



A phonetic and multimodal exploration of the BA construction of the Chinese imperative

--- A case study of making dumplings in dyadic interaction

Guohua Hu

Supervisors:

Prof. Dr. Marianne Gullberg

Senior lecturer. Dr. Mechtild Tronnier

Centre for Languages and Literature

Master's Thesis, 30 ECTS (SPVR01)

General Linguistics

December 2020

Abstract

Previous studies of the Chinese imperative have treated syntax and prosody separately, have not considered gestures, and have looked at sentences in isolation and not in discourse. This study therefore investigated how the relationship between grammatical components, intonation (perceived stress), and gesture are co-organized and interact in the 把 *bǎ* ‘Ba’ construction (i.e., BA) of Chinese imperatives, specifically in natural spontaneous Chinese discourse in a face-to-face task-driven dyadic activity. The results reveal that the verb phrase (VP) bears the perceived sentence stress of BA and gestural strokes also mostly align with the VP. These results suggest that there is, in fact, an interaction between the three modalities: syntax, prosody, and gesture. Future studies might investigate other gestural articulators, for example head movements, and BA outside of the imperative.

Keywords: Chinese imperative, 把 *bǎ* ‘Ba’ construction, syntax, prosody, gesture, multimodality

TABLE OF CONTENTS

ABSTRACT.....	2
LIST OF EXCERPTS	5
LIST OF TABLES.....	6
LIST OF FIGURES.....	7
TRANSCRIPTION CONVENTIONS.....	9
LIST OF GLOSSING CONVENTIONS.....	10
ACKNOWLEDGEMENT.....	11
1. INTRODUCTION.....	12
2. THEORETICAL BACKGROUND.....	13
2.1. CHINESE IMPERATIVES AND THE BA IN PREVIOUS STUDIES.....	13
2.2. THE PROSODY OF CHINESE IMPERATIVES AND THE BA.....	18
2.3. THE GESTURES OF CHINESE IMPERATIVES AND THE BA.....	22
2.4. SUMMARY	25
3. THE CURRENT STUDY	25
4. METHODS	26
4.1. PARTICIPANTS	26
4.2. MATERIAL AND THE DESIGN OF THE TASKS	28
4.2.1. <i>Materials for the production task</i>	28
4.2.2. <i>The production task design</i>	28
4.2.3. <i>The perception task design</i>	29
4.3. PROCEDURES	29
4.3.1. <i>The production task</i>	29
4.3.2. <i>The perception task</i>	32
4.4. DATA TREATMENT AND CODING.....	33
4.4.1. <i>Production data: speech</i>	33
4.4.1.1. <i>Transcription</i>	33

4.4.1.2.	<i>Further linguistic coding of BA</i>	34
4.4.1.3.	<i>Gesture annotation</i>	34
4.4.1.4.	<i>Intonation annotation</i>	35
4.4.1.5.	<i>Alignment</i>	37
4.4.2.	<i>Speech perception treatment</i>	38
4.5.	ANALYSES	39
5.	RESULTS	39
5.1.	GESTURAL STROKES ALIGNED WITH BA	40
5.2.	PERCEIVED STRESS WITH BA	45
5.3.	THE ALIGNMENT OF GESTURAL STROKES AND PERCEIVED STRESS IN BA	51
6.	GENERAL DISCUSSION, CONCLUSIONS, AND FUTURE STUDY	55
	BIBLIOGRAPHY	60
	APPENDIX I CONSENT FORM	64
	APPENDIX II ANSWER SHEET	65
	APPENDIX III	66
	APPENDIX IV	69

List of Excerpts

Excerpt 1 Group 1, Male instructor, line 003-004, 00:04.31-00:07.55	33
Excerpt 2 Group 6, Female, line 039, 02:20.78-02:31.91. The arrow in line 039 indicates the gesture illustrated in Figure 9.	41
Excerpt 3 Group 2, Female, line 009, 00:10.55-00:16.06. The arrow in line 009 indicates the gesture illustrated in Figure 10.	43
Excerpt 4 Group 5, Female, line 098, 04:44.71-04:50.23. The arrow in line 098 indicates the gesture illustrated in Figure 11.	44
Excerpt 5 Group 3, Female, line 104, 05:47.69-05:49.55. The arrow in line 104 indicates the example illustrated in Figure 12.	47
Excerpt 6 Group 7, line 036, 01:46.20-01:49.89. The arrow in line 036 indicates the example illustrated in Figure 13.	48
Excerpt 7 Group 8, line 051, 01:52.98-01:57.32. The arrow in line 051 indicates the example illustrated in Figure 14.	49
Excerpt 8 Group 5, line 016, 00:48.25-00:51.54. The arrow in line 016 indicates the example illustrated in Figure 15.	50
Excerpt 9 Group 2, Female, line 054, 04:15.42-04:21.90. The arrow in line 052 indicates the example illustrated in Figure 16.	52
Excerpt 10 Group 4, Female, line 069, 03:11.85-03:13.70. The arrow in line 069 indicates the example illustrated in Figure 17.	54

List of Tables

Table 1 The demographic information of the participants in the production study..	27
Table 2 Main ingredients, stuffing, and spices	28
Table 3 Details of recording for analysis	32
Table 4 The distribution of stress of Excerpt 1, Group1, line 003.....	39
Table 5 Distribution of gestural strokes over grammatical elements in the BA- construction	40
Table 6 Distribution of perceived sentence stress over grammatical elements in the BA-construction (BaP = Ba-Phrase, Ba = 把 <i>bǎ</i> ‘Ba’, Ba-O = Ba-Object, Ba- O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, H = high, LH = rising, L = dip, HL = falling)	46
Table 7 Distribution of perceived sentence stress aligned with lexical tones in the Verbs (H = high, LH = rising, L = dip, HL = falling; H* = local prominent high, L* = local prominent low)	46
Table 8 Distribution of strokes aligning with perceived sentence stress within BA- construction (BaP = Ba-Phrase, Ba = 把 <i>bǎ</i> ‘Ba’, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, H = high, LH = rising, L = dip, HL = falling).....	52

List of Figures

Figure 1 An example of gestural stroke preceding the stressed syllable (Kendon, 2004:114). ~ = gestural preparation; * = stroke; -. = recovery; ... = pause; SMALL CAPS = stress.	24
Figure 2 An example of gestural stroke aligned with the stressed syllable (Kendon, 2004:117). ~ = gestural preparation; * = stroke; -. = recovery; ... = pause; SMALL CAPS = stress.	24
Figure 3 Schematic representation of the experimental set-up	31
Figure 4 Sample of hand gestures from ELAN, produced by male instructor from Group 1, in line 003 to 004	35
Figure 5 F ₀ contour of line 003 to 004 in Excerpt 1 displayed by Praat picture.....	36
Figure 6 The utterance produced by the male instructor from Group 1 in line 003 to 004 in Excerpt 1	37
Figure 7 Gestural stroke aligns with the grammatical component perceived as the stressed syllable, male instructor, Group 1, in line 003 to 004 in Excerpt 1	38
Figure 8 Gestural stroke alignment (***) with the entire BA-construction of line 039 in Excerpt 1, transcription (top) and export filmstrip (bottom), respectively from ELAN.	42
Figure 9 Gestural stroke alignment (***) with only BaP within the BA-construction of line 009 in Excerpt 2, transcription (top) and export filmstrip (bottom), respectively from ELAN.....	43
Figure 10 Gestural stroke alignment (***) with only VP within BA-construction of line 098 in Excerpt 3, transcription (top) and export filmstrip (bottom), respectively from ELAN.....	45
Figure 11 F ₀ contour of line 104 in Excerpt 5	47
Figure 12 F ₀ contour of line 036 in Excerpt 2	48
Figure 13 F ₀ contour of line 051 in Excerpt 3	49
Figure 14 F ₀ contour of line 016 in Excerpt 4	50

Figure 15 Gestural stroke (***) aligned with Ba-O as perceived sentence stress with tone HL of line 052 in Excerpt 8, transcription (top) from Praat and export filmstrip (bottom) from ELAN, respectively.	53
Figure 16 Gestural stroke (***) aligned with verb within VP as perceived sentence stress with HL of line 069	54

Transcription conventions

The transcription system used for discourse in this study is adapted from Ochs (1979) and Atkinson and Heritage (1999) with modifications.

- (0.2) Pauses of 0.2 second
- [Overlap
- Specific line in the transcript which is referred to in the text
- (noise) Paralinguistic and non-linguistic actions

The transcription system used for prosody in Praat can be found in Pierrehumbert and Beckman (1988) with modification.

- <SIL> Pause
- H* High pitch accent
- L* Low pitch accent

The transcription system used for hand movements in ELAN can be found in Kendon (2004)

- ↵ Preparation of gesture
- * Gestural stroke
- . Recovery of gesture

List of Glossing Conventions

The glossing of transcription basically falls in the Leipzig Glossing Rules (Comrie et al., 2008) with modifications.

- 1PL First person plural
- 2SG Second person singular
- BA The marker 把 *bǎ* ‘Ba’
- GEN Genitive
- PFV Perfective aspect (了 *le* ‘le’)

Acknowledgement

First and foremost, my faithful and deep gratitude goes to my supervisor Prof. Dr. Marianne Gullberg. Without her help, I could not have completed this master dissertation. I also thank Dr. Mechtild Tronnier for good advice concerning phonetics.

Without the generous Rausing scholarship, this work would not exist. Thanks, Dr. Dehong Meng for the hospitality at Beijing Foreign Studies University. Without the participation of students in Beijing and Hebei, this field will not have been explored yet.

1. Introduction

As in any other language, native Chinese speakers have various communicative strategies for expressing imperatives. A coherent imperative unit not only conveys the speaker's intent but also shapes an action, or a command. The imperative has been examined widely from different angles such as syntax and pragmatics and also prosody.

Syntactically, the Chinese imperative appears to be a simple structure due to the poor morphology of the language. There does, however, exist an inverted ordering, Object-Verb in the so-called 把 *bǎ* 'Ba' construction (henceforth BA), which is often used as an imperative (e.g., Thompson, 1973). Pragmatically, the main feature of an imperative is to make others do something with different degrees of strength, from strong to soft orders: commands, requests, consulting, and asking for something (e.g., Fan, 1998; Ma, 1995; Shi, 1980; Yin, 2013). Despite previous contributions to understand BA, much is unknown, for example about the prosody of BA and gestures that accompany it. We are auditorily aware of prosody, for instance rising or falling pitch, which might be interpreted as having different meanings when the conversation is taking place. Also, we are visually aware of gestures, and these might convey various information. Indeed, communication also consists of intentions, motivations, and emotions of the speaker's performance beyond the neat syntax when sociocultural and interactive instructions occur. Additionally, in previous studies of BA, syntax and pragmatics have been treated separately and have not been based on natural discourse. No study to date has explored BA from a multimodal approach. Therefore, this study attempts to provide a basic understanding of the BA where prosody and gestures are taken into consideration as they occur in natural discourse.

The organization of this dissertation is as follows: section 2 provides a theoretical background and information on earlier studies on BA in Chinese imperatives. Section 3 contains the aim and research questions of the current study. The methods will be given in section 4, including data collection, data treatment, etc. The results will be summarized in section 5. Finally, section 6 discusses the results and makes suggestions for further research.

2. Theoretical background

A speech act is an utterance that not only communicates but also shapes an action like asserting, asking, and commanding. The actions are generally widely accepted as three purposes corresponding to the declarative, interrogative, and imperative structures (Velupillai, 2012:345-346). Chinese is no exception.

This section first presents a brief outline of previous studies on the syntax, semantics, and pragmatics of the Chinese imperative, and in particular the BA imperative (2.1). Some perspectives of BA (2.2) concerning prosody and gestures (2.3) will be presented. After a brief conclusion of previous descriptive studies, we look at a new approach – multimodality – to analyse BA in imperative discourse with prosody and gestures (2.4).

2.1. Chinese imperatives and the BA in previous studies

Aikhenvald (2010) categorizes the imperatives into canonical and non-canonical types. The canonical imperatives convey a command given to a second person while the non-canonical imperatives orient toward a first or a third person. Commands addressed to a first person are often referred to with the term hortative (Aikhenvald, 2010:4; Chapter 12 in Velupillai, 2012), with a particular function of encouraging and spurring the addressee on. Syntactically, the imperative has a “predicator realized by a verb in the simple form (unmarked for aspect or phase), often preceded by the personal pronouns” (Halliday & McDonald, 2004:337). However, the second person pronoun *you* is often omitted (Kroeger, 2005:199-203).

Chinese does not employ morphosyntactic inflection to mark the grammatical components that form an imperative. This section will explain us to understand that pragmatic factors rather than syntactic rules determine whether or not a subject orientating toward a person is needed in Chinese imperatives. It also shows us to know that the Verb Phrase (VP) of Chinese imperatives vary. The inverted ordering Object Verb (OV) in BA as a communicative strategy is used in the imperative.

Additionally, Chinese is a topic-prominent language (Li & Thompson, 1989), which

also allows for ellipsis of arguments, like the subject (cf. Cheng, 2011; Osborne & Liang, 2015), in particular in face-to-face dyadic communication (Jing-Schmidt, 2005:46). It has also been pointed out that pragmatic factors rather than syntactic rules determine whether or not a subject is needed in Chinese imperatives (see also Chapter 9 in Jing-Schmidt, 2005).

Therefore, the Chinese imperative appears to be a very simple structure due also to the poor morphology of the language. There can be either a VP consisting only of a bare verb root as in (1) or various VPs regardless of absence of the personal pronoun. VPs can be preceded or followed by different components such as an adverbial element or a nominal element of degree, like (2)-(3). The adverb 慢 *màn* ‘slowly’ is placed before the verb 走 *zǒu* ‘go’ and 点儿 *diǎn’r* ‘a little’ as a “quasi-comparative expression” (Halliday & McDonald, 2004:337) is placed after the adjectival verb 快 *kuài* ‘quick’. VPs also can be extended after the main verb by different complements (C) signalling the degree, result, direction, or destination of the action, which will be given further concerning BA.

(1) 吃

chī

‘Eat!’

(Li & Thompson, 1989:451, characters added in all examples)

(2) 慢慢走

Màn mǎn zǒu

‘Leave slowly’

(Halliday & McDonald, 2004:337)

(3) 快点儿

Kuài diǎn’r

‘Hurry up’

(Halliday & McDonald, 2004:337)

(4) 你请喝酒

Nǐ qǐng hē jiǔ

‘Please drink some wine’

(Li & Thompson, 1989:454)

Let us now turn to the more complicated VP structure, or VO, as in example (4). The verb 喝 *hē* ‘drink’ simply bears the object 酒 *jiǔ* ‘wine’. There is an inverted ordering, or OV, which also has a close relationship to the imperative, or 把字句 *bǎzìjù*, literally *bǎ* character sentence (i.e., BA).

Historically, 把 *bǎ* ‘Ba’, like other standard verbs, has its lexical meanings, such as ‘take’, ‘hold’, ‘handle’, and ‘grasp’ (cf. Thompson, 1973). It bears a BA-object, as in (5).

(5) 他把你害了

Tā bǎ nǐ hài le

‘He hurt you.’

(Huang et al., 2009:163)

It has the meaning of ‘he disposes of you in a way described as hurting’. In this way, 把 *bǎ* ‘Ba’ was named first 处置式 *chùzhìshì* ‘disposal structure’, which “states how a person is handled, manipulated, or dealt with; how something is disposed of; or how an affair is conducted” (Li & Thompson, 1989:468). However, the lexical meaning of 把 *bǎ* ‘Ba’ as a main verb has lost its properties since it fails the “traditional verbhood tests” (Huang et al., 2009). That is that *他把你害了 *tābǎnǐ hàile* will be ungrammatical when 把 *bǎ* ‘Ba’ bears an aspect marker 了 *le*.

Because of the loss of verb function, there are two main basic conditions to form a 把 *bǎ* ‘Ba’ sentence: definite/generic objects and serial verb constructions. The object after 把 *bǎ* ‘Ba’ should not be indefinite (Li & Thompson, 1989:465), as in (6).

(6) 她自己把药吃了。

tāzìjǐ bǎyào chīle

‘She took the medicine herself.’

(Po-Ching & Rimmington, 2015:201)

Notice that there is no definite article in Chinese, but in the English translation it is marked as a definite/generic object. Generally, a 把 *bǎ* ‘Ba’ sentence cannot end with a monosyllabic verb, which will be explained in § 2.2. Instead, it ends in a serial verb construction or ditransitive verb.

In fact, the serial verb constructions as a complement strategy (Dixon, 2010:405-407; Halliday & McDonald, 2004:312) is widely spread, so the verbs work together as one unit and therefore are assumed to be describing a single action. This construction varies depending on the grammatical relationship and the semantic roles (i.e., four complements customarily termed “extensive, resultative, locative, and directive”, see Jing-Schmidt, 2005:18) for different predicates.

An action verb preceding 点 *diǎn* ‘a little’, see also in (3), indicates that it is semantically extensive. This structure is also referred to as reduplication, as in (7).

(7) 尝尝/尝一尝这个菜。

chángcháng/chángyīcháng zhègècài

‘try this dish’

(Lee-Wong, 1994:496)

The example copies all 尝尝 *chángcháng* or a part of the base form 尝一尝 *chángyīcháng*. This can be seen as a specification of the degree of an activity sliding from only once to regularly happening (Dixon, 2010:36-37). This structure also indicates semantically a momentary action (i.e., 尝下这个菜 *chángxià zhègècài* ‘try this dish now’), which also is termed as “brief durational” complement (Po-Ching & Rimmington, 2015:202).

The resultative complement focuses on the end result of an action rather than on the action itself. The end result of an action might also be the transformation or creation of something else, such as 成 *chéng* ‘become’ in (8).

(8) 我打算把这本英文书翻译成中文。

Wǒ dǎsuàn bǎ zhè běn Yīngwén shū fānyì chéng Zhōngwén

‘I plan to translate this English book into Chinese’

The semantic term of location means “an action that causes something (of a definite nature) to be fixed in the location” (Teng, 2016:203), for example, 在...下 *zài ... xià* ‘under...’ in (9). Note that 在 *zài* has the meaning of ‘existing’. However, the term of direction “indicates the action of shifting something or someone (of a definite nature) from one location to another” (Teng, 2016:204), for instance, 到 *dào* ‘to’ as in (10). There is no distinctive nuance between these two categories; location signals “rather the change of location that occurs to the bǎ-NP caused by the action” (Jing-Schmidt, 2005:23).

- (9) 我叫你别把我的车停在树下，
... *Wǒ jiào nǐ bié bǎ wǒ de chē tíng zài shù xià*, ...
‘I told you not to park my car under the tree, ...’

- (10) 请把这张桌子搬到楼上。
Qǐng bǎ zhè zhāng zhuōzi bān dào lóushàng.
‘Please move this table upstairs’.

All aforementioned examples feature the VP as a matrix verb followed by a complement. There can be semantic overlap between these four notions extentative, resultative, locative, and directive, but the distinctions are not absolute (Jing-Schmidt, 2005). The VP in BA has been examined. 把 *bǎ* ‘Ba’ is not a main verb since it has lost the lexical meaning and instead become a direct object marker. That is to say that the BA-construction comprises the marker of 把 *bǎ* ‘Ba’ and its object Ba-O in BaP together with the VP. It has a *disposal* meaning which is often used in imperatives:

“an imperative is a command to a listener to do something, and if a direct object known to the speaker and the hearer is involved, the listener is commanded to do something to the entity referred to by that direct object. The disposal function of the ba construction naturally fits such a role”. (Li & Thompson, 1989:475)

So far, this section has provided an account of the properties of the Chinese imperative. However, there are two other characteristics concerning imperatives which have received much less attention, namely intonation and gesture (Aikhenvald, 2010:287). As Aikhenvald points out, the intonation in the imperative may differ from that in declaratives and interrogatives. Intriguingly, in a pilot acoustic phonetic experiment concerning BA native Chinese participants reported that they needed more information than syntax, especially intonation and gesture, to complete the communication when they had to listen to sentences with BA in citation form. Additionally, a controversial discussion on Chinese is obviously the issue of interaction between lexical tones and intonation in speech.

2.2. The prosody of Chinese imperatives and the BA

In this section we start with a short presentation of the Chinese tone system, where we will discuss the relationship between tone, stress, and intonation. Also, we will see the phonological and phonetic properties of Chinese imperatives and BA.

Each Chinese syllable (in general equal to one character) contains in principle three components: consonant, vowel, and tone (Wang et al., 2004:81). The syllable <ba> distinguishes the meanings from <pa> and <bi> due to the initial consonants /p, b/ and the different vowels /a, i/. Moreover, by using contrastive tones, that is changing the F0 (measured in Hertz, or Hz, the acoustic correlate of perceived pitch) the syllable <ba> receives different meanings: 八 *bā* ‘eight’, 拔 *bá* ‘pull’, 把 *bǎ* ‘Ba’, and 爸 *bà* ‘father’; as well as a neutral tone 吧 *ba* ‘mood-tag particle’. The diacritic symbols ^ˉ, ^ˊ, ^ˇ, and ^ˋ above the syllable <ba> is an orthographic way to show how the tone contour performs (Duanmu, 2007:225-229) in ordering: high level (H), rising (LH), falling-rising (L), and falling (HL). However, these monosyllabic tones can affect each other in circumstances in spontaneous communication.

In Chinese linguistics the definition 字调 *zìdiào* ‘character tone’ is generally accepted for a monosyllabic, which distinguishes it from the other term, 词调 *cídiào* ‘word tone’. Even though a solitary monosyllable can be accounted for as a word, more than 70% of modern Chinese words are disyllabic (Duanmu, 2007:160; Wang et al., 2004:81). Phonetically, concerning so called 正常重音 *zhèngcháng zhòngyīn* ‘unspecified

stress' disyllabic words (Lin et al., 1984) the second syllable normally has more stress than the first one. However, in their study contrastive tonal pairs like 大意 *dàyi* 'main idea' vs. 大意 *dàyi* 'careless' and a default neutral tone in the second syllable of disyllabic words (i.e., specified stressed in the first syllable) like 月亮 *yuèliang* 'moon', 地方 *dìfang* 'place', 办法 *bànfā* 'method' are not included. Concerning trisyllabic words (Yan & Lin, 1988) the third syllable has more stress than the other two ones. They also found that the duration of a stressed syllable is longer than unstressed syllable(s). However, there are two stress patterns concerning Chinese tonal phonology, iambic and trochaic (Duanmu, 2007), respectively. He has assumed that trochaic is favoured since neutral tone never has occupied the first syllable. However, where might the stress be placed when all syllables in a polysyllabic word contain the same high tone, for example 高中生 *gāozhōngshēng* 'high school student'? In this case, modern Chinese lexicology must be considered. Thus, it is very hard to judge which syllable, in particular concerning the polysyllabic words or phrases, bears the stress (Jin, 1996).

A tone might be changed once it combines with another one. This tonal phenomenon is called tone sandhi (Chen, 2000:xi) and occurs at non-monosyllabic tone levels by combining different tones. In particular, it occurs when a L combines with another L; the preceding L will change into a LH as in (11).

(11) 你把我原来的意思讲走了。

Nǐ bǎ wǒ yuánlái de yìsi jiǎngzǒu le

'You distorted/did not convey my original meaning'

The preceding 把 *bǎ* 'Ba' is changed into *bá*, or L+L → LH+L due to the following 我 *wǒ* which also is pronounced with a L. In this case, tone sandhi influences even the tone of 你 *nǐ*, or *ní* (Po-Ching & Rimmington, 2015:201, the tone sandhi analysis is mine).

Apart from the tone sandhi, a monosyllabic bare verb in the colloquial BA is never allowed except for in poetry and lyrics (Feng, 2001). That is to say that *把他打 *bǎtā dǎ* 'hit him' (Feng, 1996, my English translation) is ill-formed due to the monosyllabic verb 打 *dǎ* 'hit', which makes the BA (two syllables) heavier than the VP (one syllable). Feng (1996) proposed that BA can be revised to a well-formed structure, for instance, by adding

adverbs 一个一个地 *yīgèyīgēde* ‘one by one’ before the verb 刷 *shuā* ‘wash up’ as in (12).

(12) [把碗] [一个一个地刷]

bǎwǎn yīgèyīgēdeshuā

‘wash up the bowls’

A complement could also be added, such as 干净了 *gānjìngle* ‘clean’ after the verb 洗 *xǐ* ‘wash up’ in [把脸] [洗干净了] *bǎliǎn xǐgānjìngle* ‘wash up (your) face clean’. It is plausible that the VP could bear the stress of intonation since it has become heavier by adding syllables. Feng concludes that such restrictions on the bare verb effect are governed neither by syntax nor by semantics, but primarily by prosody. However, VPs bearing stress through the addition of syllables might not match all cases since BaP [把碗] *bǎwǎn* ‘bowl’ in (12) can also be extended such as [把那些桌子上没洗的碗] *bǎnàxiē zhuōzǐshàng mǎixǐde wǎn* ‘those unwashed bowls on the table’. Adding syllables in order to lengthen VP alone without acoustic empirical evidence is not sufficient to judge the stress of the sentence. In spontaneous conversation, one might be less likely to add syllables with the special intention to stress the VP. Thus, it is possible that one syllable may become a pitch accent within a phrase due to tonal circumstances.

Feng also adds that the BaP might not be assigned to stress since it is not new information for either speaker or listener, which seems to match with pitch accents tending to be high on “new” information in the discourse (cf. Debreslioska & Gullberg, 2019). Unfortunately, there is no phonetic support for Feng’s claim. In addition, his examples are not based on natural discourse. No doubt that Feng has paid much attention to the stress theory, but not to the relationship between stress and focus. If BaP is assigned stress, it may well trigger a conflict to the inverted order OV as communicative strategy, which perhaps achieves an “emphasis on intention and specific action [...], either to make requests or to give orders” (Po-Ching & Rimmington, 2015:206). Also, their data give no phonetic empirical data as to how the emphasis is realized in spontaneous speech. It is unclear which of the grammatical components (BaP or VP) will bear the stress of the sentence and how these two main grammatical components might anchor the sentence stress in the case of

BA.

It is assumed that the intonation in the imperative mode may differ from that in the declarative and the interrogative (Aikhenvald, 2010:29). Many studies (e.g., Jin, 1992; Lu, 2006; Lu & Sun, 2008; Shen, 1994; Shi & Jiao, 2016; Sun, 2008; Xu, 1999; Zhou, 2006) are concerned with the contrast between imperative intonation and intonation of other clause types like declaratives. Xu (1999) has investigated focus and non-focus concerning declaratives in very strictly controlled conditions, for instance each monosyllable (i.e., character) in the whole utterance must have the same high tone in 貓咪摸貓咪 *māomīmōmāomī* ‘Kitty strokes kitty’. He has found that the tonal contour in the end of the whole intonational contour (i.e., second ‘kitty’) shows a clear declination. Interestingly, when it comes to common commanding, not all lexical tones keep their pitch prominence (Zhou, 2006). Jin (1992) had informants read imperative sentences which showed that the stressed syllable in polite imperative sentences showed longer and more prominent acoustic properties. Shen (1994) concludes that the duration of the focus tones is shortened in imperative clauses, which implies an absolute order; the listeners experience a softer command if the duration is not shortened. Further, Lu (2006), Sun (2008), and Zhou (2006) report that the duration of an imperative clause is shorter than that of a declarative one; a shortening, as a preparation, usually occurs directly before the last accented syllable. Lu & Sun (2008) have found that the whole imperative intonation rises when related to a strong mood; especially, the tone of the focus part in the material rises dramatically; thus, the pitch range is compressed.

In short, previous studies differ due to various research purposes and data collections, but some kind of consensus among them on the Chinese imperative intonation can be discerned. Firstly, there is an important interplay between tones and intonation in imperatives. Secondly, the aforementioned studies are examples of pure acoustic research and the connection between grammatical components and intonation is still missing. Recently, Shi & Jiao (2016) have found that commanding verbs and negative words bear the focus stress of commanding sentences where the duration and the intensity of these two components are prominent. Thirdly, it is still not known whether high pitch accents assigned to “new” information in discourse might be applied to Chinese. Finally, most of

the previous accounts of imperative intonation are carried out by using acoustic data that are very much oriented to constructed sentences or citation forms.

Very little is known about intonation of BA in natural discourse. In the current study, we will neither predict nor investigate how intonation of BA is manifested due to the lack of acoustic research and further because of the uncontrollable spontaneous dyad interaction. For example, BA might not exist alone as test sentence in earlier studies; it might combine with discourse particles, sequential utterances, pauses, and so on. The intonational contour in the end of BA might not appear as a declination. Also, it is impossible that the whole BA contains one and the same tone in spontaneous speech. Additionally, it is very hard to judge where the prominence is placed in Chinese due to the complex effect of acoustic parameters concerning tones, such as lengthening, F0 range, intensity, and vowel quality beyond F0 although it will be analysed with acoustic tool. Therefore, this study is limited to find where the stress of BA is placed according to perception. Once the stress is determined the acoustic analysis will be considered. That is to say that the stress is based on the information where the hearer first perceives the stress. Beyond intonation, visual effects such as eye gaze or gestures may play a role for the imperative.

2.3. The gestures of Chinese imperatives and the BA

This section will introduce the notion of (manual) gestures using the definition of gesture based on Kendon (2004) in large part. Next, we will examine gestural studies on Chinese imperatives and the BA.

Gesture is “a core of visible bodily movement phenomena” (Kendon, 2004:16), which are produced by kinesic body articulators such as the head, the eyes and eyebrows, upper limbs like arms, hands, and fingers, and the torso (I here disregard movement of speech organs such as the tongue, lips, etc.)

The movements can occur without speech as in sign language, pantomime, and emblems, or with speech. For instance, the V sign is accepted by different cultures to signify victory or peace as an emblematic gesture (Knapp et al., 2014). The movements involved while actually performing practical tasks are not treated as gestures; for instance, in the current study dusting flour over the chopping board when making dumplings. But

hand movements produced by the instructor that “allude to and represent the performance of such behaviours are definitely considered gestures” (Knapp et al., 2014:200). Also, gesture “refers to visible bodily activity that is regarded as serving as an utterance or as a component of an utterance” (Kendon, 2004:110). Utterances here include speech and gestures. The meanings of co-speech gestures, in contrast to sign language, are not “full-fledged linguistic systems with segmentation, compositionality, a lexicon, a syntax, distinctiveness, arbitrariness, standards of well-formedness, and a community of users” (McNeill, 1992:38). Rather, co-speech gestures take their meaning from co-occurring speech. Also, co-speech gestures tend to align with new rather than old information (Debreslioska & Gullberg, 2019; Levy & McNeill, 1992).

The current study draws on two theoretical frameworks as outlined in Kendon (2004:108-157) one is the definitions of gesture unit and gesture phrase; another is the mutual co-occurrence of gestures and speech in communication. According to his view, an entire gesture expression is depicted as an “excursion”, where the articulator moves away from one determined spatial position of rest (i.e., preparation represented by ~) to another space where a sudden fundamental change is created (i.e., stroke represented by *), and then moves back to the initial position where the articulator started (i.e., recovery represented by -). Such an excursion is defined as a gesture unit consisting of a gesture phrase and recovery, whereas preparation and stroke are comprised in a gesture phrase. The essential component of a gesture phrase is the stroke which may be preceded by an optional preparatory phase, in which the articulator (e.g., the hand) is brought to the point in space where the stroke occurs, and is followed by an optional retraction phase, in which the articulator is brought either back to its starting point or to another point to begin the next gesture.

Further, Kendon crystalizes how gesture and speech interact by defining these two aforementioned analytic units. Gestural strokes may occur after speech or vice versa or simultaneously by reforming, adding, adjusting, and changing respective components in communication. A key point here as in (Kendon, 2004) is how a gestural stroke is aligned with the primary stressed syllable. The following figures summarize these key notions of Kendon’s view.

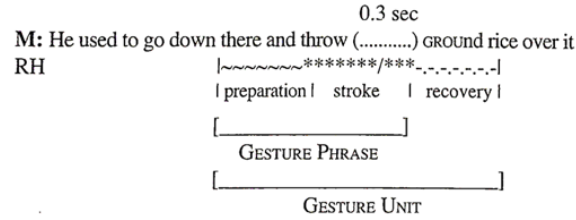


Figure 1 An example of gestural stroke preceding the stressed syllable (Kendon, 2004:114). ~ = gestural preparation; * = stroke; - = recovery; ... = pause; SMALL CAPS = stress.

For instance, in Figure 1 the gestural stroke performed by M's right hand (RH) is taking place along with and crossing the verb *throw* followed by a 0.3 second pause; immediately after the stroke is finished, the modifier *GROUND* of the object "ground rice" is stressed aligning with the gestural recovery. In contrast, another example (Kendon, 2004:117) shows that the gestural stroke produced by RH/LH referring to the subject's right and left hands starts at "WINDOW" momentarily synchronised with the stress as in Figure 2.

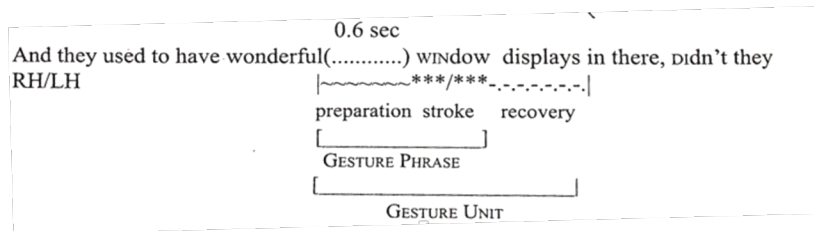


Figure 2 An example of gestural stroke aligned with the stressed syllable (Kendon, 2004:117). ~ = gestural preparation; * = stroke; - = recovery; ... = pause; SMALL CAPS = stress.

Kendon's research on multimodality has drawn great attention in languages other than English. Li (2014) investigates how grammatical components (topic-comment and copula-complement), prosody (relationship between tones and intonation), and body movements interact in turn-taking in Mandarin conversation. She concludes that these components are much more distinguished than they are in English; the interaction between Chinese lexical tones and local intonational pitch track plays a role; body movements often occur at the boundaries of turn construction units. However, the criteria of intonation units governed in her study are constrained to the acoustic standardizations, which differs from the pitch

accent analytic method established by Pierrehumbert and Beckman's (1988) annotation mentioned in the last section.

So far, empirical studies like that of Li (2014) concerning the interaction between BA-construction, prosody, and gestures in conversation by Chinese native speakers have not been found since this research field is still a virgin landscape at least in Mainland China (cf. Chen et al., 2013).

2.4. Summary

In short, most of the influential investigations concerning BA in Chinese imperatives have focused on the complex relationships between syntax, semantics and pragmatics in speech and writing, chiefly in traditional clausal analyses, as evidenced by the abundant literature reviewed above. Concerning intonation, BA and Chinese imperatives have not always been examined in conjunction in natural discourse. Most of the previous accounts of intonational phenomena has been conducted using acoustic data where typically non-prosodic restrictions are imposed on the data to be examined and are very oriented to constructed sentences.

The triple relationship between sentence stress, gestures, and BA in native Chinese imperative discourse has been sadly lacking. To this end, the current study attempts to fill this gap. It will therefore embrace a multimodality-based approach to examine how gesture, prosody, and BA align with each other in spontaneous discourse.

3. The current study

The aim of the study is to observe how grammatical, prosodic, and gestural components interact in the BA construction in Chinese imperatives in spontaneous discourse, applying a multimodality approach. The study focuses only on manual gestures and only on perceived sentence stress concerning prosody. The enormous corpus of non-BA imperatives, their movements, and pitch tracks will be left to future research. Further, due to the limited space, all aspects relevant to the understanding of the multimodality, such as sociocultural aspects must also be excluded.

Therefore, the questions to be addressed in this study are the following:

1. Do hand movements occur with BA?
2. How are they aligned with the grammatical components of BA in communication?
3. Which grammatical components, i.e., BaP or VP, bear perceived the sentence stress of BA?
4. Do gestures synchronize with the stress of the BA?
5. Are gestures aligned with the VP rather than the BaP since the BaP does not express new information (cf. Levy & McNeill, 1992)?

Based on these questions, this study makes the following predictions:

Concerning gestures, the hand movements will align with both BaP and VP since gestures and speech interact mutually.

Concerning prosody, the sentence stress will correspond to the verb in VP due to the fact that the BaP expresses old information.

Concerning interaction between sentence stress and hand gestures, this will be temporally aligned with the VP.

The research questions are examined in dyadic interactive activities, designed for the current study, in which participants are making dumplings. The analyses therefore draw on multimodal production data from these dyads. The production data is coded and analysed for linguistic and intonational properties and gesture. In addition, the speech data is subjected to a perceptual study to determine perceived stress.

4. Methods

4.1. Participants

The participants were recruited with the help of Dr. Dehong Meng, lecturer at the School of Chinese Language and Literature (SCLL), Beijing Foreign Studies University (BFSU), when the current study was conducted there in November 2016. The participants were students at SCLL at the time. The recruitment was announced after lectures on different occasions at the very beginning of the project. On each occasion Dr. Meng introduced me to the participants for the sake of acquaintance. Finally, 16 undergraduate students with

Chinese language as major (14 female) and a mean age of 21 years (range 19-21, $SD = .602$), participated in the production study. Their demographic information is shown in Table 1.

Table 1 The demographic information of the participants in the production study

Gender		Age		Ethnicity	F M		Language profile	F M	
Female	Male	Min	Max						
		19		Han	13	2	Beijing Mandarin	2	
14	2	21		Mongolian	1		Northern Mandarin	9	
							Southern dialects	3	2
16				16				16	

One female belonged to the Mongolian minority; the remaining people were Han people. They came from different areas in China with various regional dialects. Two females spoke Peking Mandarin and nine females reported they use northern Mandarin. Five participants had different southern dialects. The Mongolian was not functionally bilingual; according to her own report she doesn't even understand Mongolian. None of the southern dialect participants spoke their dialects. All had passed the national enrolment examination for college, and in particular their Standard Chinese got extremely high credits. Additionally, all had passed (国家) 普通话水平测试 (*Guójiā Pǔtōnghuà Shuǐpíng Cèshì* (PSC) '(National) Standard Chinese Proficiency Tests' at the time and all were at Level 1-A (97-100 of 100 full credits). They were freshmen. The only major difference among the participants was that, although all spoke standard Chinese fluently, they came from different dialectal regions. Dialectal differences have, for the time being, been ignored since they all spoke standard Chinese at the time of recording. The participants were compensated for their participation with a 100-yuan bonus check (approximately 150 SEK in accordance with the exchange rate at the time).

An additional 35 freshmen (29 female) and a mean age of 19 years (range 18-22, $SD = .79$), were recruited at Hebei Foreign Studies University (HFSU) for a subsequent perception test (see Data treatment and coding) in November 2019. They had also passed the national enrolment examination for college and national Standard Chinese examination. They come from the same dialectal region (Hebei Province) where Beijing is embedded

geographically. Importantly, none of the participants in both universities had studied linguistics at the time.

4.2. Material and the design of the tasks

The material of the study included not only the catering facilities and ingredients used in the production task but also video recording instruments. The dyadic production task was intended to consist of a natural, task-driven interactive discourse. The perception task was also designed in order to avoid conservative phonetic mislabelling.

4.2.1. Materials for the production task

The catering facilities (e.g., a chopping board, a rolling pin, mixing bowls, chopsticks, a jug, and side plates) and the ingredients used for the task were made available (see Table 2).

Table 2 Main ingredients, stuffing, and spices

Main ingredients	Stuffing	Spices
Plain flour	Minced chicken	Salt
Flour for dusting	Fine-chopped Chinese cabbage leaves	Monosodium Glutamate (MSG)
Water	Fine-chopped spring onion and ginger	Pepper
		Sesame oil

4.2.2. The production task design

The production task should ensure the occurrence of gestures and prosodic variations and the circumstances of the discourse should be as natural as possible. Therefore, an interactive face-to-face task was considered. A teaching-learning task of making dumplings was chosen for the following three reasons.

Firstly, imperatives are good for advice and instruction (Aikhenvald, 2010:1). Cooking instructions are a good domain, when one individual as instructor teaches another person. Written imperative clauses are also frequently used in cooking recipes. Secondly,

it was assumed, following Kendon (2004) as mentioned in 2.3, that certain main action verbs (e.g., 撒点扑面 *sǎdiǎnpúmiàn* ‘dust flower over the chopping board’) may be accompanied by gestures (e.g., shaking the hand as if scattering flour) in such an instructive setting. Thirdly, an instructor does not need to follow a recipe, which in turn might lead to gestures being performed unconsciously in a relatively natural way.

4.2.3. The perception task design

As described in section 2.2, Chinese stress is still a controversial issue. It is a very subjective task to judge stress by only observing pitch tracks and amplitude contours in Praat. Therefore, it was considered more objective to present a perception task where many listeners determine where the sentence stress is placed. Participants who did not take part in the production task were presented with the transcriptions of the recordings and asked to mark the placement of stress in the discourse.

4.3. Procedures

4.3.1. The production task

During recruitment, participants were orally informed in Chinese that the intention of the project was to study language teaching interaction. Once the participants had confirmed participation, the author met them separately, in their dyads. They had then signed a consent form (Appendix I). They were also informed that even at this stage they could withdraw at any time and that their data would then not be used. Their anonymity is guaranteed. They were allowed to choose roles as instructor or learner, and they were told how the process of making dumplings works. The instructor ought to have some prior knowledge of making dumplings, but this was not absolutely necessary. They were provided with a brief outline of the material, the production task, and the experimental process.

The process contains six events: (1) prologue, (2) mixing, (3) ripping, (4) rolling, (5) wrapping, and (6) epilogue. The recipe used in the study was created after consulting Chinese online cooking resources. The main four steps for making dumplings are described

below. The instructor in each dyad had to find her/his own way to communicate with the learner.

Mixing

Pour the flour into a bowl and stir the water gradually into the bowl. Mix with chopsticks or hands until it is elastic and becomes a dough. Scatter a little flour on the chopping board and knead the dough well. Leave the dough to rest. Place the stuffing items in another bowl. Mix them with spices together thoroughly. Put it aside.

Ripping

Roll the dough out so it becomes like a thin and long roll. Hold the roll in one hand and leave a little bit dough as big as an egg yolk, then tear it off sharply. Roll the small dough into a small round ball. Dust the ball and flatten it softly.

Rolling

Take a small piece of the flattened dough ball in one hand. Roll over it with the rolling pin. Keep the dough rotating while rolling slowly. Shape the dough into a circular wrap.

Wrapping

Place the wrap on one palm, adjust the mince with chopsticks into the middle of it. Fold one side to the opposite side and pinch it with the fingers.

To save time, the participants were presented with a demo dyad (soundless) where Dr. Meng plays the instructor role and the author that of the learner. The actual recording took place in a classroom at the SCLL, BFSU. Before the recording started, all catering facilities and ingredients for making dumplings were prepared and placed on a table in the middle of the classroom, illustrated in the set-up in Figure 4.

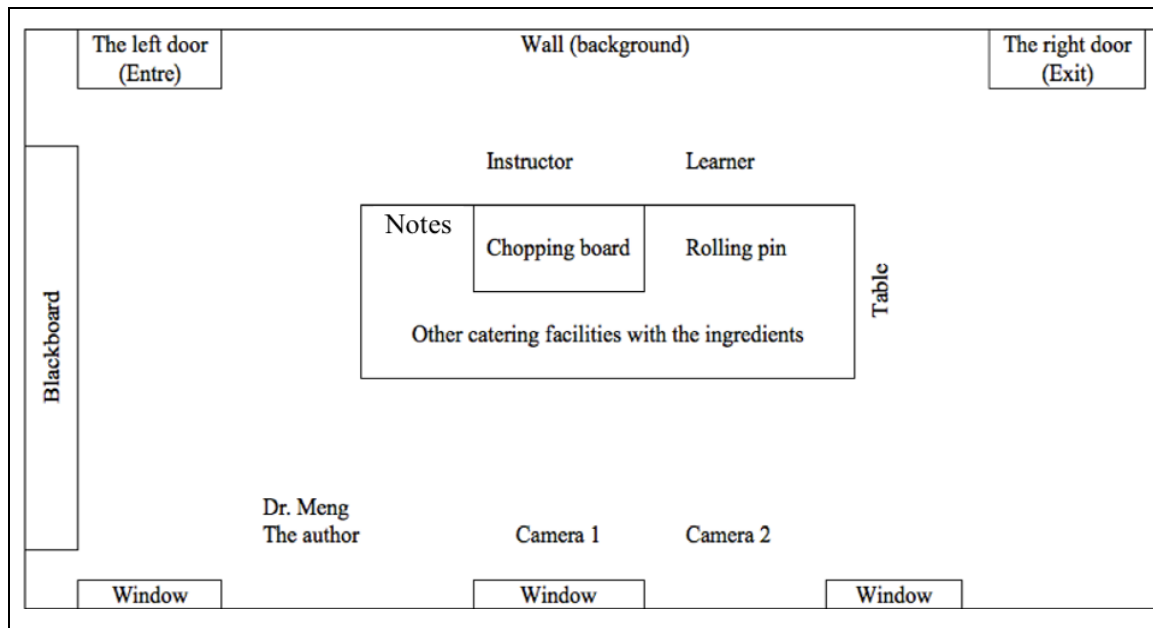


Figure 3 Schematic representation of the experimental set-up

The instructor stood to the left and the learner stood to the right in front of the table where the ingredients were prepared (Figure 3). A printed note with four steps of making dumplings was put on the chopping board in order to remind the participants if/when they did not remember the process. Camera 1 (recording a large front view) was handled by one technician and was placed on the left side, and Camera 2 (recording of a narrow side view) handled by another technician was displayed on the right side in front of the subjects. Both cameras were placed as unobtrusively as possible, approximately 1.5 meters away, so that the dyad participants were in view.

Two cameras were used to capture a large front view (camera 1, Sony PXW150) and a narrow side view (camera 2, Sony FDR AX30) of instructor and learner. When participants came into the room, there was generally a few minutes of relaxed small talk before the real action, so the participants could get more comfortable in front of the visible cameras. An extra sound recording as “backup” (Jun & Fletcher, 2014:512) was prepared for further prosodic analysis.

The aim of the project was revealed after the recording was finished. None of the participants had guessed the precise intention or identified gesture as being the study target. All subjects expressed surprise at how well they had performed and also said they had enjoyed the experiment. Table 3 provides the details of the recording for each dyad.

Table 3 Details of recording for analysis

Dyad	M	F	Duration
1	2		00:08.49
2		2	00:05.13
3		2	00:05.56
4		2	00:05.11
5		2	00:05.41
6		2	00:05.24
7		2	00:04.22
8		2	00:04.42

There was only one male dyad. One in group 6 was Mongolian. The sessions lasted approximately 5.5 minutes on average (range 04:22 – 08:49 mins, $SD = 1.314$).

4.3.2. The perception task

The sound files of the video-recordings of the 16 participants in the production task from SLCC at BFSU in 2016, were presented auditorily to 35 other participants in a classroom at the Department of International Finance, at HFSU in 2019. No especially quiet environment was needed since the task was not orientated to phonetic or phonological precision. The 艾本 200 A 无线 FM 调频耳机 *Àiběn 200 A wúxiàn FM tiáopín ěrjī* ‘Aiben 200 A wireless FM earphone’ was used for the listening.

The transcriptions (see further concerning data treatment and coding) of the video-recordings were presented only in Chinese, which were printed out as an answer sheet (Appendix II). The participants were tested in a group and they listened to each utterance at a time and then mark prominence as speech went along. They listened, on one occasion, to the complete dyad group recording. The participants were asked to only encircle the character of the most prominent placement in the given discourse lines produced by the instructor. The answer sheets were handed in when they were satisfied with their own judgements. Each session took one hour.

4.4. Data treatment and coding

4.4.1. Production data: speech

The videos were saved as video files with incorporated audio. They were also saved as independent sound files in order to annotate intonation. All sound files were listened to by Dr. Meng and two other lecturers of Chinese in order to make sure their Standard Chinese proficiency meets the requirement of the current speech production task.

4.4.1.1. Transcription

The speech from the video was first automatically rendered into Chinese characters by 大洋视频编辑系统 *dàiyáng shìpín biānjí xìtǒng* ‘Dayang Video Editor System’. Beyond the major job made by this editor system, the remaining (pauses, spurts, discourse particles, or speech disfluencies, the Pinyin, and English translation) had to be transcribed manually. The videos were also saved as independent sound files in order to annotate the intonation. The transcription system used for the discourse in the study is adapted from Ochs (1979) and Atkinson and Heritage (1999) with a few small modifications (cf. Excerpt 1).

Excerpt 1 Group 1, Male instructor, line 003-004, 00:04.31-00:07.55

Line	Instructor	Learner
	...	
→ 003	嗯要把面：倒入 (0.2) 倒	
	en yào bǎmiàn: dào rù (0.2), dào	
	mhm shall Ba pour into (0.2), pour,	
	<i>mhm, pour the flour into, pour,</i>	
004	面粉倒入碗里面 (0.6)	
	miànfěn dào rù wǎn lǐ miàn (0.6)	
	flour pour into bowl in (0.6)	
	<i>pour the flour into the bowl.</i>	
	...	

007	哦，把面倒入碗中
	o bāmian dǎorù wǎnzhōng
	mhm, Ba flour pour into bowl in
	<i>Oh, pour the flour into the bowl</i>

Each line equal to a “spurt” (Gee, 2011:128-129) is numbered in the leftmost column, which is judged by obvious pauses between spurts (Tao, 1996). The lines are transcribed in Chinese characters, its Pinyin representation (Romanization system), literal translation into English, and the idiomatic translation (*italics*). The content of the discourse by instructor and learner is separated, using a modified version of discourse transcription (Atkinson & Heritage, 1999; Ochs, 1979). The arrow (→) is assigned to the specific line of the discourse produced by the instructor, which will be analysed for the study and be referred to in the context.

Discourse particles, for example 嗯 *en* ‘mhm’ in line 003, 哦 *o* ‘oh’ in line 007 in Excerpt 1, have been labelled but not been accounted for the analysis. The segments that constituted truncated utterances, reactive tokens, as well as onomatopoeic words that do not conform to any analysable grammatical structure are also ignored.

4.4.1.2. Further linguistic coding of BA

Utterances containing only 把 *bǎ* ‘Ba’ were selected and the grammatical components were labelled after the selection. They consist of BaP including the marker 把 *bǎ* ‘Ba’ (Ba), the modifiers of the Ba-object (Ba-O-M), and the Ba-object itself (Ba-O) and the VP including verb (V), its complement (V-C), its object (V-O), and the modifiers of the verb-object (V-O-M). Some utterances are a complete BA containing both BaP and VP meanwhile others are either a BaP or VP. Some contain only the marker 把 *bǎ* ‘Ba’.

4.4.1.3. Gesture annotation

In the production segments already selected containing 把 *bǎ* ‘Ba’, gestures occurring in those segments were annotated in ELAN (<https://archive.mpi.nl/tla/elan>) without reference to the sound files, as in Figure 4.

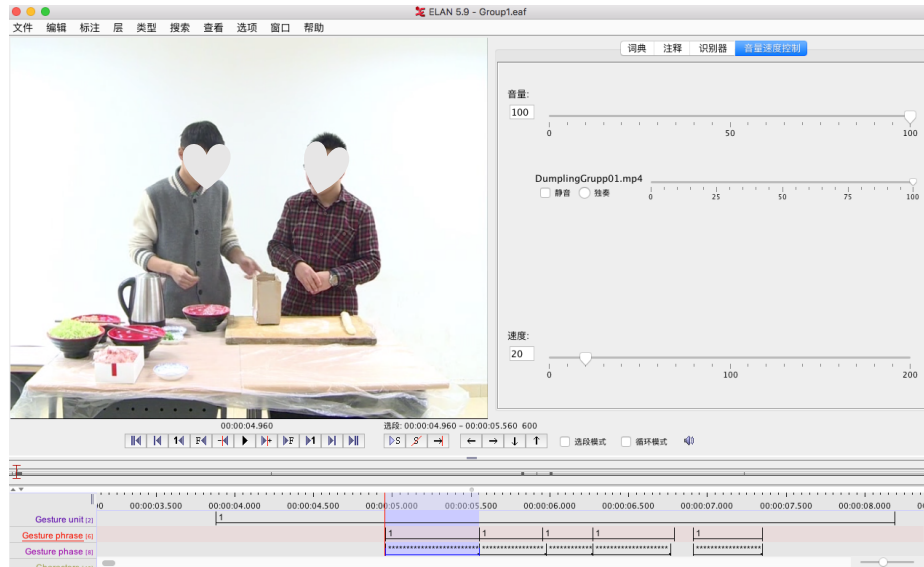


Figure 4 Sample of hand gestures from ELAN, produced by male instructor from Group 1, in line 003 to 004

Figure 4 is taken from ELAN illustrating how hand gestures produced by the male instructor from Group 1 in line 003 to 004 in Excerpt 1 were annotated. The participants' faces are masked. From the top to the bottom in the annotation tiers the gesture unit, the gestural phrase, and the gestural phrase are labelled in turns. The entire gestural unit starts at around 00:03.90 and ends at 00:08.20. This gesture contains five gestural phrases where each phrase in its turn only consists of one gestural phrase, the stroke. The phase coding is the key element in the analyses, and especially strokes (as opposed to preparations, holds, or recoveries). Few discrete gestural strokes are only aligned with pauses, spurts, discourse particles, or speech disfluencies are eliminated.

4.4.1.4. Intonation annotation

The production data sound files were imported into Praat (<https://www.fon.hum.uva.nl/praat/>). The autocorrelation was used to generate a pitch track in order to obtain the accurate pitch. Figure 5 illustrates an analysed BA from Excerpt 1 drawn in Praat Picture.

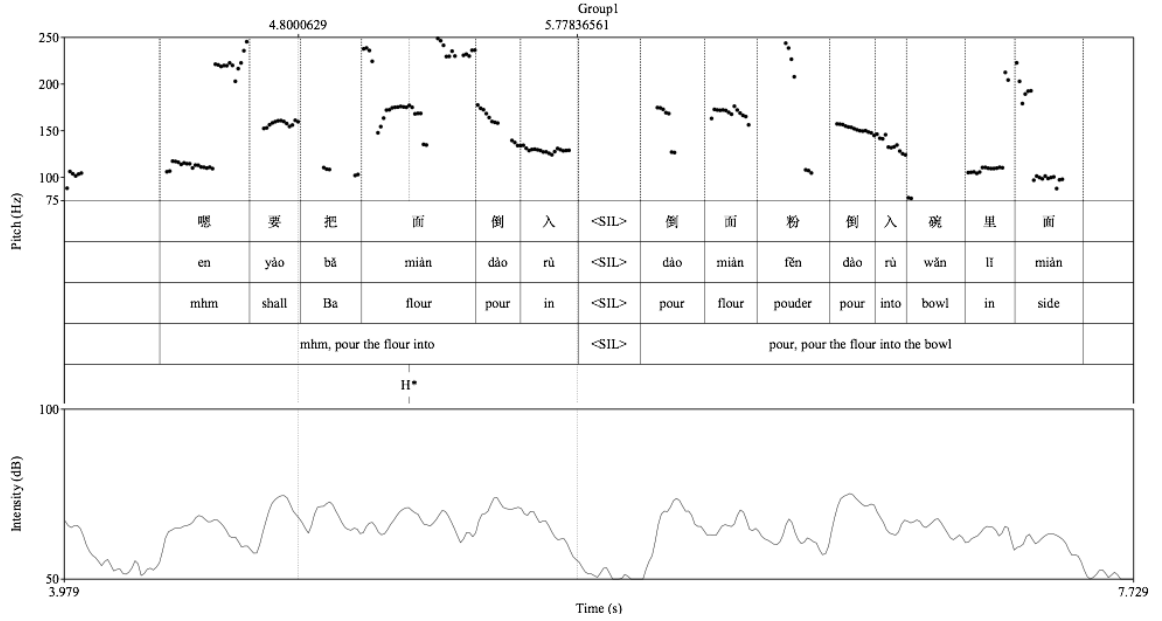


Figure 5 F₀ contour of line 003 to 004 in Excerpt 1 displayed by Praat picture

In Figure 5, the topmost layout is the pitch track and the tiers of characters, Pinyin, and literal and idiomatic translation, respectively, were created and labelled (<SIL> stands for silence). As the focus of the current study was to investigate where sentence stress might be located in BA-constructions, only prominence within the vowel of the associated stressed syllable had to be marked in accordance with the perceptual results. Therefore, once the perceived sentence stress was identified (see further section 4.4.2), the phonetic prominence (either a local high, H* or low, L*) of *Tones* (the *To* in ToBI, Tones and Break Indices, see Pierrehumbert & Beckman, 1988) was labelled in the last tier. The amplitude track (the bottom) was also used as a reference to determine the stress when two prominences were reported in term of the perceptual results. For instance, the Ba-O 面 *miàn* ‘flour’ is the most prominent in BA in Excerpt 1, line 003, shaping acoustically a simple local high peak within the BaP although the amplitude value of the Ba-O 面 *miàn* ‘flour’ is slightly lower than that of the sequential verb 倒 *dào* ‘pour’. In this way, it ensures the accuracy of the acoustic labelling.

4.4.1.5. Alignment

In order to observe how the grammatical components, align with the gestural strokes the Praat TextGrid was imported into ELAN as in Figure 6, used for the analysis in the study.

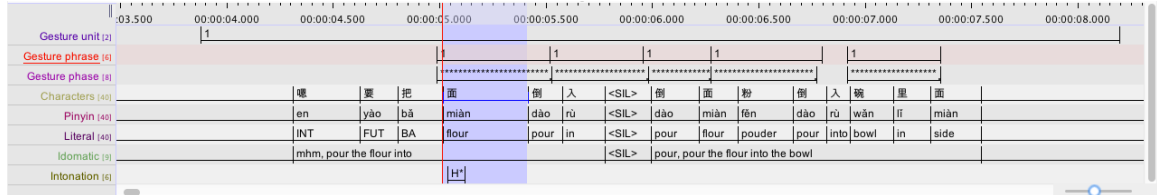


Figure 6 The utterance produced by the male instructor from Group 1 in line 003 to 004 in Excerpt 1

Figure 6 shows how hand gestures align with the grammatical components after the import. The labellings from the topmost to the bottom are the gesture unit, the gestural phrase, the gestural phase, characters, Pinyin, literal and idiomatic translation, and intonation. Only the first gesture is aligned with the BA-construction, starting very late in the phase of the marker 把 *bǎ* ‘Ba’, going through the whole Ba-O 面 *miàn* ‘flour’ and crossing but ends within the verb 倒 *dào* ‘pour’.

In order to observe how the pitch track is manifested when both the prominence and gestural stroke are aligned with the grammatical components, the gesture annotations were exported as Praat TextGrid into Praat, so they could align with the intonation annotations synchronously shown in Figure 7.

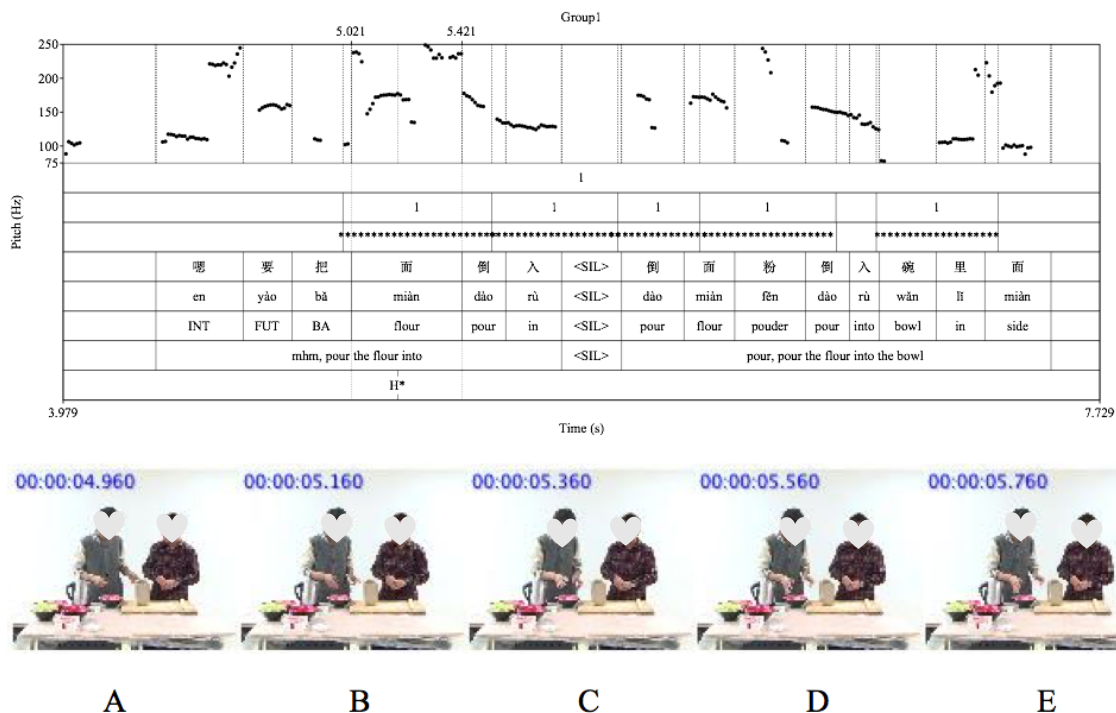


Figure 7 Gestural stroke aligns with the grammatical component perceived as the stressed syllable, male instructor, Group 1, in line 003 to 004 in Excerpt 1

Figure 7 demonstrates how the gestural stroke, the Ba-O 面 *miàn* ‘flour’, and the perceived stress of the utterance 嗯要把面倒入 *en, yào bǎmiàn dào rù* ‘mhm, pour the flour into’, which is phonetically manifested as a pitch accent (H*), are aligned together. The top is annotation of both intonation, gestural strokes, and transcription and the bottom is the filmstrip showing how the gestural stroke is co-occurs with the Ba-O 面 *miàn* ‘flour’ when it is perceived as the stressed syllable. The speaker’s open palm inward hands, with the left forefinger separated from the left thumb pointing forward, are posing extended and shrank swinging once (see C in Figure 7).

4.4.2. Speech perception treatment

The response sheets from the perception task were gathered, and the perceived stress of sentence indicated on those sheets was calculated and sorted according to grammatical components circled. Appendix III summarizes the general performance of the 35 participants with respect to their responses to all collected BA produced by the 8 instructors

that vary in focus. Table 4, extracted and simplified from the Appendix III, shows how stress of the utterance in Excerpt 1, Group 1, line 003 is distributed.

Table 4 The distribution of stress of Excerpt 1, Group1, line 003.

(BaP = Ba-Phrase, Ba = 把 *bǎ* ‘Ba’, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, V-C = Verb Compliment)

Group	Lines	BaP			VP			Total
		Ba	Ba-O	Ba-O-M	V	V-O	V-C	
1	003		23		7	5		35

The perception test was intended to establish more objectively where the stress lies. That is to say that 23 of 35 subjects reported that they perceived the Ba-O 面 *miàn* ‘flour’ as the most prominent in this BA. Once the stress placement has been determined by the participants, it was more easily observed in Praat.

4.5. Analyses

The selected observations are overall presented in descriptive statistics. In total, the eight video recordings contain 74 discourse lines where BA are included, produced by the instructors (see Appendix IV). 47 of those turns (i.e., 64% of turns) with 把 *bǎ* ‘Ba’ are aligned with gestural strokes; 3 of them aligning with spurt and discourse particle are excluded. For 70 turns (95%) perceived sentence stress is reported; and of these only 29 (39%) gestural strokes are aligned with perceived sentence stress.

The relationship between intonation, grammatical components, and gestural strokes in BA was tested with χ^2 -tests, which were conducted in SPSS 27.0 (Statistical Package for the Social Sciences).

5. Results

As presented in section 3, three predictions were made based on research questions, which will be covered in turn below. Each section ends a short interim summary.

5.1. Gestural strokes aligned with BA

In this section the results of the analysis of gestural stroke alignment with BA-constructions is presented. Table 5 shows the 44 cases of gestural strokes occurring with BA broken down by grammatical element with which they co-occur. As can be seen, most strokes cover both the BaP and the VP jointly.

Table 5 Distribution of gestural strokes over grammatical elements in the BA-construction

(BA = BA-construction, BaP = Ba-Phrase, Ba = 把 *bǎ* ‘Ba’, Ba-O = Ba-Object, VP = Verb Phrase, V = Verb, V-O = Verb-Object)

Types of gestural strokes align with		BA				Total
		BaP		VP		
		Ba	Ba-O	V	V-O	
Both	(1)	Ba	Ba-O	V	V-O	4
	(2)	Ba	Ba-O	V		9
	(3)		Ba-O	V		9 22
Only BaP	(1)	Ba	Ba-O			8
	(2)		Ba-O			5
	(3)	Ba				4 17
Only V	(1)			V	V-O	1
	(2)			V		2
	(3)				V-O	2 5
Total						44 44

A Chi-square test of independence was performed to examine the relation between the gestural strokes and grammatical components. The relationship between the two variables was significant ($\chi^2 = (1, N = 44) = 88.00, p = .000$). The test revealed that gestural strokes were significantly more likely to occur with the whole BA-construction than with either BaP or V.

The following will exemplify how gestural strokes align with BA-constructions in detail. Excerpt 2 is taken from Group 6 where the female instructor (left) is telling her learner (line 039) how to roll the small dough. The transcription (top) of the relationship between the gestural strokes and the grammatical components in this example is illustrated in Figure 8; the export filmstrip for the occurrence between these two modalities is extracted (bottom).

Excerpt 2 Group 6, Female, line 039, 02:20.78-02:31.91. The arrow in line 039 indicates the gesture illustrated in Figure 8.

Line	Instructor	Learner
038	对，然后你 duì ránhòu nǐ right then 2SG <i>alright, then you</i>	
→ 039	你现在要把它弄成一个就是 ránhòu nǐ xiànzài yào bǎ tā nòngchéng yīgè jiùshì then 2SG now shall BA it make one just <i>now you shape the dough into a</i>	
040	圆柱体的面 yuánzhùtǐde miàn cylindrical flour <i>cylindrical form</i>	
041	[对 [duì [right [alright	弄成一个[圆柱体的..哈哈..面 nòngchéng yīgè [yuánzhùtǐde...haaaa...miàn make one [cylindrical (laughing) flour <i>make a [cylindrical (laughing) form</i>

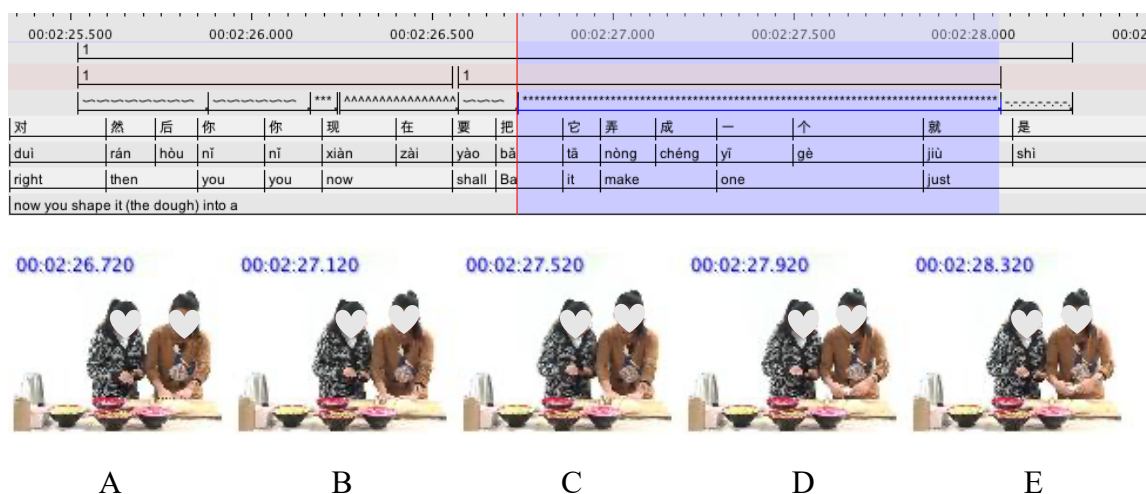


Figure 8 Gestural stroke alignment (***) with the entire BA-construction of line 039 in Excerpt 1, transcription (top) and export filmstrip (bottom), respectively from ELAN.

The gestural preparation starts with 要 *yào* ‘shall’ and continues through the marker 把 *bǎ* ‘Ba’ (see A in Figure 8). Half way through 把 *bǎ* ‘Ba’ the speaker’s arms slightly lift, her hands with separated fingers oriented with open palm down, move up and down, quickly, and repeatedly with each syllable during the production of the rest of the grammatical components (see C-E in Figure 8) until the discourse particle 就是 *jiùshì* ‘just’. The action slows down in the end of 就 *jiù* ‘just’, and the stroke goes into the recovery phase as her hands relax and return to the rest position. This illustrates how a single gestural stroke covers several spoken grammatical components.

In contrast, the next two examples illustrate how a single gestural stroke aligns with only one grammatical component, either BaP (Excerpt 3) or VP (Excerpt 4). Excerpt 3 is extracted from Group 2 where the female instructor (right) in line 009 is teaching the learner how to make the dough. The transcription (top) and filmstrip (bottom) of relationship between the gestural stroke and the grammatical unit are demonstrated in Figure 9.

Excerpt 3 Group 2, Female, line 009, 00:10.55-00:16.06. The arrow in line 009 indicates the gesture illustrated in Figure 9.

Line	Instructor	Learner
008	好, 首先我们来 (0.3) 和面 hǎo, shǒuxiān 1PL lái (0.3) huómian well, first we come (0.3s) mix flour <i>well, we are going to (0.3s) mix flour</i>	
→ 009	然后先把 (0.3) 把面倒到碗里 ránhòu xiānbǎ (0.3) bǎmiàn dàodào wǎnlǐ then first Ba (0.3s) Ba flour pour to bowl in <i>mhm first... (0.3s) pour the flour into the bowl</i>	[小声嘀咕 [xiǎoshēngdīgū] [muttering] Muttering]

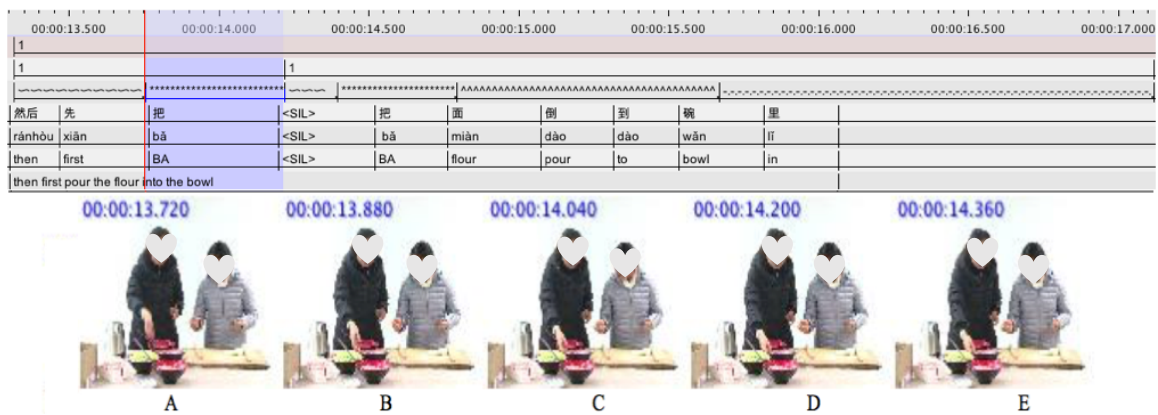


Figure 9 Gestural stroke alignment (***) with only BaP within the BA-construction of line 009 in Excerpt 2, transcription (top) and export filmstrip (bottom), respectively from ELAN.

The except contains two strokes; the focus here is on the first gesture produced by her right hand. The preparation starts at the discourse particle 然后 *ránhòu* ‘then’ and ends at the adverb 先 *xiān* ‘first’ when she is speaking. As she continues saying 把 *bǎ* ‘Ba’ the right hand forms a fist where the tip of the thumb is resting on the first joint of the index finger, the inward palm is extended forward to shape a clear gestural stroke (see A-E, in Figure 9). The hand is then retracted during the short silence that follows. This example

demonstrates how a single gestural stroke occurs with only the marker 把 *bǎ* ‘Ba’ within BaP.

Excerpt 4 is taken from Group 5 where the female instructor (left) is introducing, in line 098, the learner how to wrap dumplings. The relationship between the gestural stroke and the grammatical unit is shown in transcription (top) and filmstrip (bottom) in Figure 10.

Excerpt 4 Group 5, Female, line 098, 04:44.71-04:50.23. The arrow in line 098 indicates the gesture illustrated in Figure 10.

Line	Instructor	Learner
→ 098	啊，然后把那个面片对折 a, ránhòu bǎ nàgè miànpian duì zhé ahh, then BA that skin opposite fold <i>ahh, then fold in half that skin</i>	
099		对折? duì zhé? oppositely fold? <i>fold it in half?</i>

0	00:04:45.000	00:04:45.500	00:04:46.000	00:04:46.500				
		1						
		1						
		~~~~~	*****	~~~~~				
啊	然后	把	那个	面片	对折			
a	ránhòu	bǎ	nàgè	miàn	piàn	duì	zhé	
ahh	then	Ba	that	flour	slice	opposite	fold	
ahh, then fold in half that dumpling wrapper								



Figure 10 Gestural stroke alignment (***) with only VP within BA-construction of line 098 in Excerpt 3, transcription (top) and export filmstrip (bottom), respectively from ELAN.

Here, the instructor's left arm is extended toward the learner with open palm down outward, with the thumb resting under the other separated fingers to shape preparation. The left hand, so formed, then closes the tip of the thumb suddenly under the tips of the other flexed fingers as a bird's beak. This stroke is exactly aligned with the complement 对 *duì* 'opposite' in the verb 对折 *duìzhé* 'fold in half'.

We have seen in this section that the alignment of gestural strokes with BA-construction vary. Sometimes one single gestural stroke is aligned with all grammatical components within BA-construction, while in other cases gestural strokes are aligned with either BaP or V. The next section will illustrate how perceived stress is distributed across BA.

## 5.2. Perceived stress with BA

This section presents the results on the distribution of perceived sentence stress in BA-constructions. Table 6 illustrates which grammatical element is perceived as most prominent in BA-constructions.

Table 6 Distribution of perceived sentence stress over grammatical elements in the BA-construction (BaP = Ba-Phrase, Ba = 把 *bǎ* ‘Ba’, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, H = high, LH = rising, L = dip, HL = falling)

BA-construction		Perceived sentence stress aligned with lexical tones				Total
		H	LH	L	HL	
	Ba			2		2
BaP	Ba-O		2		6	8
	Ba-O-M			1	4	5
VP	V	6	10	3	35	54
	V-O			1	1	2
Total		6	12	7	46	71

As shown in Table 7, the perceived sentence stress mainly falls on the verb ( $n = 54$ , or 76%) in VP regardless of what kind of lexical tone the verb bears, although HL is dominant.

Table 7 Distribution of perceived sentence stress aligned with lexical tones in the Verbs (H = high, LH = rising, L = dip, HL = falling; H* = local prominent high, L* = local prominent low)

Acoustic realization	Perceived sentence stress aligned with lexical tones in the Verbs				Total
	H	LH	L	HL	
H*	6	3	1	30	40
L*		7	2	5	14
Total	6	10	3	35	54

Table 7 further shows that when the verbs (V) in BA, with any of the four lexical tones (H, LH, L, and HL), are perceived as the syllable bearing the sentence stress, they have a tendency to be acoustically realized as a phonetic prominence, be it as a local peak (H*) or a valley (L*) due to the tonal sequences.

We can also see that the lexical tone H is still realized as a H*; however, the lexical tones LH, L, and HL are acoustically manifested either as H* or as L*. The acoustic

realization of the perceived stress and lexical tones as control variables have been examined with the  $\chi^2$ -test. The results show that there is a significant relationship between the lexical tones and the acoustic realization of the perceived stress on the verb,  $\chi^2 = (1, N = 54) = 54.00, p = .000$ . The test revealed that H and HL were significantly more likely to be acoustically realized as H* than LH and L to be phonetically realized as L*, which will be illustrated in Excerpts 5-8.

Excerpt 5 Group 3, Female, line 104, 05:47.69-05:49.55. The arrow in line 104 indicates the example illustrated in Figure 11.

Line	Instructor	Learner
103		<p>再掐一下，按一下</p> <p>zài qiā yīxià, àn yīxià</p> <p>then pinch little, click little</p> <p><i>just pinch it, and press it again</i></p>
→ 104	<p>嗯，然后把那个收个尾儿就可以了</p> <p>en, ránhòu bǎ nàgè shōugèwěi'r jiù kěyǐ le</p> <p>mhm, then BA that close one end just ok PFV</p> <p><i>mhm, then close the end of the edge, so it will be fine</i></p>	

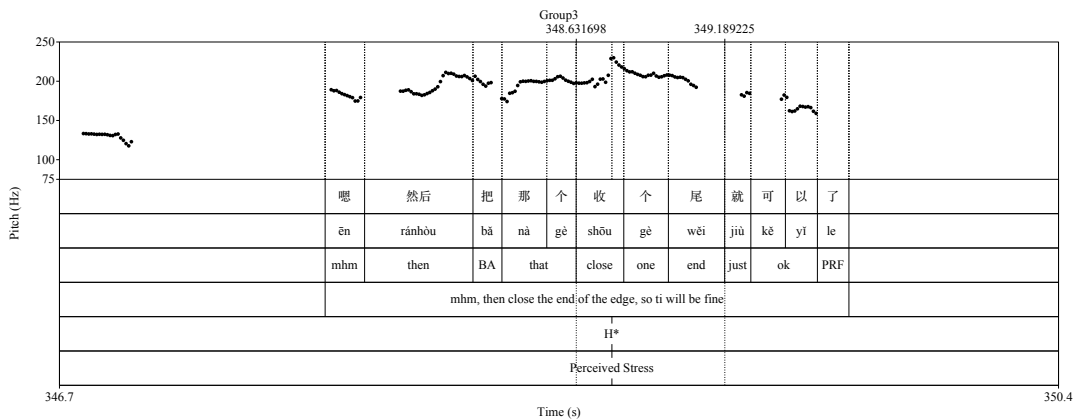


Figure 11 F₀ contour of line 104 in Excerpt 5

As illustrated in line 104, produced by the instructor from Group 3, the most prominent syllable in this utterance ...把这个收个尾儿... *bǎ zhègè shōugè wěi'r* ‘close the end of the edge’ is the verb 收 *shōu* ‘close’ with lexical tone H in the VP 收个尾儿 *shōugè wěi'r* ‘close the end of the edge’. The  $F_0$  contour shapes a rise that peaks at the end of 收 *shōu* ‘close’ (348.631-349.189), which is an example of canonical H* or High pitch accent.

When the verb is exposed as lexical tone HL it also tends to be perceived as H*. The production of the interaction between the lexical tone and the intonation aligned with the verb produced by the instructor from Group 7 is shown line 036 in Excerpt 6 and Figure 12.

Excerpt 6 Group 7, line 036, 01:46.20-01:49.89. The arrow in line 036 indicates the example illustrated in Figure 12.

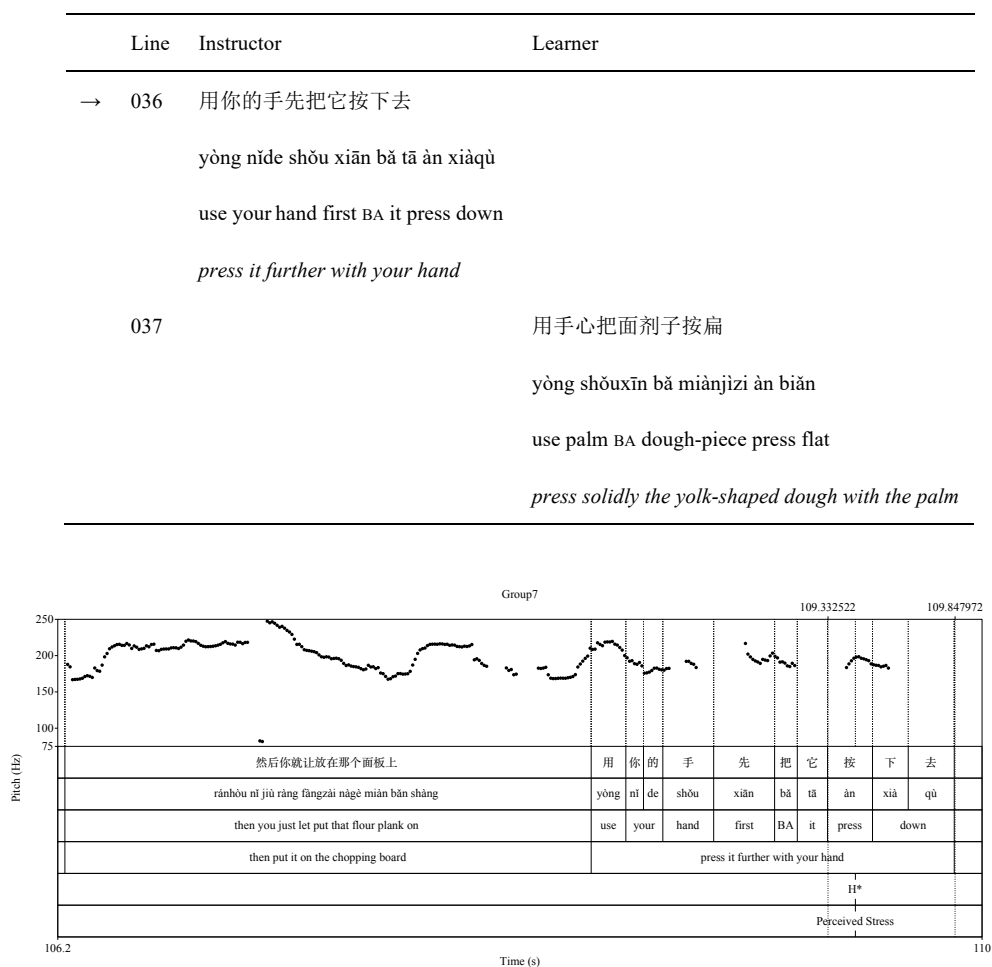


Figure 12  $F_0$  contour of line 036 in Excerpt 2



The analysis the verb 按 *àn* ‘press’ in the VP 按下去 *ànxìàqù* ‘press down’ within the BA-construction demonstrates that the F₀ contour shows a convex curve within the stressed syllable. The convex is identified as a canonical H*.

In contrast to the H*, the L* appears in another context; Excerpts 7 and 8 are such cases in point.

Excerpt 7 Group 8, line 051, 01:52.98-01:57.32. The arrow in line 051 indicates the example illustrated in

Figure 13.

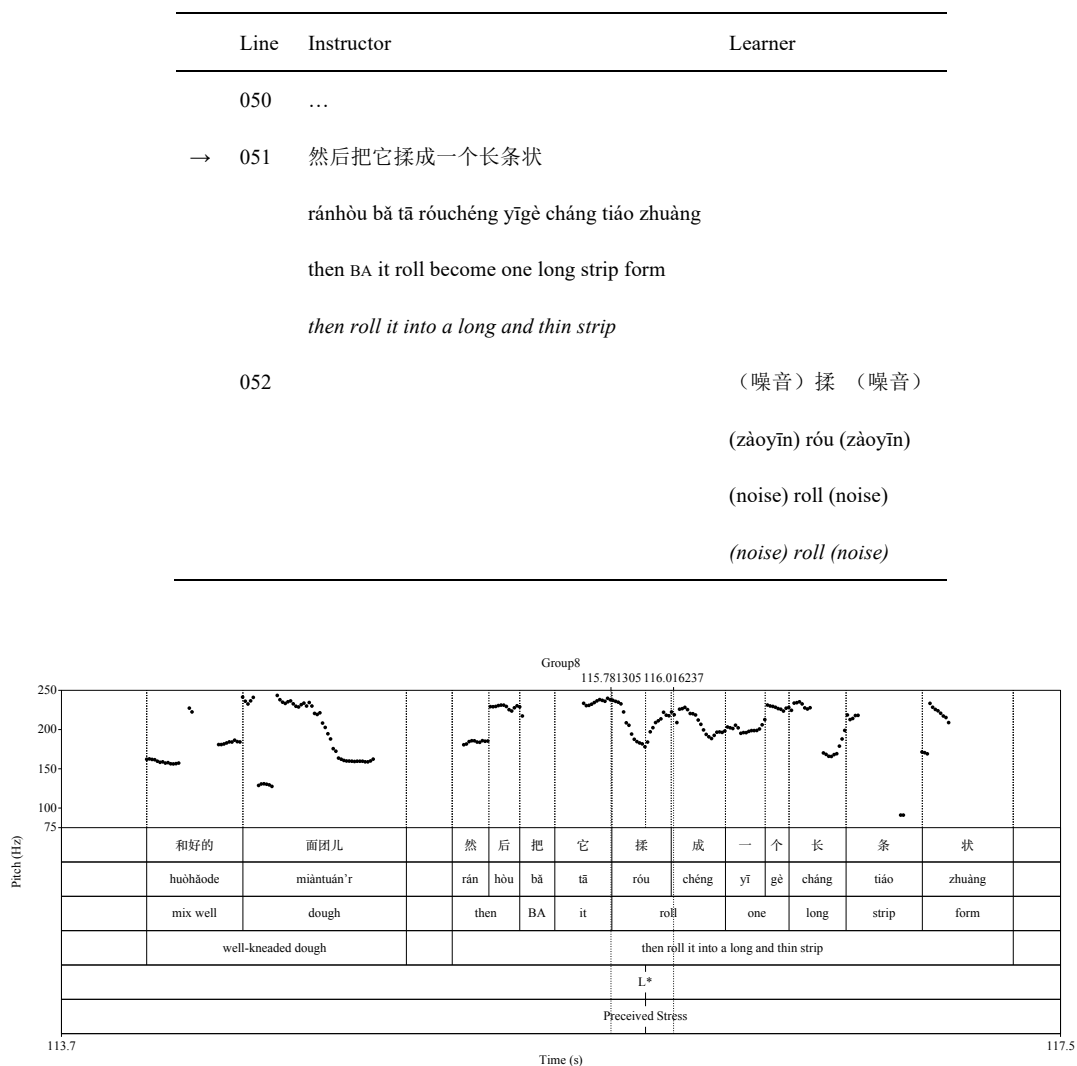


Figure 13 F₀ contour of line 051 in Excerpt 3

In excerpt 7, the instructor is telling the learner, in line 051, how to form a long and thin strip. Figure 13 shows that the perceived stress of the sentence of the BA-construction falls on the monosyllabic verb 揉 *róu* ‘roll’ (115.781-116.016) with the lexical tone LH within the VP 揉成 *róuchéng* ‘roll’. Phonetically, the F₀ contour of the lexical tone is realised as a local valley which is identified as a canonical L*. The L* can also be acoustically manifested when the verb is exposed as a lexical tone L as in Excerpt 8 and Figure 14.

Excerpt 8 Group 5, line 016, 00:48.25-00:51.54. The arrow in line 016 indicates the example illustrated in

Figure 14.

Line	Instructor	Learner
015	...	
→ 016	然后，接下来把面团放旁边醒一醒就可以 ránhòu, jiēxiàláí bǎmiàntuán fāngpángbiān xǐngyīxǐng jiùkěyǐ then, next BA dough place beside awake just okay well, then leave the dough aside to rest okay	
017		哦，把面团放旁边醒一下 o, bǎ miàntuán fāng pángbiān xǐng yīxià ohhh, BA dough put beside awake once ohhh, leave the dough aside to rest

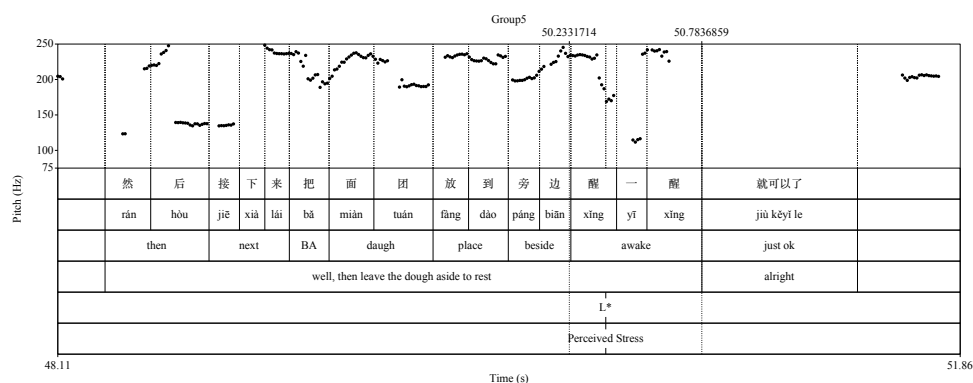


Figure 14 F₀ contour of line 016 in Excerpt 4

The utterance in the excerpt is produced by the instructor from Group 5 and here she suggests to her learner just to leave the dough aside for resting. As shown in the figure, the first 醒 *xǐng* ‘awake’ of the VP 醒一醒 *xǐngyīxǐng* ‘awake’ (50.233-50.783) is perceived as prominent. Acoustically, the pitch trace of the VP shapes a valley falling at the end of the first 醒 *xǐng* ‘awake’.

To sum up, this section has shown how perceived sentence stress is phonetically manifested when aligning with either BaP or V in the BA-construction. It has clearly outlined that the perceived sentence stress mainly falls on the verb in VP regardless of what kind of lexical tone the verb bears. The perceived sentence stress of the verbs (V) with their lexical tones (H, LH, L, and HL) tend to be acoustically realized as a phonetic prominence either a local peak (H*) or valley (L*) due to the tonal context. Furthermore, H and HL tend to be acoustically realized as H*, while LH and L seem to be phonetically realized as L*. The following section will present how gestural strokes are aligned with intonation in BA.

### 5.3. The alignment of gestural strokes and perceived stress in BA

This section illustrates how gestural strokes, intonation, and grammatical components are interacted. Table 8 presents the combined data.

Table 8 shows which grammatical elements are perceived as prominent in BA-constructions where there is also a gestural stroke. A Chi-Square test showed that gestures are most likely to occur with perceptually stressed Vs in the BA-construction with HL lexical tones ( $\chi^2 = (1, N = 29) = 23.14, p = .001$ ). Gestures occasionally occur with Ba-O with HL tone in BaP and also V with HL in VP. These patterns are exemplified below.

Excerpt 9 is taken from Group 2 where the female instructor (right) is teaching, in line 054, the learner how to mix the stuffing. How the gestural stroke, intonation, and the BA-construction align is shown in transcription (top) and filmstrip (bottom) in Figure 15.

Table 8 Distribution of strokes aligning with perceived sentence stress within BA-construction (BaP = Ba-Phrase, Ba = 把 *bǎ* ‘Ba’, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, H = high, LH = rising, L = dip, HL = falling)

Strokes aligned with Perceived sentence stress						
BA-construction		corresponded with lexical tones				Total
		H	LH	L	HL	
BaP	Ba			2		2
	Ba-O			1	4	5
VP	V		4		17	21
	V-O			1		1
Total			4	4	21	29

Excerpt 9 Group 2, Female, line 054, 04:15.42-04:21.90. The arrow in line 052 indicates the example illustrated in Figure 15.

Line	Introducer	Learner
→ 052	先把菜和肉按 1 比 1 的比例加到碗里 xiān bǎ cài hé ròu àn yībǐyīde bǐlì jiā dào wǎn lǐ first BA vegetable and meat follow one to one GEN scale add to bowl in <i>first, pour the vegetable and meat in the bowl as one to one proportion</i>	
053		1 比 1 的比例到进碗里 yī bǐ yī de bǐlì dào jìn wǎn lǐ one to one GEN scale add bowl in <i>pour them in the bowl as one to one proportion</i>

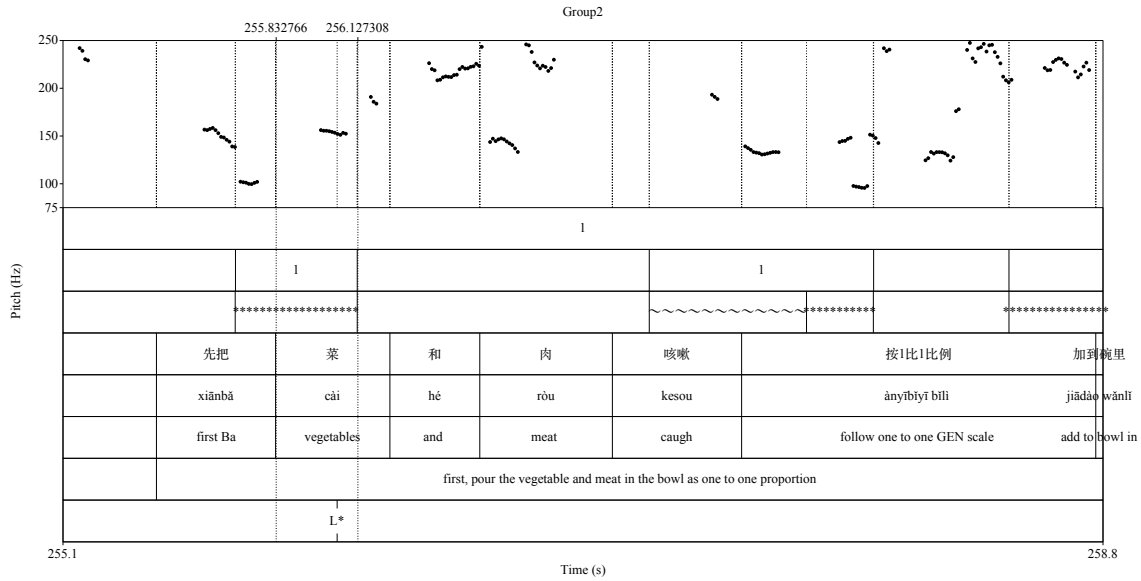


Figure 15 Gestural stroke (***) aligned with Ba-O as perceived sentence stress with tone HL of line 052 in Excerpt 8, transcription (top) from Praat and export filmstrip (bottom) from ELAN, respectively.

This single gestural stroke posed by her hands exactly occurs as the instructor is saying the Ba-O 菜 *cài* ‘vegetables’ with lexical tone HL within the BaP. The Praat picture shows the F0 contour of 菜 *cài* ‘vegetables’ is realized as a local low perceived as sentence stress. The filmstrip shows her separated finger hands with open palm down moving outward and inward, rapidly, once; and then the hands rest as shaped in a hold ready for the next gesture.

The next example extracted from Group 4 shows the interaction between these three modalities within the VP. The female instructor (left) is teaching, in line 069, the learner how to mix the stuffing as in Excerpt 10. How the gestural stroke, intonation, and the BA-construction interact is shown in transcription (top) and filmstrip (bottom) in Figure 16.

Excerpt 10 Group 4, Female, line 069, 03:11.85-03:13.70. The arrow in line 069 indicates the example illustrated in Figure 16.

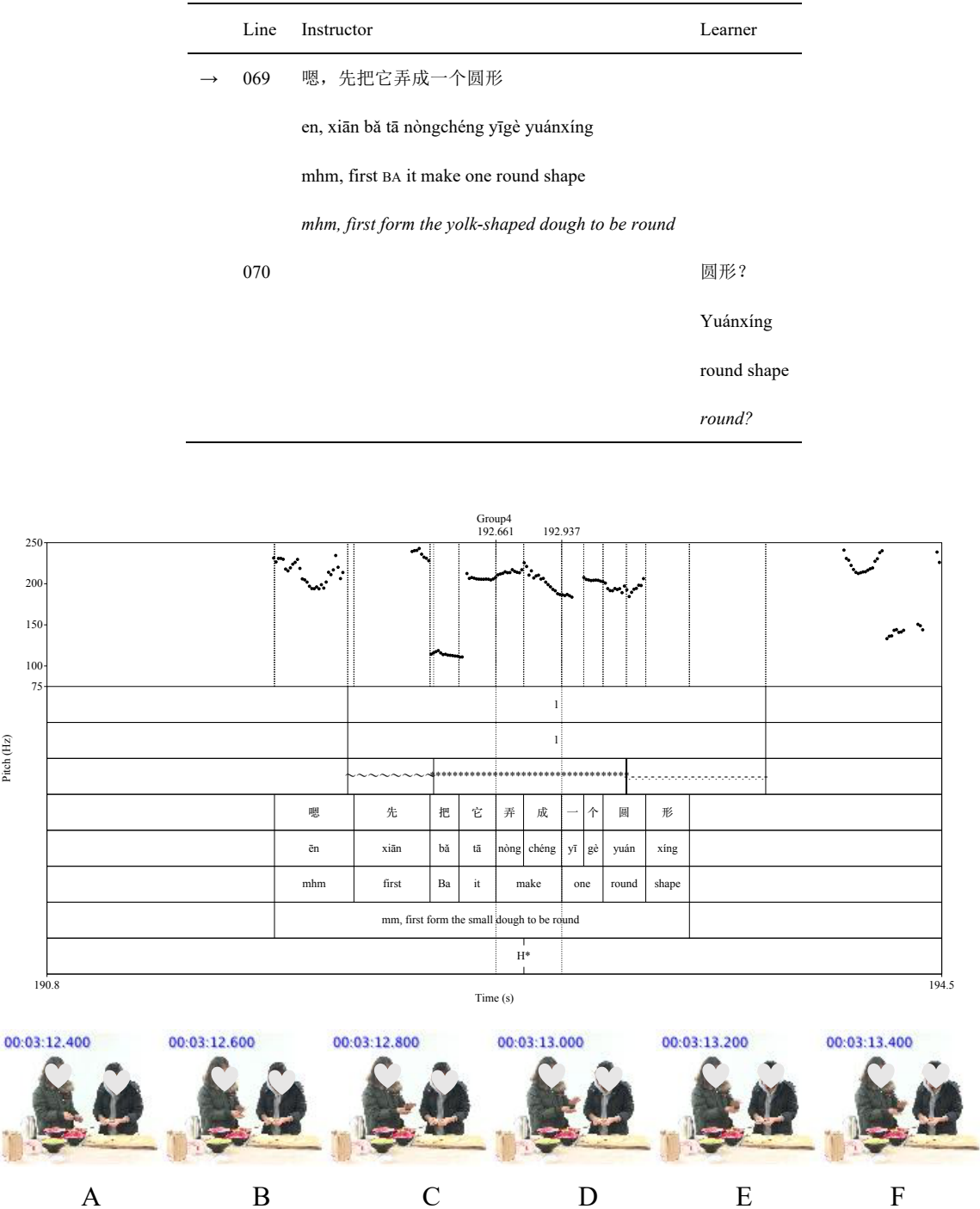


Figure 16 Gestural stroke (***) aligned with verb within VP as perceived sentence stress with HL of line 069 in Excerpt 10, transcription (top) from Praat and export filmstrip (bottom) from ELAN, respectively.

The gestural stroke covers both BA and VP. It starts precisely before 把 *bǎ* ‘Ba’ and persists through the Ba-O 它 *tā* ‘it’. At this moment, her hand palms come together as if grasping a ball. As she continually says ... 弄成一个圆形 ... *nòngchéng yīgè yuánxíng*... ‘...form a round pie’ her left-hand palm over the right one is moving away towards the learner and drawing back, quickly, once. When she commences the VP it is also constructed by a phonetic prominence within V and V-C 弄成 *nòngchéng* ‘form’, where 弄 *nòng* with lexical tone HL and 成 *chéng* LH are acoustically manifested as a local peak. This acoustic realization is also confirmed by the perceptual results.

In sum, we have seen how gestural strokes, intonation, and grammatical components interact in this section. That is to say that gestures are most likely to occur with perceptually stressed Vs in the BA-construction with HL lexical tones. So far, the BA-construction as one expressing one imperative form in Chinese discourses has been examined multimodally (i.e., syntax, intonation, and hand gestures) in order to explore how they interact in the discourses, which will be discussed in the next section.

## 6. General discussion, conclusions, and future study

The current study aimed to examine how grammatical components, intonation (perceived stress), and gesture are organized and interact in Chinese imperatives, specifically in natural spontaneous Chinese BA discourse in a face-to-face task-driven dyadic activity. The results of the current study revealed that hand movements do occur in BA. It is the VP that bears sentence stress of BA. It also shows that hand movements and the sentence stress are temporally synchronized with the VP.

The alignment between hand gestures and grammatical components of BA vary concerning the relationship between these two modalities. Overall, gestures most often align with both BaP and VP. This reflects that “gesture expression is a fully integrated component of the utterance’s construction”, as Kendon (2004:125) points out. More than one third ( $n = 17$ , or 38%) of the gestures also align with BaP only. This might be

interpreted as a mutual adaptation of gesture and speech (Kendon, 2004:135). For example, there is a pause of 0.3 seconds inserted between two 把 *bǎ* ‘Ba’ produced by the female instructor in Group 2 and both are accompanied by two clear gestural strokes, as in Figure 9. In other words, her utterance is stopped in order to complete the gestural stroke. Another example of mutual adaptation of these two modalities is the case of Excerpt 1 as in Figure 7. Here we can see that gesture continues when a pause is inserted in the spontaneous utterance produced by the male instructor. Gestural strokes only very rarely occur with the VP. Could this suggest that prosody might have an advantage over gesture since the verb ( $n = 54$ , or 76%) in VP bears the sentence stress of the BA? Note also that only an auditory perception task was conducted – visual perception is still perception.

Turning to stress, it seems that BaP is not new information generally since it is marked as a definite/generic object (Li & Thompson, 1989:465). Still, there are two cases where only 把 is stressed, which might be interpreted as the instructors starting the BaP while seeking the stuffing. This might have resulted in the Ba marker was unconsciously being produced as longer or heavier. Ba-O in BaP can also bear the stress (see Table 6, as in § 5.2). This might probably be due to the coincidence of the object exposed with HL, for instance, 面 *miàn* ‘flour’, 菜 *cài* ‘vegetable’, 肉 *ròu* ‘meat’, 馅 *xiàn* ‘stuffing’. Interestingly, it therefore seems to be reasonable that the verb bears the stress since most of the monosyllabic verbs produced by the participants bear HL, i.e., falling tone, phonetically manifested as H* ( $n = 30$ , or 68%, as in Table 7). There are 9 exceptions (3 LH are acoustically realized H*, 5 HL as L*, and 1 L as H*), which don’t correspond with the result that that H and HL lexical categories, in spontaneous speech, will be acoustically realised as exclusive prominent H* vs. LH and L as L*. Statistically, this can be explained that the expected count might be very low but deviations will exist always in reality. One might presume that the choice of the verbs had been influenced. However, no manuscript was used at the time; one pre-recorded mute demo was shown before the task started. It is a remarkable coincidence that 68% of the verbs bear the HL occasionally, which might be a plausible reason for why HL is overrepresented. However, no Ba-O bearing H is stressed, for example, 把它 *bǎtā* ‘it’ in 把它弄成一个圆形 *bǎtā nòngchéng yīgè yuán* ‘form it a round pie’ as in Excerpt 10. There are two possible reasons: one is that BaP is not stressed



due to it representing old information; another is that the very short duration of 它 *tā* concerning tonal sequences hinders it from reaching an actual high after 把 *bǎ* with L since it has to prepare for seamless connection to the forthcoming 弄 *nòng* ‘make’ with HL (see Figure 16). Beyond this, it is more illustrative that there is a close interaction between monosyllabic tones and local intonation compared to topic-comment and copula-complement (Li, 2014) which is not the focus of the current study.

Finally, concerning a possible interaction between grammatical components, prosodic units, and gestural strokes, the results reveal that VP shows a higher frequency of alignments with gestures and sentence stress than does BaP. Consider now again the last Excerpt as in Figure 16; note that the stroke has begun and is ongoing through the BA, differing from the earlier studies mentioned in Kendon (2004:125) where their strokes were “always completed either before or at the same time as the tonic syllable of the co-occurrent tone unit”. This difference raises the question cross-linguistically whether gestural strokes vary along with different intonational patterns. Additionally, this instructor’s hand palms merge together as if grasping a ball, which might be semantically prompted by the meaning of the verb 弄 *nòng* ‘form’. The monosyllabic tone was perceived as an H* with regard to sentence stress due to the interaction between tone and intonation. This local intonation, accompanied by a clear hand gesture, may also convey a new idea – 圓 *yuán* ‘be round’, as in transcript of Excerpt 10. This new idea does not seem to be understood by the learner, which is clearly reflected by the question 圓形 *yuánxíng* ‘round’ in her turn-taking, as in line 070 in Excerpt 10. This finding is supported by Kendon’s (2004:125) view that “the stroke of the gesture phrase, the tonic syllable, and the information centre of the speech phrase, [are] all co-occurrent”. One might also question if the stroke will fall on a verb with a very salient meaning when the role of lexical tones is irrelevant since HL is dominant. This perspective has not been taken into account in the current study.

In sum, the results suggest that there is in fact an interaction between syntax, prosody, and gesture in Chinese BA imperatives. Previous studies of the Chinese imperative have treated the three modalities separately (e.g., Fan, 1998; Po-Ching & Rimmington, 2015; Shi & Jiao, 2016). Aikhenvald (2010) mentioned the role of gesture but there was no

empirical study to support the claims. The current study is therefore the first to consider all three aspects and to show that there is an interaction among the three modalities.

There might be innumerable other research topics one could investigate in the arena of multimodality. The analysis has only paid attention to the alignment of the complete BA-construction and gesture; single grammatical components (e.g., 把 *bǎ* ‘Ba’ or Ba-O in BaP, V or V-O in VP) aligned with gestures have been ignored. In particular, when the marker 把 *bǎ* ‘Ba’ in spontaneous speech combines with a discourse particle like 然后 *ránhòu* ‘then’, as in Excerpt 7, or with an adverb 先 *xiān* ‘first’, as in Excerpt 9, they form so-called prosodic words based on acoustic criteria. For example, all these four syllables are compressed under one and the same tone contour envelop. In other words, the interface between these three modalities is a not one-to-one correspondence. What kind of role do the prosodic words play in spontaneous discourse when they are aligned with gestures? Even though this study has discussed about the integrity of the gesture-speech ensemble, no systematic investigation has yet been done concerning BA, which might be considered in a future study.

Methodologically, a post-interview with the participants in a production task could have shed light on why their gestures were performed in a certain way, which might help us to obtain more details about how the information is decoded. The reason for this is that we have seen that not all components in BaP are non-stressed, which can imply that BaP could be stressed and could convey a new idea. In terms of the perception task, the future the participants will be asked which syntactic component is perceived as stressed and which character will be stressed in the sentence as soon as the heavy part is identified. Also, the four tonal categories of the objects used in the tasks will be carefully considered.

Finally, only hand movements have been investigated in the current study and BA has been studied only as an imperative. Other gestural articulators, such as head movement, could be considered, and BA outside of the imperative could also be investigated. Last but not least, we have noticed that the verb in BA cannot exist alone and must be combined with a complement (see §2.1). How these Chinese verb-particles indicating motion events are constructed has still not been touched upon in terms of multimodality. For example, the instructor is requiring the learner 用你的手先把它按下去 *yòng nǐde shǒu xiān bǎtā àn*

*xiàqù* ‘press it further with your hand’, as in line 036 in Excerpt 6 (Here ‘it’ refers to the dough). The learner gets a suggestion to perform a motion, 按 *àn* ‘press’ so that she will cause the dough in a placement or situation by path 下 *xià* ‘down’ and direction 去 *qù* ‘away’. It would be worth exploring how these Chinese verb-particles are encoded and decoded since this verb category differs cross-linguistically, especially comparing to placement and removal events concerning Swedish phrasal verbs (Gullberg & Burenhult, 2012). This might also help us to understand how old or new information interact and are processed between speakers and listeners in communication as a repertoire rather than as segments like syntax, pragmatics, and prosody studied separately.

# Bibliography

- Aikhenvald, A. Y. (2010). *Imperatives and commands*. Oxford University Press.
- Atkinson, J. M., & Heritage, J. (1999). Jefferson's transcript notation. In A. Jaworski & N. Coupland (Eds.), *The discourse reader* (pp. 158-166). Routledge.
- Chen, H., Gu, Y., & Zuo, H. (2013). Shǒushì yǔ yányǔde guānxìlùn - Jiāntán duìrènzhī yǔyánxuéde qǐshì *Wànguóyǔ*5, 50-58. [Theories on the relationship between gesture and speech - Also on implications for cognitive linguistics *Journal of Foreign Languages*5, 50-58] . 陈桦, 顾琰, 左瀚, 2013, 手势与言语的关系论——兼谈对认知语言学的启示 *外国语* 5, 50-58.
- Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge University Press.
- Cheng, H.-T. J. (2011). Argument ellipsis in Mandarin Chinese. Proceedings of the 23rd North American Conference on Chinese Linguistics (NACCL-23),
- Comrie, B., Haspelmath, M., & Bickel, B. (2008). *The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses* <https://doi.org/https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>
- Debreslioska, S., & Gullberg, M. (2019). Discourse reference is bimodal: how information status in speech interacts with presence and viewpoint of gestures. *Discourse Processes*, 56(1), 41-60.
- Dixon, R. M. (2010). *Basic linguistic theory volume 2: Grammatical topics* (Vol. 2). Oxford University Press on Demand.
- Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford University Press
- Fan, X. (1998). *Hànyǔjùzide lèixíng Shūhǎi chūbǎnshè* [Types of Chinese sentences Shuhai Publishing House] . 范晓, 1998, 汉语句子的类型. 书海出版社
- Feng, S. (2001). Prosodically constrained bare-verb in ba constructions. *Journal of Chinese Linguistics*, 29(2), 243-280.
- Feng, S. (1996). Lùn hànyǔde "yùnlǜcí" *Zhōngguó shèhuìkēxué*1, 161-176. [On Chinese prosodic words *Social Sciences in China Press*1, 161-176.] . 冯胜利, 1996, 论汉语的“韵律词” *中国社会科学* 1, 161-176.
- Gee, J. P. (2011). *An introduction to discourse analysis: Theory and method* (3rd ed.). Routledge.

- Gullberg, M., & Burenhult, N. (2012). Probing the linguistic encoding of placement and removal events in Swedish. In A. Kopecka & B. Narasimhan (Eds.), *Events of Putting and Taking* (pp. 167-182). John Benjamins Publishing Company.
- Halliday, M. A. K., & McDonald, E. (2004). Metafunctional profile of the grammar of Chinese. In C. Alice, M. James Robert, & M. Christian MIM (Eds.), *Language Typology: A Functional Perspective* (pp. 305-396). John Benjamins Publishing Company.
- Huang, C.-T. J., Li, Y.-h. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press.
- Jin, S. (1996). *An acoustic study of sentence stress in Mandarin Chinese* [Doctoral dissertation, The Ohio State University].
- Jin, S. (1992). Běijīnghuàde yǔqì hé yǔdiào *Zhōngguó yǔwén*2, 113-123. [Mood and intonation of Beijing Mandarin Chinese2, 113-123.] . 劲松, 1992, 北京话的语气和语调 *中国语文* 2, 113-123.
- Jing-Schmidt, Z. (2005). *Dramatized Discourse: The Mandarin Chinese «BA»-Construction*. John Benjamins Publishing Company.
- Jun, S.-A., & Fletcher, J. (2014). Methodology of studying intonation: From data collection to data analysis. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 493-519). Oxford University Press.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Knapp, M. L., Hall, J. A., & Horgan, T. G. (2014). *Nonverbal communication in human interaction* (8th ed.). Cengage Learning.
- Kroeger, P. R. (2005). *Analyzing grammar: An introduction*. Cambridge University Press.
- Lee-Wong, S. M. (1994). Imperatives in requests: Direct or impolite—observations from Chinese. *Pragmatics*, 4(4), 491-515.
- Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse processes*, 15(3), 277-301.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, X. (2014). *Multimodality, interaction and turn-taking in Mandarin conversation*. John Benjamins Publishing Company.
- Lin, M., Yan, J., & Sun, G. (1984). Běijīnghuà liǎngzìzǔ zhèngchángzhòngyīnde chūbùshíyàn *Fānyán*1, 57-73. [Experiment on stress of Mandarin disyllabic words *Dialect*1, 57-73.] . 林茂灿, 颜景助, 孙国华, 1984, 北京话两字组正常重音的初步实验 *方言* 1, 57-73.

- Lu, J. (2006). Hànyǔ mìnglǐngjùde yǔdiào tèdiǎn. Dìqījiè zhōngguó yǔyīnxué xuéshù huìyì jì yǔyīnxué qiányán wèntí guójì lùntán lùnwénjí [Properties of intonation in Chinese imperatives in Proceeding of the 7th Chinese Academic Conference on Phonetics and International Forum on The Frontiers of Phonetics] . 路继伦, 2006, 汉语命令句的语调特点. 第七届中国语音学学术会议暨语音学前沿问题国际论坛论文集
- Lu, J., & Sun, J. (2008). Hànyǔ mìnglǐngjù yīngāo, shíchángyǔ yīnxì móshì. Dìbājiè zhōngguó yǔyīnxué xuéshù huìyì jì qīnghè Wú Zōngjì xiānshēng bǎisui huádàn yǔyīnxué qiányán wèntí guójì lùntán lùnwénjí [On pitch, duration, and phonological pattern of Chinese imperative sentences in Proceedings of the 8th Chinese Phonetics Academic Conference and The International Symposium Celebrating Mr. Wu, Zongji's centenary on the Frontiers of Phonetics] . 路继伦, 孙佳, 2008, 汉语命令句音高, 时长与音系模式. 第八届中国语音学学术会议暨庆贺吴宗济先生百岁华诞语音科学前沿问题国际研讨会论文集.
- Ma, Q. (1995). Lùn Hànyǔ qǐshǐjùde tèzhēng wèntí *Yǔyán yánjiū* 1, 44-51. [On Chinese imperative properties *Language Research* 1, 44-51.] . 马清华, 1995, 论汉语祈使句的特征问题 *语言研究* 1, 44-51.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.
- Ochs, E. (1979). Transcription as theory. *Developmental pragmatics*, 10(1), 43-72.
- Osborne, T., & Liang, J. (2015). A survey of ellipsis in Chinese. Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015),
- Pierrehumbert, J., & Beckman, M. (1988). Japanese tone structure. *Linguistic inquiry monographs*(15), 1-282.
- Po-Ching, Y., & Rimmington, D. (2015). *Chinese: A comprehensive grammar*. Routledge.
- Shen, J. (1994). Hànyǔ yǔdiào gòuzào hé yǔdiào lèixíng *Fāngyán* 3, 221-228 [Intonation construction and types in Chinese *Dialect* 3, 221-228] . 沈炯, 1994, 汉语语调构造和语调类型 *方言* 3, 221-228.
- Shi, F., & Jiao, X. (2016). Pǔtōnghuà mìnglǐngyǔdiàode shícháng hé yīliàngfēnxī *Hànyǔxuéxí*42(1), 93-108. [The Duration and Intensity of Commanding Sentence Intonation of Mandarin Chinese *Journal of Chinese Study*] . 石锋, 焦雪芬, 2016, 普通话命令语调的时长和音量分析 *汉语学习* 42(1), 93-108.

- Shi, P. (1980). Sìzhǒng jùzide yǔdiào biànhuà *Yǔyánjiàoxué yǔ yánjiū* 2, 71-81. [Intonational changes of four types sentences *Language teaching and researching* 2, 71-81.] . 石佩雯, 1980, 四种句子的语调变化 *语言教学与研究* 2, 71-81.
- Sun, J. (2008). *Hànyǔ mìnglǐngjù yǔdiàode shíyàn yánjiū* Shuòshìlùnwén, Tiānjīn Shīfàn Dàxué [Experimental intonation of Chinese imperatives Master dissertation, Tianjin Normal University]. 孙佳, 2008, 汉语命令句语调的实验研究 硕士学位论文, 天津师范大学.
- Tao, H. (1996). *Units in Mandarin conversation: Prosody, discourse, and grammar*. John Benjamins Publishing Company.
- Teng, W.-H. (2016). *Yufa! A practical guide to Mandarin Chinese grammar*. Taylor & Francis.
- Thompson, S. A. (1973). Resultative verb compounds in Mandarin Chinese: A case for lexical rules. *Language*, 361-379.
- Velupillai, V. (2012). *An introduction to linguistic typology*. John Benjamins Publishing Company.
- Wang, L., Lu, J., & Fu, H. (2004). *Xiàndài Hànyǔ*. Shāngwù yìnshūguǎn [Modern Chinese Commercial Press] . 王理嘉, 陆俭明, 符淮青, 2004, 现代汉语. 商务印书馆
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of phonetics*, 27(1), 55-105.
- Yan, J., & Lin, M. (1988). Běijīnghuà sānzìzǔ zhòngyīnde shēngxué biǎoxiàn *Fāngyán* 3, 227-237. [Stress acoustic realisation of Beijing Mandarin trisyllabic words *Dialect* 3, 227-237.] . 颜景助, 林茂灿, 1988, 北京话三字组重音的声学表现 *方言* 3, 227-237.
- Yin, X. (2013). *Xiàndài Hànyǔ qǐshǐjù fànchóu yánjiū* Bóshì lùnwén, Fùdàn Dàxué. [Research on category of the imperatives in modern Chinese Doctoral dissertation, Fudan University]. 尹相熙, 2013, 现代汉语祈使范畴研究 博士学位论文, 复旦大学.
- Zhou, X. (2006). *Hànyǔ mìnglǐngjù yǔdiào yánjiū* Shuòshì lùnwén, Tiānjīn Shīfàn Dàxué [Research on intonation of Chinese imperatives Master dissertation, Tianjin Normal University]. 周昕, 2006, 汉语命令句语调研究 硕士学位论文, 天津师范大学.

# Appendix I Consent Form



## CENTRE FOR LANGUAGES & LITERATURE LUND UNIVERSITY

PO BOX 201, 221 00 Lund, Sweden

### Consent Form

I hereby give my permission to Guohua HU (胡国华), Centre for Languages and Literature, Lund University, Sweden, to use today's recordings (audio and video) for the following purposes:

(Please tick the appropriate box, "☐", if you give your permission.)

☐ 1. analyses for scientific research;

2. as illustrations of the above scientific research in professional seminars, lectures, conferences, and in scientific publications;

☐ as still photographs;

☐ as video clips.

My anonymity is guaranteed. Under no circumstances will my personal identity be revealed to anybody other than the above mentioned scientific researchers (e.g. no names will be used in presentations of the recording).

Name

Signature

Date

Participant number: _____



# Appendix II Answer Sheet

This answer sheet (in Chinese) was used in the perception task. English translation is added here. The participants were asked to only encircle the character of the most prominent placement in the given discourse lines produced by the instructor.

日期: _____ (Date: _____) 编号: _____ (Nr. _____)

话轮	教者	学者
Line	Instructor	Learner
001	一起来包饺子	
002	然后我们首先第一步就是	
003	嗯, 要把面倒入倒	
004	面粉倒入碗里面	
005	唉, 不能用手抓	
006	我刚刚说的是倒对	
007		哦, 把面倒入碗中
008	对, 然后再接下来, 嗯, 加,	
009	你加点水	
010	好, 你来加	
011	好! 可以太多吗, 可以... 太多吗?	
012		不可以
013	对, 然后, 接下来	
014	我们来和面	
015	和面怎么和	
016	不行, 不可以这样和	
017	我们要用手	
018	这样顺时针...	
019	一起和	
020	知道了吗	

Appendix III

The perception results of perceived sentence stress (H = hight, LH = low high, L = low, HL = high low; BaP = Ba-Phrase, Ba = the marker of 把 bǎ, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, VP = Verb Phrase, V = Verb, V-O = Verb-Object, V-O-M = Verb-Object-Modifier, V-C = Verb-Compliment; DP = Discourse Particle, Adv = Adverb). Note that all participants have not responded to all stimuli.

Nr.	Group	Lines	Instructor	Syllable	Tone	Glossing								DP/Adv	Total
						BaP				VP					
						Ba	Ba-O	Ba-O-M	V	V-O	V-O-M	V-C			
1.	1	3	嗯，要把面倒入	倒 <i>dào</i> ‘pour’	HL		7		23	5				35	
2.	1	88	然后我们要把它做成皮儿	做 <i>zuò</i> ‘make’	HL	1	4	1	20	7		1	1	35	
3.	1	97	然后把我说过的话重复一遍	重 <i>chóng</i> ‘again’	LH	1	1		17	2				21	
4.	2	7	然后先把	把 <i>bǎ</i> ‘Ba’	L	20			13	2				35	
5.	2	8	把面倒到碗里	倒 <i>dào</i> ‘pour’	HL		1		18	1				20	
6.	2	19	先揉，把面揉一揉	揉 <i>róu</i> ‘roll’	LH	4	10		20					34	
7.	2	25	把这个面剂子，	面 <i>miàn</i> ‘flour’	HL	2	12	2	9	2				27	
8.	2	43	然后我们把饺子皮儿对折	对 <i>duì</i> ‘opposite’	HL	2	2		27	4				35	
9.	2	49	然后把饺子给按紧	按 <i>àn</i> ‘press’	HL		2		31		1			34	
10.	2	50	把皮按紧了	按 <i>àn</i> ‘press’	HL		1		28	3				32	
11.	2	52	先把菜和肉按<SIL>1 比 1 的比例加到碗里	菜 <i>cài</i> ‘vegetable’	HL		18		5				2	25	
12.	3	4	把那个和好的面切成剂子	和 <i>huò</i> ‘mix’	HL		13		7					29	
	3	4	把那个和好的面切成剂子	剂 <i>jì</i> ‘daugh’	HL				9						
13.	3	9	...把面倒入碗中	倒 <i>dào</i> ‘pour’	HL		10		20			5		35	
14.	3	47	先把面搓好	搓 <i>cuō</i> ‘roll’	H				30				1	35	
15.	3	52	就把不平的地方揉一下	揉 <i>róu</i> ‘roll’	LH		4		25				3	32	
16.	3	63	先把它按压一下	按 <i>àn</i> ‘press’	HL		3		27				1	31	
17.	3	68	把它拿起来	拿 <i>ná</i> ‘take’	LH		2		27			2		31	
18.	3	100	...把这个也按上	按 <i>àn</i> ‘press’	HL		10		25					35	
19.	3	104	...把那个收一个尾儿...	收 <i>shōu</i> ‘fold’	H		3		25					28	
20.	4	9	先把这个面呢，	面 <i>miàn</i> ‘flour’	HL		21	10					4	35	

21.	4	26	把面先放在这	放 <i>fàng</i> ‘place’	HL	3	32			35
22.	4	69	先把它弄成一个圆形	弄 <i>nòng</i> ‘make’	HL		31	2	2	35
23.	4	83	...把这个转一下	转 <i>zhuàn</i> ‘turn’	HL		5 30			35
24.	4	98	...把那个面片对折	对 <i>duì</i> ‘opposite’	HL		27	8		35
25.	5	5	...把面粉倒入碗中	倒 <i>dào</i> ‘pour’	HL	10	21	4		35
26.	5	16	...把面团放旁边醒一醒	醒 <i>xǐng</i> awake	L	5	26	4		35
27.	5	22	把这些							
28.	5	23	把配料放到碗里	放 <i>fàng</i> ‘place’	HL	4	31			35
29.	5	36	我们要把面团搓成长条状的捻子	搓 <i>cuō</i> ‘roll’	H	7	21	3		32
	5	36	我们要把面团搓成长条状的捻子	捻 <i>niǎn</i> ‘wick’	L			2		
30.	5	46	然后剂子做好了之后，把它拿在手上	拿 <i>ná</i> ‘take’	LH	5	21	4		31
31.	5	52	<SIL> 你要把它	把 <i>bǎ</i> ‘BA’	L	1				8
32.	5	68	就把面皮放在手上	放 <i>fàng</i> ‘place’	HL	7	22		3	32
33.	5	69	把馅放在<SIL>皮中间	放 <i>fàng</i> ‘place’	HL	10	18	3		31
34.	6	7	...把那个面粉放在这个碗里头	放 <i>fàng</i> ‘place’	HL	3	31			34
35.	6	12	再把水倒进去	倒 <i>dào</i> ‘pour’	HL		31			31
36.	6	22	然后你先把菜放在那个碗里	菜 <i>cài</i> ‘vegetable’	HL	31	4			35
37.	6	25	然后你，你再把这些，根据你自己想要的口味	想 <i>xiǎng</i> ‘will’	L	31			3	34
38.	6	29	然后现在把这个<SIL>馅儿和均匀	和 <i>huò</i> ‘mix’	HL	5	27			32
39.	6	33	...把这个菜放在一边	放 <i>fàng</i> ‘place’	HL	4	31			35
40.	6	39	...要把它弄成一个，就是	弄 <i>nòng</i> ‘make’	HL	2	33			35
41.	6	42	把它弄圆	弄 <i>nòng</i> ‘make’	HL		35			35
42.	6	50	右手把它像这样	它 <i>tā</i> ‘it’	LH	31				31
43.	6	51	每个都要把它弄均匀	弄 <i>nòng</i> ‘make’	HL	2	31	2		35
44.	6	59	你要把它放到手上	放 <i>fàng</i> ‘place’	HL	3	30	2		35
45.	6	64	你就把它放在面板上	放 <i>fàng</i> ‘place’	HL	3	31			34
46.	6	87	对，你把它放在<SIL>左手	放 <i>fàng</i> ‘place’	HL	5	21		3	29
47.	6	90	对，把馅儿放在这个饺子皮中间	馅 <i>xiàn</i> ‘stuffing’	HL	21	8		2	31
48.	6	93	你把这个饺子中间对折	对 <i>duì</i> ‘opposite’	HL	2	8 21			31
49.	6	97	再把左边右边都捏紧	捏 <i>niē</i> ‘pinch’	H		31		4	35
50.	7	6	就是先把面倒到碗里	倒 <i>dào</i> ‘pour’	HL		31			31

51.	7	22	我们先把那个和好的面团	和 <i>huò</i> ‘mix’	HL	21	6			4	31
52.	7	23	把它揉成一个长条形	揉 <i>róu</i> ‘roll’	LH	10	21		4		35
53.	7	25	就是用手把它揉成长的	揉 <i>róu</i> ‘roll’	LH	3	27		5		35
54.	7	32	你把它搓好后								
55.	7	33	就可以把它								
56.	7	37	你把它揪得再均匀一点	揪 <i>jiū</i> ‘pinch’	H	7	21		3	4	35
57.	7	41	把那个揪下来的那个剂子给它搓	剂 <i>jì</i> ‘daugh’	HL	21	3	6			30
58.	7	43	...把它按下去	按 <i>àn</i> ‘press’	HL		31				31
59.	7	52	就把那个边擀得均匀一点	擀 <i>gǎn</i> ‘roll’	L		2	31			33
60.	7	56	然后右手把馅放上去	放 <i>fàng</i> ‘place’	HL		31				31
61.	7	61	然后就要用手把这个皮儿对折一下	对 <i>duì</i> ‘opposite’	HL		11	21			32
62.	7	66	把两边都								
63.	7	71	就是要把菜和肉混在一起	混 <i>hùn</i> ‘mix’	HL		31			4	35
64.	7	87	等你把它搅匀了	搅 <i>jiǎo</i> ‘mix’	L	2	31				33
65.	8	10	把面粉倒在一	倒 <i>dào</i> ‘pour’	HL		31				31
66.	8	32	我们可以自己把								
67.	8	33	把和好的面先放在一边儿	和 <i>huò</i> ‘mix’	HL	21	9			1	32
68.	8	48	我们把和好的馅也先放在一边	和 <i>huò</i> ‘mix’	HL	19	8			4	31
69.	8	51	然后把它揉成一个长条状	揉 <i>róu</i> ‘roll’	LH	7	21		2	5	35
70.	8	64	然后把它拿起来	拿 <i>ná</i> ‘take’	LH	2	31				33
71.	8	74	把这个面放下	放 <i>fàng</i> ‘place’	HL	4	31				35
72.	8	89	把它放在面板上	放 <i>fàng</i> ‘place’	HL	4	28		3		35
73.	8	107	然后把皮儿	皮 <i>pí</i> ‘skin’	LH	31					31
74.	8	131	把它捏紧	捏 <i>niē</i> ‘pinch’	H	4	31				35

Appendix IV

The perception results of perceived sentence stress (H = hight, LH = low high, L = low, HL = high low; Ba = the marker of 把 bǎ, Ba-O = Ba-Object, Ba-O-M = Ba-Object-Modifier, V = Verb, V-O = Verb-Object; H* = local prominent high, L* = local prominent low)

Nr.	Group	Lines	Instructor	Syllable	Alignment of									
					Stroke and grammar	Intonation and grammar								Stroke, grammar, and perceived stress
						Perceived stress						Acoustics		
						Tone	Ba	Ba-O	Ba-O-M	V	V-O	H*	L*	
1.	1	3	嗯，要把面倒入	倒 <i>dào</i> ‘pour’	1	HL					1		1	
2.	1	88	然后我们要把它做成皮儿	做 <i>zuò</i> ‘make’		HL					1		1	
3.	1	97	然后把我说过的话重复一遍	重 <i>chóng</i> ‘again’		LH					1			1
4.	2	7	然后先把	把 <i>bǎ</i> ‘Ba’		L	1							1
5.	2	8	把面倒到碗里	倒 <i>dào</i> ‘pour’		HL					1		1	
6.	2	19	先揉，把面揉一揉	揉 <i>róu</i> ‘roll’	1	LH					1		1	1
7.	2	25	把这个面剂子，	面 <i>miàn</i> ‘flour’		HL		1						
8.	2	43	然后我们把饺子皮儿对折	对 <i>duì</i> ‘opposite’		HL					1		1	
9.	2	49	然后把饺子给按紧	按 <i>àn</i> ‘press’		HL					1		1	
10.	2	50	把皮按紧了	按 <i>àn</i> ‘press’		HL					1		1	
11.	2	52	先把菜和肉按<SIL>1 比 2 的比例加到碗里	菜 <i>cài</i> ‘vegetable’	1	HL		1						1
12.	3	4	把那个和好的面切成剂子	和 <i>huò</i> ‘mix’		HL		1						1
	3	4	把那个和好的面切成剂子	剂 <i>jì</i> ‘daugh’		HL								
13.	3	9	...把面倒入碗中	倒 <i>dào</i> ‘pour’		HL					1		1	
14.	3	47	先把面搓好	搓 <i>cuō</i> ‘roll’		H					1		1	
15.	3	52	就把不平的地方揉一下	揉 <i>róu</i> ‘roll’		LH					1			1
16.	3	63	先把它按压一下	按 <i>àn</i> ‘press’		HL					1		1	1
17.	3	68	把它拿起来	拿 <i>ná</i> ‘take’		LH					1		1	1
18.	3	100	...把这个也按上	按 <i>àn</i> ‘press’		HL					1		1	

19.	3	104	...把那个收一个尾儿...	收 <i>shōu</i> ‘fold’		H		1	1	
20.	4	9	先把这个面呢，	面 <i>miàn</i> ‘flour’		HL	1			
21.	4	26	把面先放在这	放 <i>fàng</i> ‘place’		HL		1	1	
22.	4	69	先把它弄成一个圆形	弄 <i>nòng</i> ‘make’		HL		1	1	1
23.	4	83	...把这个转一下	转 <i>zhuàn</i> ‘turn’	1	HL		1	1	1
24.	4	98	...把那个面片对折	对 <i>duì</i> ‘opposite’	1	HL		1	1	1
25.	5	5	...把面粉倒入碗中	倒 <i>dào</i> ‘pour’		HL		1	1	
26.	5	16	...把面团放旁边醒一醒	醒 <i>xǐng</i> awake		L		1	1	
27.	5	22	把这些							
28.	5	23	把配料放到碗里	放 <i>fàng</i> ‘place’		HL		1	1	
29.	5	36	我们要把面团搓成长条状的捻子	搓 <i>cuō</i> ‘roll’		H		1	1	1
	5	36	我们要把面团搓成长条状的捻子	捻 <i>niǎn</i> ‘wick’	1	L				1
30.	5	46	然后剂子做好了之后，把它拿在手上	拿 <i>ná</i> ‘take’		LH		1	1	
31.	5	52	<SIL> 你要把它	把 <i>bǎ</i> ‘BA’		L	1			1
32.	5	68	就把面皮放在手上	放 <i>fàng</i> ‘place’	1	HL		1	1	1
33.	5	69	把馅放在<SIL>皮中间	放 <i>fàng</i> ‘place’		HL		1	1	1
34.	6	7	...把那个面粉放在这个碗里头	放 <i>fàng</i> ‘place’		HL		1	1	1
35.	6	12	再把水倒进去	倒 <i>dào</i> ‘pour’		HL		1	1	
36.	6	22	然后你先把菜放在那个碗里	菜 <i>cài</i> ‘vegetable’	1	HL	1			1
37.	6	25	然后你，你再把这些，根据你自己想要的口味	想 <i>xiǎng</i> ‘will’		L	1			1
38.	6	29	然后现在把这个<SIL>馅儿和均匀	和 <i>huò</i> ‘mix’		HL		1	1	1
39.	6	33	...把这个菜放在一边	放 <i>fàng</i> ‘place’		LH		1	1	
40.	6	39	...要把它弄成一个，就是	弄 <i>nòng</i> ‘make’		HL		1	1	1
41.	6	42	把它弄圆	弄 <i>nòng</i> ‘make’		HL		1	1	1
42.	6	50	右手把它像这样	它 <i>tā</i> ‘it’		LH	1			
43.	6	51	每个都要把它弄均匀	弄 <i>nòng</i> ‘make’		HL		1	1	
44.	6	59	你要把它放到手上	放 <i>fàng</i> ‘place’		HL		1	1	1
45.	6	64	你就把它放在面板上	放 <i>fàng</i> ‘place’	1	HL		1	1	1
46.	6	87	对，你把它放在<SIL>左手	放 <i>fàng</i> ‘place’		HL		1	1	1
47.	6	90	对，把馅儿放在这个饺子皮中间	馅 <i>xiàn</i> ‘stuffing’		HL	1			1
48.	6	93	你把这个饺子中间对折	对 <i>duì</i> ‘opposite’		HL		1	1	1

49.	6	97	再把左边右边都捏紧	捏 <i>niē</i> ‘pinch’	H		1	1					
50.	7	6	就是先把面倒到碗里	倒 <i>dào</i> ‘pour’	HL		1	1	1				
51.	7	22	我们先把那个和好的面团	和 <i>huò</i> ‘mix’	HL	1							
52.	7	23	把它揉成一个长条形	揉 <i>róu</i> ‘roll’	LH		1	1					
53.	7	25	就是用手把它揉成长的	揉 <i>róu</i> ‘roll’	LH		1	1					
54.	7	32	你把它搓好后										
55.	7	33	就可以把它										
56.	7	37	你把它揪得再均匀一点	揪 <i>jiū</i> ‘pinch’	H		1	1					
57.	7	41	把那个揪下来的那个剂子给它搓	剂 <i>jì</i> ‘daugh’	HL	1							
58.	7	43	...把它按下去	按 <i>àn</i> ‘press’	HL		1	1					
59.	7	52	就把那个边擀得均匀一点	擀 <i>gǎn</i> ‘roll’	L		1		1				
60.	7	56	然后右手把馅放上去	放 <i>fàng</i> ‘place’	HL		1	1					
61.	7	61	然后就要用手把这个皮儿对折一下	对 <i>duì</i> ‘opposite’	HL		1	1					
62.	7	66	把两边都										
63.	7	71	就是要把菜和肉混在一起	混 <i>hùn</i> ‘mix’	HL		1		1				
64.	7	87	等你把它搅匀了	搅 <i>jiǎo</i> ‘mix’	L		1		1				
65.	8	10	把面粉倒在一	倒 <i>dào</i> ‘pour’	HL		1	1					
66.	8	32	我们可以自己把										
67.	8	33	把和好的面先放在一边儿	和 <i>huò</i> ‘mix’	HL	1							
68.	8	48	我们把和好的馅也先放在一边	和 <i>huò</i> ‘mix’	HL	1							
69.	8	51	然后把它揉成一个长条状	揉 <i>róu</i> ‘roll’	LH		1		1				
70.	8	64	然后把它拿起来	拿 <i>ná</i> ‘take’	LH		1		1				
71.	8	74	把这个面放下	放 <i>fàng</i> ‘place’	HL		1	1					
72.	8	89	把它放在面板上	放 <i>fàng</i> ‘place’	HL		1		1				
73.	8	107	然后把皮儿	皮 <i>pí</i> ‘skin’	LH	1							
74.	8	131	把它捏紧	捏 <i>niē</i> ‘pinch’	H		1	1					
					44	2	8	5	54	2	40	14	29