- (5) 各々眷属百千万数と与に、而も自(ら) 围绕せられて、仏の所に来詣しぬ。
 (Kabutogi & Nakada 1979: 78)
Onoono kenzoku pyakusenmanzyu to tomo ni, sikamo midukara wineu.serarete, potoke no tokoro ni raikēi.si-nu.
 'Each, being surrounded with countless followers, **approached** the place of the Buddha.'

Although scenes often end with characters in motion, and thus governed by *-nu*, like what we see in example (1), when the scene is primarily the telling of a parable or an explanation of the dharma, we occasionally find *-tu* in the resolution, such as what we see in example (6).

- (6) 余時世尊、已に大衆の為に、此の十千の天子の往昔の因縁を説(き)たまひつ。
 (Kasuga 1942a: 187)
 So no toki seson, sude ni daisyu no tame ni, ko no jissen no tensi no wauzyaku no innen wo **toki-tamapi-tu.**
 'Then the lord, already for the sake of the crowd, **explained** the origin of the history of these ten thousand divine sons.'

In narrative sequences among the examined texts, these auxiliaries are found in complications 98% of the time. Resolutions containing predicates of motion or speech are not the rule. Furthermore, *kundokubun* narration is most often framed with the modal auxiliary *-ki*, which indexes established facts. Thus, when perfectives are found in resolutions, they are most often themselves governed by *-ki*, as we see in example (7).

- (7) 妙幢菩薩は佛の足を礼したてまつり已（り）て、（從）座ヨリして（而）起（ち）て、其の本處に還（り）にキ。(Kasuga 1942a: 19)
Meudau bosatu pa potoke no asi o rei.si-tatematuri-wopari-te, za yori.site tati-te, sono ponsyo ni kaperi-ni-ki.
 ‘Kṣitigarbha finished paying obeisance to the Buddha’s feet, stood up from his seat, and returned to his original place.’

This preponderance of *-ki* marking on resolutions means the perfectives are most often found sentence-finally governing predicates in complications, or events that progress the story. As perfectives that index changes of state, the events depicted with *-tu* and *-nu* are both dynamic and central to the narrative progression. This study finds the narrative function of these perfectives in early Heian *kundokubun* narration to be that of foregrounding events that progress the stories within these texts.

References:

- Hopper, Paul and Sandra A. Thompson. 1980. “Transitivity in Grammar and Discourse,” *Language* 56-2: 251–299.
- Kasuga, Masaji. 1942a. *Saidaiji-bon Konkōmyō saishōkyō koten no kokugogakuteki kenkyū: Honbun-hen* (‘A Linguistic Study of the Old Diacritics of the Saidaiji Golden Light Sutra: Text Volume’). T. kyō: Benseisha.
- Kasuga, Masaji. 1942b. *Saidaiji-bon Konkōmyō saishōkyō koten no kokugogakuteki kenkyū: Kenkyū-hen* (‘A Linguistic Study of the Old Diacritics of the Saidaiji Golden Light Sutra: Research Volume’). Tōkyō: Benseisha.
- Kenny, Anthony. 1963. *Action, Emotion, and Will*. London: Routledge and Kegan Paul.
- Kabutogi, Shōkō & Norio Nakada. 1979. *Kotenbon shiryō sōkan 1: Muryōgikyō koten*. Tōkyō: Benseisha.
- Kobayashi, Yoshinori. 2012. *Heian jidai no bussho ni motozuku kanbun kundokushi no kenkyū 3: Shoki kundokugo taikai* (‘Research on the History of Kanbun Kundoku Based on Heian Period Buddhist Texts 3: The Early-period Kundokugo System’). Tōkyō: Kyūko Shoin.
- Labov, William. 1972. *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.
- Nara, Hiroshi. 2011. “Aspect and Discourse in Tense-Switching: A Case Study of Natsume Sōseki’s *Botchan*.” *Japanese Language and Literature* 45: 273–305.
- Nakada, Norio. 1958. *Kotenbon no kokugogakuteki kenkū: Yakubun-hen* (‘A Linguistic Study of Old Glossed Texts: Translation Volume’) Tōkyō: Kōdansha.
- Nakada, Norio. 1980. *Kotenbon shiryō sōkan 2: Shōsōinbon Jizō jūrinkyō kango, shichi Gangyō ten* (‘. Tōkyō: Benseisha.
- Ōtsubo, Heiji. 1992 *Ishiyamadera-bon Daihō kōbutsu kegon kyō koten no kokugogakuteki kenkyū* (‘A Linguistic Study of the Old Diacritics of the Ishiyamadera Flower Ornament Sutra’). Tōkyō: Kazama Shobō.
- Quinn, Charles. 1987. *A Functional Grammar of Predication in Classical Japanese*. Dissertation, University of Michigan.
- Tabuchi Masao. 1987. “Tōdaiji toshokanbon Konkōmyō saishōkyō chūshaku no kunten (‘The Glosses of the Tōdaiji Library Annotated *Golden Light Sutra*’).” *Kuntengo to kunten shiryō* 76: 53–90.
- Tsukishima, Hiroshi. 1967. *Kōfukiji-bon Daijionji Sanzōhōshiden koten no kokugakuteki kenkyū: Kenkyū hen* (‘A Linguistic Study of the Old Diacritics of the Kofukiji Daci’en Temple Biography of Xuanzang: Research Volume’). Tōkyō: Tōkyō Daigaku Shuppankai.
- Watase, Shiger. 2013. *Ōchō jodōshi kinō ron: Anatanaru ba, waku kōzō, enkinhō* (‘The Functions of Heian Period Auxiliaries: Removed Ground, Framing Construction, Distancing’). Tōkyō: Izumi Shoin.
- Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

John Bundschuh: bundschuh.8@osu.edu

Loanword adaptation in Japanese Kansai dialect

Yuki Hattori

The Ohio State University

Purpose:

Like many world languages, Japanese has many loanwords borrowed from different languages: Chinese, English, French, German, Portuguese, and so forth. These foreign loanwords are adopted using native phonological rules which determine pronunciation. Much like native words, Japanese loanwords are shortened or clipped using truncation styles. Regional dialects cause variations in these loanword abbreviations. The most distinctive regional dialect in Japan is Kansai-ben, also known as the Western dialect. The purpose of this study is to observe and determine if Japanese native speakers from the Kansai region apply their dialect's unique features when adopting loanwords. Japanese words have high and low pitches. Lexical accent is where pitch falls from high to low. Ramsey (1979) mentions how a low pitch in the Kansai dialect is more distinctive compared to Japanese standard speech. He also asserts pitch accentual patterns in Kansai dialect have a pitch fall in the middle of the word. The Kansai dialect is likely to place a pitch accent on the penultimate mora for 3-mora clippings (Kubozono 2008; Fudano 2019, Table 1). The antepenultimate mora is usually accented in 4-mora words in both standard and the Kansai dialects (Cutler and Otake 1998; Fukumori 2020); however, just like 3-mora words, the low pitch on the first syllable is more distinctive in the Kansai dialect (Ramsey 1979:158, Table 2). Therefore, it can be hypothesized that native Japanese speakers from the Kansai area should know how to apply their dialect's features to new loanwords.

Table 1: Accent Pitch in Kansai and Tokyo-speech: 3-mora words (Kubozono, 2012; bold type indicates the location of accent)

	Television	Hawaii
Kansai	tere bi	hawa i
Tokyo	tere bi	h awai

Table 2: Accent Pitch in Kansai and Tokyo-speech: 4-mora words (Fukumori, 2020)

	Somen (white noodle)	Funny (slang)
Kansai	soome N	o moroi
Tokyo	s oomeN	omoro i

Design:

A speech production test was administered to 64 Japanese college students in the Kansai area. Although they were required to take basic English courses as a part of their university's curriculum, none of them majored in English. The participants' task was to shorten 20 English words that have not yet been adopted in Japanese. All English words on the list were written in Japanese based on Japanese phonological (transliteration) rules and sounds applied. All participants' performances were recorded and analyzed.

Results:

The data were examined in regard to two differing aspects: the number of morae and accentual patterns. As for the number of morae in clipped loanwords, no strong preference for 3-mora or 4-

mora clippings were found (45.31% for 3-mora words vs 48.37% for 4-mora words). However, when looking at the original loanwords' length, a clear tendency was observed: The current research discovered a tendency regarding how the Kansai dialect treated loanwords of a particular length. The original loanwords with applied Japanese phonological rules were divided into two categories: <7-mora and >=7words. Forty-nine out of 64 participants (76.6%) made 3-mora clippings consistently (over 68%) for the loanwords with 5 or 6 morae. Yet only 23 out of 64 participants (35.9%) made 4-mora clippings (at least over 68% of the time) for the loanwords with 7 or more morae (Table 3). Regarding the accentual patterns, the productions were also divided into 3-mora and 4-mora clippings. In 3-mora clippings, LHL (penultimate) patterns were observed in 19 out of 20 words (Table 4). However, when looking at each word's LHL consistency, no clear tendency was detected. Four-mora clipping data evoked both antepenultimate and penultimate accentual patterns (Table 5), and the participants did not show a strong preference to either of them. Overall, the data revealed that Kansai speakers might not be conscientious of accentual patterns for both 3-mora and 4-mora clippings.

Table 3: Preferences based on the full-length word

Original Words	Japanese phonological features applied	3-mora (%)	4-mora (%)
<7 morae			
Apothecary	[aposekari:]	73.44	23.44
Eclipse	[ekurɪpʉsɯ]	65.63	32.81
Kidnap	[krɔnap:ɯ]	90.63	6.25
Nobleman	[no:bɯrɯmaN]	6.25	90.63
Pavement	[pɛɪbɯmɛNto]	62.50	26.56
Witness	[wit:ɔnesɯ]	89.06	10.94
Infection	[ɪNɸɛkɯ{ɔN]	53.13	46.88
Landlord	[raNdoro:do]	46.88	50.00
Yashmak	[jaʃumak:ɯ]	68.75	9.38
		49 out of 64 people (76.6%)	
>=7 morae			
Heterogeneous	[heteroɔɹɪmasɯ]	67.19	26.15
Manipulation	[manɪpjɯre:ʃɔN]	60.94	35.94
Catastrophe	[katastoroɸi:]	9.38	85.94
fairground	[ɸɛagɯrauNdo]	17.19	73.44
Registration	[rɛɔɹɪsɯtoɹe:ʃɔN]	35.94	59.38
Recommendation	[rɪkɔmɛNde:ʃɔN]	26.56	66.15
Widespread	[wɪɔɹɔsɯpɯɹɛd:ɔ]	28.13	65.63
Orthography	[o:sogɯɹaɸi:]	37.50	46.88
Dermatologist	[da:matoroɔɹɪsɯto]	32.81	51.56
Governorship	[gabana:ʃɪp:ɯ]	23.44	73.44
Zoologist	[zɯ:oroɔɹɪsɯto]	10.94	85.94
			23 out of 64 people (35.9%)

Table 4: Penultimate accentual pattern in 3-mora clippings

	Penultimate (%)
Apothecary	38.30
Catastrophe	0
Dermatologist	9.52
Eclipse	37.21
Fairground	27.27
Governorship	20.00
Heterogenous	23.26
Infection	8.82
Kidnap	20.69
Landlord	8.70
Registration	26.09
Manipulation	30.77
Nobleman	25.00
Orthography	12.50
Pavement	17.50
Recommendation	23.53
Witness	1.75
Widespread	5.56
Yashmak	25.00
Zoologist	14.29

Table 5: Accentual patterns in 4-mora clippings

	Antepenultimate (%)	Penultimate (%)
Apothecary	26.67	28.57
Catastrophe	10.91	1.81
Dermatologist	9.52	3.03
Eclipse	9.52	9.52
Fairground	0	0
Governorship	8.51	2.13
Heterogenous	15.38	46.15
Infection	3.33	3.33
Kidnap	25	0
Landlord	6.25	6.25
Registration	7.89	0
Manipulation	4.35	0
Nobleman	3.45	3.45
Orthography	3.33	3.33
Pavement	0	13.33
Recommendation	11.63	13.95
Witness	0	0
Widespread	0	0
Yashmak	16.67	16.67
Zoologist	14.29	40

Conclusion:

Kansai speakers are likely to clip shorter English loanwords into 3 morae. However, 4-mora clippings were also observed, though the patterns were inconsistent. As for pitch accent patterns, no consistent pitch accent patterns were seen on both 3-mora and 4-mora clippings. This study revealed how in the modern society's rapidly changing vocabulary, the Kansai speakers retain the accent patterns of their traditional dialect. The standardization of Japanese dialects might be a reason for the lack of consistent trends from this study's results (Sanada 2018). In particular, the "standard" Japanese used in younger generations are heavily influenced by the Internet and digital multimedia, leading to an inconsistent use of dialectal phonological rules.

Selected References:

- Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Acoustical Society of America*, 105(3), 1877-1888.
- Fudano (2019)- Okayama University of Science professor, specializing in dialectology; these were personal communications that took place at Kansai Foreign Language University.
- Fukumori, T. (2020). The List of Accent in Idiomatic Osaka Dialect, *The Hokkaido Linguistic Circle*, 18, 183-196
- Kubozonzo, H. (2008). Japanese accent. In Shigeru Miyagawa and Mamoru Saito (eds.), *The \ Oxford handbook of Japanese Linguistics*, 165-191. Oxford: Oxford University Press.
- Ramsey, S. (1979). The Old Kyoto Dialect and The Historical Development of Japanese Accent. *Harvard Journal of Asiatic Studies*, 39(1), 157-175. doi:10.2307/2718816
- Sugito, M. (2000). CD-ROM Accent Dictionary of Spoken Osaka and Tokyo Japanese. *Journal of the Phonetic Society of Japan*, 4(2), 36-43.

Yuki Hattori: hattori.12@osu.edu

Session II-B
Abstracts

A usage-based construction approach to Korean DO-causatives

Eunhye Kim Hess

Oklahoma State University

Cross-linguistically, DO-causative is a type of syntactic or periphrastic causative, which involves a DO-verb (Moreno, 1993). In Korean, DO-causative involves the combination of a lexical item, either *key* or *tolok*, with the DO-verb *ha*. (See (1) and (2)). Previous studies on Korean causatives have described the differences between the two forms from the traditional structural views. The known differences are that *key* occurs more frequently than *tolok*; the directness of causation is higher with *key* than with *tolok*; the causee NP is marked with the accusative case particle (*ul/lul*) with *key* while it is the nominative particle (*i/ka*) with *tolok* (Seo, 1987; Song, 2015; Yeon & Brown, 2011).

- (1) Mina-ka Yumi-lul mul-ul masi-key ha-es-ta.
Mina-NOM Yumi-ACC water-ACC drink-CAUS-PST-DEC.
'Mina had Yumi drink water.'
- (2) Mina-ka Yumi-ka mul-ul masi-tolok ha-es-ta.
Mina-NOM Yumi-NOM water-ACC drink-CAUS-PST-DEC.
'Mina had Yumi drink water.'

However, in naturally occurring language, the suggested syntactic and semantic features of each construction seem to be inconsistent. Furthermore, previous studies of Korean syntactic causatives were largely based on sentence-level syntactic rules and researcher-generated sentences without considering the discourse and contexts. However, since Korean is a discourse-prominent language (Kim 1997; Clancy 1996), the indisputable relation of a syntactic form and its meaning must be captured through discourse (Strauss, Lee, & Ahn, 2006).

This study aims to examine different syntactic and semantic features of the Korean DO-causatives within the usage-based construction grammar framework (Bybee, 2006; Goldberg, 1995). The usage-based linguistics (Barlow & Kemmer 2000; Bybee and Hopper 2001) argues that the knowledge of a language is informed by the actual use of language and posits that utterances form the basis for constructions. Each construction, as a symbolic unit of form and meaning, carries its own meaning. Within this approach, the two Korean syntactic DO-causatives are considered as individual constructions, *-key HA* and *-tolok HA*.

This study focuses on exploring the prototypical form and meaning of each construction by using naturally occurring language a written corpus. The Sejong Corpus, provided by the National Institute of Korean Language, is used for the data collection. The Sejong Corpus, also known as the Korean National Corpus, is a general corpus, offering both contemporary written and oral Korean corpus data (Kim, 2006). Total 514,107¹*ecel* was collected from the written corpus, which consists of newspapers and books from the year 2002 to 2003. The data was analyzed using a concordancing tool, *Hanmaru*.

A total of 224 tokens of the *-key HA* construction and 70 tokens of the *-tolok HA* construction were found from the corpus. As for the prototypical form of each construction, the causer and causee NPs along with causee NP's case marking were analyzed. The results showed that the prototype of each construction does not have explicit causer or causee NPs in the same clause. Furthermore, the causee NP's case marking in each construction varied. This result

¹ *ecel* is similar to word unit but space-based.

suggests that the traditional approach based on sentences with fully realized causer and causee NPs do not reflect the real usages of the constructions and the pro-drop feature of Korean. When it comes to the meanings, verbs co-occurring with each construction were categorized, and further discourse analysis was conducted. The *-key HA* construction frequently co-occurred with verbs of action and states, while the *-tolok HA* construction frequently co-occurred with verbs of action. Furthermore, the discourse analysis showed that each construction delivers not only the causative meaning but also related senses, such as more purposive meaning for the *-tolok HA* construction and more resultative meaning for the *-key HA* construction. Thus, the different meanings of the two constructions can also be understood through the related senses, not just through the directedness of causation.

Example (3) from the corpus illustrates these findings. Previously in the discourse, reparations made by war crime countries were discussed. The ongoing topic was Germany, and their reparation efforts were discussed for victim countries. After reviewing Germany in the previous discourse context, in this example, the discourse topic now switches to Japan. Here, Japan is also the causee, which is morphologically marked with a topic particle *un*, occurring at the beginning of the sentence. Thus, we observe the topicality effect, which not only results in the scrambled order of causer and causee NPs but also no use of normative nor accusative particle for the causee's case marking. Further, even though *key* is used with the causative event, 'to maintain the emperor system,' the directedness of causation is not clear to be direct. Instead, the conveyed meaning is better understood as the final resultative state meaning.

- (3) panmyen, ilpon-un yenhapkwun-i ku-tul-uy chenhwang-cheycey-lul
 whereas Japan-TOP Allied Force-NOM that-PUL-GEN emperor-system-ACC
 kutaylo yuciha-key ha-y kwake-lul chengsanha-l swu eps-nun kil-ul
 as it is maintain-CAUS-as past-ACC settle-FUT case CP:not exist-REL-path-ACC
 kel-e wa-ss-ta.
 walk-CONN come-PST-DEC.
 'On the other hand, as for Japan, because the Allied Forces had/made (Japan) maintain
 their system of emperorship, (Japan) has walked toward the path that cannot settle the
 past.'

These findings reveal that the previous traditional formal approach to the DO-causatives in Korean does not adequately describe their usages in naturally occurring language. This study argues that the application of the usage-based construction grammar approach would provide an insightful theoretical framework and analysis for a discourse-prominent language, Korean.

References

- Barlow, M., & Kemmer, S. (2000). *Usage-based models of language*. Stanford: Center for the Study of Language and Information.
- Bybee, J.L., & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamin Publishing Company.
- Bybee, J.L. (2006). From usage to grammar: the mind's response to repetition. *Language*, 82, 711-733.
- Clancy, P. (1996). Referential strategies and the co-construction of argument structure in Korean acquisition. In Fox, B. (Ed.), *Studies in anaphora* (pp. 33-68). Amsterdam: John Benjamin Publishing Company.
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago and London: University of Chicago Press.
- Kim, Y. (1997). The acquisition of Korean. In Solbin, D. (Ed.), *The crosslinguistic study of language acquisition* (pp. 335-444). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

- Kim, H. (2006). Korean national corpus in the 21st century Sejong project. In *Proceedings of the 13th NIJL International Symposium* (pp. 49-54). National Institute for Japanese Language Tokyo.
- Moreno, J. C. (1993). "Make" and the semantic origins of causativity: a typological study. In Comrie, B., and Polinsky, M. (Eds.), *Causatives and transitivity* (pp. 155-164). Amsterdam/Philadelphia: John Benjamin Publishing Company.
- Seo, J. (1987). On the verb endings keh and torok. *The Journal of Humanities*. 14, 41-67.
- Song, J. (2015). Causatives. In Brown, L. & Yeon, J. (Eds.), *The Handbook of Korean Linguistics* (pp. 98-115). The Atrium: John Wiley & Sons, Inc.
- Strauss, S., Lee, J. & Ahn, K. (2006). Applying conceptual grammar to advanced-level language teaching: the case of two completive constructions in Korean. *The Modern Language Journal*, 90(2), 185-209.
- Yeon, J., & Brown, L. (2011). *Korean: A comprehensive grammar*. New York, NY: Routledge.

Eunhye Kim Hess: eunhye.kim@okstate.edu

Korean writing system: Adaptation and creation

Seojin Yang

The Ohio State University

The Korean writing system, Hangeul, is a delicately devised system for representing syllable system and sounds of the Korean spoken language. One thing special about Hangeul is that this writing system was created by King Sejong in 1443. Then how did Korean people record their colloquial language before the creation of Hangeul? What system or tool was used to record the sounds of Korean before the creation of Hangeul?

The answer: Before the creation of Hangeul, Chinese characters were used to meet the written needs of the Korean people. This study focuses on two questions: One, what are the various creative ways that the Koreans use to ‘adapt’ foreign letters to represent their native sounds? And two, what are the strategies that the Koreans use to ‘create’ the numerous new Sinitic graphs on their own? In addition, this study also aims to introduce how Hangeul is used creatively to transcribe foreign sounds in modern times.

Over 2,000 years ago, during the Three Kingdom Period, Chinese characters were borrowed and used in various ways mixed with the native Korean language. More specifically, to represent the sounds, meanings, and grammar of colloquial Korean, whole or parts of Chinese characters were ‘adapted’ in many forms, such as *hyangchal* (鄉札), *idu* (吏讀), and *gugyeol* (口訣),¹ This is illustrated as follows.

(1) a. *hyangchal* (鄉札)

	去	隱	春		去隱春
訓 :	go	hide	spring	=>	Passed spring
音 :	geo	eun	chun	=>	gan bom
Native Korean :	ga (da)	sum (da)	bom	}	

b. *idu* (吏讀)

	告	目		告目
訓 :	say	eye	≠>	letters to the noble
音 :	go	mok	=>	go mok
Native Korean :	go(hada)	nun		

c. *gugyeol* (口訣)

	學而時習之不亦說乎 ²	Isn't it a pleasant to learn and often practice?
口訣 :	學而時習之 <u>面</u> 不亦說乎 <u>牙</u>	
	면 myun	아 a
	conditional clause	sentence-closing ending

¹*hyangchal* (鄉札): A typical method of recording 鄉歌 (Hyang ga), poetic songs of Shilla, during the Shilla period (BC 57 – AD 935). Sounds/meanings of Chinese characters were borrowed and used to transcribe 鄉歌. *idu* (吏讀): A generic term for using Chinese characters in various ways to represent Spoken Korean. The method of *idu* had been broadly in use from the Three kingdom period (57 BC) to early 20th century.

gugyeol (口訣): The way of reading Chinese sentences out by adding grammatical particles of Korean spoken language.

² <論語> 學而篇

In (1a), ‘去隱春’ has the meaning of ‘passed spring’ (go + spring) and the pronunciation of ‘gan bom’ (ga + n + bom). Likewise, (1b) shows how 告目 (gomok) is formed by combining the sounds of 告 (go) and 目 (mok) to represent a whole new meaning in *idu*. In (1c), the conditional clause 면 (myun) and a sentence-closing ending 아 (a) are added to the Chinese sentence to make it easier to read aloud in spoken Korean.

In addition, new characters called *guk ja* 國字, which refers to special characters that are only used in Korea, were ‘created’ to supplement the Chinese script to represent the Korean language.

(2) *guk ja* (國字)

	石	乙		𠬪
訓 :	stone	new	=>	stone
音 :	suk	eu <u>l</u>	=>	<u>dol</u>
Native Korean :	<u>dol</u>	sae	}	

A *gukja* character 𠬪 has a meaning of ‘stone’ from the meaning of 石 and a pronunciation of ‘dol’ from 石 (dol) and 乙 (l ending). As provided in (1) and (2), this study aims at exploring various cases of ‘adaptation’ and ‘creation’ of characters with more examples and in-depth analysis.

The creation of the Hangeul system made it possible for Korean people to represent their spoken language almost perfectly without help of Chinese characters except for the cases of fine distinctions of homophones. Then, a new question arises as to how foreign sounds are represented with Hangeul in modern Korean. Are there as many ways to represent new foreign sounds today as there were with Chinese characters in the past? Are there any new forms of ‘creation’ in representing foreign sounds?

Hangeul is an efficient system of phonograms in that Hangeul can potentially transcribe any existing sounds of the world, within the limits of the phonological system of Korean. Therefore, the pronunciation of many loanwords are simply and easily represented by being transcribing them using Hangeul. However, some of foreign sounds cannot be fully represented since there may be no matching sounds in the inventory of Korean sounds. For example, the consonant inventory of modern Korean does not include such sounds as /r/, /v/, and /θ/. There is no alternative but to represent these sounds with similar-sounding Korean consonants, such as the use of ‘ㄹ’, ‘ㅃ’, and ‘ㅅ’. As in (3a), /r/ is transcribed using ‘ㄹ’, since there is no matching sound in Korean sounds inventory. Likewise, /v/ and /θ/ are transcribed using ‘ㅃ’ and ‘ㅅ’, which usually represent /b/ and /s/ respectively.

(3) Representation of /r/, /v/, /θ/

a. /r/	red [red]	led [led]	레드	/r/, /l/	-> ㄹ
b. /v/	vine [van]	bine [báin]	바인	/v/, /b/	-> ㅃ
c. /θ/	thin [θm]	sin [sin]	씬	/θ/, /s/	-> ㅅ

Recently, a new and creative way of representing foreign sounds has been broadly used in internet language, text messages, and TV shows. People combine foreign letters and Korean letters to represent foreign sounds which originally cannot be fully represented with Hangeul. This creative method usually involves combining a consonant of foreign language and a Korean

vowel to form a Korean syllable. For instance, the English consonants ‘r’, ‘v’ and ‘th’ are mixed directly with Korean vowels, as shown in (4).

(4) New Representation of /r/, /v/, /θ/

a. /r/	red [red]	R 드	ㄹ -> R
b. /v/	vine [vam]	V 인	ㅂ -> V
c. /θ/	thin [θm]	thin	ㅅ -> Th

In (4a-c), the Korean consonants ‘ㄹ’, ‘ㅂ’ and ‘ㅅ’ are replaced by ‘R’, ‘V’ and ‘Th’ respectively to represent the more accurate pronunciation of English words ‘red’, ‘vine’ and ‘thin’. Through this creative combination of English consonants and Korean vowels people are able to overcome the limits of representing foreign sounds with using Hangeul alone. How Korean speakers creatively represent foreign sounds by combining letters from foreign languages with Hangeul is discussed with more details in the presentation, together with some exceptional cases.

To summarize, the Korean writing system has been developed using various creative ways to represent the colloquial Korean language. Before the creation of Hangeul, Chinese characters were adapted to represent the sounds of Korean. Also, new Sinographs were invented to record the sounds and meanings of the Korean language. Later, the creation of Hangeul made it possible to record a greater range of sounds more simply and easily. However, but some foreign sounds could still not be represented with Hangeul. Today, Korean speakers have begun to use new and creative ways to transcribe foreign sounds using Hangeul. Thus, this study provides an overview, with examples, of how the Korean writing system has been developed with ‘adaptation’ of foreign characters and sounds, and the ‘creation’ of new Sinographs with various examples.

References

Encyclopedia of Korean Culture 한국민족문화대백과사전 (The Academy of Korean Studies 한국학중앙연구원, <http://encykorea.aks.ac.kr>)

Kim, Jonghun 김중훈. 1972. 고유한자 (固有漢字 (國字)) 고 (攷)(一) A Research of Go you han ja. 국어국문학 *Korean Language and Literature*, 55, 175-188.

Lee, Gunsick 이건식. 2009. 한국 (韓國) 고유한자 (固有漢字)의 발달 (發達)-지명 (地名)의 후부 요소 표기 (表記)를 중심으로 Development of Korean Go you han ja – in Transcription of Place Name. 구결연구 *Gu gyeol Studies*, 22, 219-29.

Lee, Gunsick 이건식. 2012. 한국 고유한자 자형 (字形) 구성 (構成) 방법 (方法) 연구 二題-기존 연구의 비판적 검토와 형성 (形聲)으로 만들어진 한국 고유한자의 몇 가지 사례. A Study of Composition Methods of Korean Go you han ja. 동양학 *Oriental Studies*, 52, 187-211.

Lee, Gunsick 이건식. 2015. 통합 한한대사전 (漢韓大辭典)의 국자 (國字) 처리에 대하여 On the Treatments of Korean-originated Chinese Characters in the Unified Great Chinese-Korean Dictionary. 동양학 *Oriental Studies*, 59, 143-167.

Nam, Poonghyun 남풍현. 1981. 차자 표기법 연구 A Research of Adaptation of Chinese Characters. 단국대 출판부 Danguk University Press.

Seojin Yang: yang.3170@osu.edu



GACL

漢語語言學研究生聯合會

Graduate Association of Chinese Linguistics

<http://u.osu.edu/gacl/>