THE ACQUISITION OF THE *BA*-CONSTRUCTION IN MANDARIN CHINESE: A STUDY OF HERITAGE SPEAKERS AND SECOND LANGUAGE LEARNERS

Xiaodan Wu and Bettina Spreng
University of Saskatchewan

1. Introduction

This paper summarizes some of the results of a study to compare the usage and underlying knowledge of *ba*-sentences by second-language learners, heritage speakers and monolingual L1 speakers of Mandarin Chinese (Wu 2023). The results are based on an online experimental study using grammatical judgement and limited production tasks.

The paper is structured as follows: Section 1 gives some background on the different speaker groups, outlines the properties of *ba*-sentences in Mandarin Chinese, and provides the research questions that guided the investigation. Section 2 outlines the research methodology, section 3 describes the main results, and section 4 summarizes the implications and consequences of the study.

1.1 Speaker groups

It is well known that second language acquisition (2LA) outcomes differ depending on various factors. While age plays a significant role in acquisition, it is not always the dominant factor (Montrul 2008). Heritage speakers who acquire a language as a first language in childhood may turn to a later acquired second language as the dominant language in later years, causing the first language to be considerably affected by the dominant second language (Polinsky 2018). This shift often happens in minority language contexts, such as with immigration or with minority indigenous languages. Counter to early second language acquisition research that assumes that the first language is rarely affected by a later acquired language, work on heritage language speakers has shown that there is a considerable effect, often connected to the age of acquisition/exposure to the second language **(**Polinsky et al. 2010).

On the other hand, age of acquisition and exposure is also a factor in 2LA contexts where the first language remains the dominant language. Alongside other factors such as exposure to other languages in childhood, quality and quantity of instruction, age of acquisition is a strong predictor for outcomes in 2LA.

In this study, L1 speakers of Mandarin Chinese, such as monolinguals and heritage speakers, form a group for age of acquisition (infancy) and their learning environment at home. For language dominance and fluency in English, heritage speakers of Mandarin Chinese form a group with L2 learners. Note that English is the L2 for heritage speakers, while it is the L1 for L2 learners of Mandarin Chinese. Thus, heritage speakers have a

Actes du congrès annuel de l'Association canadienne de linguistique 2023.
Proceedings of the 2023 annual conference of the Canadian Linguistic Association.

high dominance and fluency in their L2, while L2 learners have high dominance and proficiency in their L1.

1.2 Differences between *ba*-sentences and non-*ba*-sentences

Mandarin Chinese has a default SVO word order in declarative sentences. As shown in (1a), the subject is followed by the verb, and the result complement precedes the perfect marker *le* and sentence-final direct object *boli* 'glass'. The *ba*-sentence (1b) shows SOV order with the subject, the particle *ba*, and the object. The verb follows the object, and the complements are in the final position.

(1) a. SUBJECT + VERB + (COMPLEMENTS) + OBJECT
他 打 碎 了 玻璃。
ta da sui le boli.
he broken [RC] PFV[1] glass
'He broke the glass.'

b. SUBJECT + ***ba*** + OBJECT + VERB + (COMPLEMENTS)
他 把 玻璃 打 碎 了。
ta ***ba*** boli da sui le.
he **BA** glass broke [RC] SFP
'He broke the glass.'

While *ba*-sentences require the object to be interpreted as indicating a result state, SVO-order sentences do not. The *ba*-sentence in (2a) is ungrammatical because the cancellation of *ta ba wo pian le* 'he fooled me' negates the result of 'fooling me' and thus can only be combined with an SVO sentence as in (2b).

(2) a. *他 把 我 骗 了, 可 我 没有 上当。
ta ***ba*** wo pian le, ke wo meiyou shangdang.
he **BA** I lied PFV, but I not fool
'He fooled me, but I didn't fall for it.'

b. 他 骗 了 我，可 我 没有 上当。
ta pian le wo, ke wo meiyou shangdang.
he lied PFV me, but I not fool
'He fooled me, but I didn't fall for it.' (Liu and Zhao 2005: 111)

[1] Abbreviations and terminology: PVF = perfect verb aspect marker; SFP = sentence final particle; RC = Result Complement. The term complement is traditionally used as term for any predicate modifiers including adjuncts in Chinese linguistics. This usage will be followed in this paper.

Ba-sentences may include different types of postverbal complements modifying the predicate for the focus on the state of the objects. The result complement (3a) enables the result state of the object together with the *ba*-sentence structure.

(3) a. NP1+ *ba* +NP2+ V + result complement
他 把 玻璃 打 碎 了。
ta ***ba*** boli da **sui** le.
He **BA** glass broke [RC] SFP
'He broke the glass.'

The direction complement modifies the predicate, indicating the direction of an activity (4a). The *ba*-sentence must enable the focus on the object's state and is ungrammatical without it (4b).

(4) a. NP1+ *ba* +NP2+ V + direction complement.
你 把 衣服 拿 进来。
ni ***ba*** yifu na **jinlai.**
you **BA** clothes bring in [DC]
'You bring your clothes in.'

b. *你 把 衣服 拿。
ni ***ba*** yifu na.
you **BA** clothes bring
(int.) 'You bring your clothes.' (Fu 2013: 7)

The quantity complement specifies the number of event repetitions denoted by the predicate (5a). Like the direction complement, it is obligatory for grammaticality.

(5) a. NP1+ *ba* +NP2+ V + quantity complement
我 把 这本 书 读 了 三遍。
wo ***ba*** zheben shu du le **sanbian.**
I **BA** this book read PFV three times [QC]
'I read this book three times.' (Fu 2013: 7)

b. *我 把 这本 书 读。
wo ***ba*** zheben shu du.
I **BA** this book read
'I read this book.'

The prepositional complement is a prepositional phrase with different meanings. Example (6) illustrates locative with the preposition *zai*. This PP is not sentence-final.

(6) a. NP1+ *ba* +NP2+ V + zai+ NP3 (locative)

他	把	一本书	放		在	桌子	上。
ta	***ba***	yiben	shu	fang	**zai**	**zhuozi**	shang.
he	**BA**	a	book	put	on [PC]	table	upside

'He put the book on top of the table.' (Fu 2013: 7)

b. *他 把 这本 书 放。
ta ***ba*** zheben shu fang.
he **BA** this book put
(int.) 'He put the book.'

The degree complement introduced by the particle *de* 'to' denotes the degree to which the result state has been achieved (7a). It is obligatory for *ba*-sentences (7b) and incompatible with SVO sentences (8) regardless of position.

(7) a. NP1+*ba*+NP2+V+ de complement [Degree complement]?

他	把	教室	打扫	得	干干净净。
ta	***ba***	jiaoshi	dasao	**de**	**ganganjingjing.**
he	**BA**	classroom	clean	to	completely

'He cleaned the classroom completely.' (Fu 2013: 7)

b. *他 把 教室 打扫。
ta ***ba*** jiaoshi dasao.
he **BA** jiaoshi clean
(int.) 'He cleans the classroom.'

(8) a. *他 打扫 得 干干净净 教室。
ta dasao **de** **ganganjingjing** jiaoshi.
he clean to completely classroom
'He cleaned the classroom completely.'

b. *他 打扫 教室 得 干干净净。
ta dasao jiaoshi **de** **ganganjingjing.**
he clean classroom too completely
(int.) 'He cleaned the classroom completely.'

1.3 Acquisition of *ba*-sentences

Chinese native speakers usually use simple *ba*-sentences in early childhood, beginning at 24 months. By 48 months, their *ba*-constructions become syntactically complex (Chang and Zheng 2017). Thus, we can expect *ba*-sentences in speakers who have acquired Mandarin Chinese in early childhood.

The *ba*-construction necessitates a non-canonical SOV word that does not exist in English and a postverbal complement to expound on the result state of the direct object. Research has shown that learners go through stages of simpler complement structures in *ba*-sentences and exhibit errors in word order and type of complements (Fu 2013).

Heritage speakers of Mandarin Chinese with a dominant English L2 show influences of English word order and also simpler complement structures in Mandarin Chinese *ba*-sentences (Polinsky et al. 2010).

1.4 Research Questions

The above insights from previous work on the acquisition of Mandarin Chinese as L1 and L2 with differing dominance levels and ages of acquisition prompted the following research questions to be addressed in this study.

What is the relationship between the participants' responses and the dominant language environment and dominant language use? Do different dominant language environments and different language use affect the use and knowledge of *ba*-sentences?

Does the age of acquisition of the L2 affect L2 learners and HL speakers of Mandarin Chinese differently? Since the L2 is Mandarin Chinese for the L2 learner group, is it affected by English differently than with HL speakers whose L2 is English?

How do the three speaker groups differ for underlying grammatical knowledge of *ba*-sentences? Do judgements and usage show different patterns for the speaker groups?

Since L2 and L1 acquisition differ in complement types and complexity, does that affect HL and L2 speakers differently?

2. Methodology

Thirty-nine Chinese native speakers (CN), 24 second-language learners (L2), and 25 heritage Chinese speakers (HL) were subjected to data analysis. Some Heritage Chinese speakers and Chinese second learners could listen and speak but not read or write Chinese characters; therefore, they were provided with English translations and audio.

The experiment was created using the Gorilla software (https://gorilla.sc/). It included five parts: consent, biographical and demographic information, Task 1 (judgement task), Task 2 (fill in the blanks), and Task 3 (word order choices). A questionnaire was used to determine proficiency level, age of acquisition of the L2, and language and linguistic dominance.

2.1 Grammatical judgment task

Grammatical judgment tasks are used to access unconscious knowledge about an existing structure for acceptability (Mandell 1999, Polinsky and Kagan 2007). In the grammatical judgments task of this study (see Figure 1 below), the participants were presented with *ba*-sentences with different word orders and were asked to choose each sentence as acceptable, not acceptable, or not sure.

Figure 1. Task 1: Grammatical judgment task.

2.2 Multiple choice task

The second task was a multiple-choice task (see Figure 2 below), which tested the participants' understanding of complements in *ba*-sentences. In this test, the participants were given a situation to describe and select the appropriate complement phrases. There were five complements: *de* complement, direction complement, preposition complement, quantity complement, and result complement. Each sentence offered four options, each of which had a similar meaning but would be used differently in the sentences. The participants picked which complement was appropriate for the given sentence.

Figure 2. Task 2: Multiple choice task.

Option A corresponds to (9a), B to (9b), C to (9c) and D to (9d). This task simulates controlled production, focusing on the use of complements. Only (9a) is grammatical.

(9) a. 我把鸡 A 抓住了

wo	*ba*	ji	zhua	zhu	le.
I	BA	chicken	catch	[RC]	SFP

'I caught the chicken.'

b. *我把鸡 B 抓好了

wo	*ba*	ji	zhua	hao	le.
I	BA	chicken	do	well	SFP

(int.) 'I caught the chicken.'

c. *我把鸡 C 抓。
wo *ba* ji zhua.
I BA chicken catch.
(int.) 'I caught the chicken.'

d. *我把鸡 D 不确定。
wo *ba* ji buqueding.
I BA chicken not sure.
(int.) 'I caught the chicken.'

2.3 Constructing *ba*-sentences

Syntax is often mastered later than morphology and phonology in language acquisition, even if the earlier components are not fully acquired. Studies of how L2 learners acquire grammatical morphemes, negation, questions, and reference to the past (Lightbown and Spada 2006) show that language learners with different language backgrounds go through similar developmental stages in acquiring these linguistic features.

The final task assessed individuals' ability to appropriately integrate the *ba* phrases (see Figure 3 below) with direct objects ① or a second noun phrase combined with other complements ② and for negation ③. These are the only expected grammatical orders.

① Direct-object *ba*-phrase: NP 1+ *ba* + NP 2+Verb,
② *ba*-phrase with second noun-phrase: NP 1+ *ba* + NP 2+Verb+others,
③ Negation: NP 1+ no/not + *ba* + NP 2+Verb.

Figure 3. Task 3 Constructing ba-sentences task.

3. Results

This section includes descriptive statistics and the statistical analysis to assess the interaction of the different variables, which was done in R version 3.6.3 (R Core Team 2019). There are four models: two regression linear models (Task 1) and two generalized linear mixed models (Task 2 and Task 3).

Each variable was coded as follows: the first language is *L1*; the second language is *L2*; the age of acquiring the second language is *Age L2*, and the predominant language is *Pred L*. The predominant language is *Pred English and Pred M Chinese (Mandarin Chinese)*. The age of participants is divided into three levels: *Age_ Pre (18-30)*, *Age_Mid (30-40)*, *Age Post (40-50)*. The age of learning a second language is split into two groups, for instance, before 11 and after 11, coded as *Age_L2Bef11* and *Age_L2Aft11*. Snow and Hoefnagel-Höhle (1978) provide evidence that most children before puberty are eventually more effective than adults in L2 acquisition (SLA), albeit they are not necessarily faster.

3.1 Task 1: Grammaticality judgments

3.1.1 Descriptive summary

The first task tested whether the CN group differs from the HL and L2 groups in underlying knowledge. This task included 26 *ba*-sentences divided into three types.

- Type 1: Correct word order.
- Type 2: Correct word order, but there are other problems.
- Type 3: Wrong word order.

Figure 4 shows that the CN and the HL groups have the same trend in judging sentences; they chose acceptable more than unacceptable.

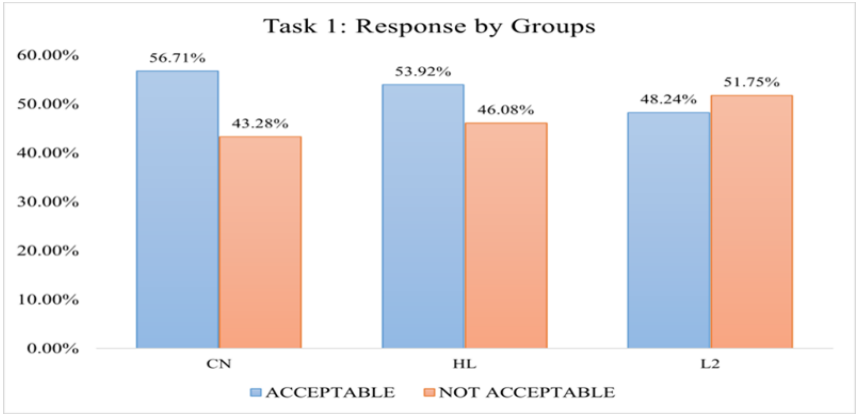

Figure 4. Overall responses by participant groups.

Figure 5 below illustrates the acceptance rate of each group for various patterns. For Type 1, HL's acceptance rate was similar to that of the CN group and higher than that of L2. The L2 group had a greater acceptance rate for sentence pattern 3 with the wrong order than the other two groups. Except for sentence pattern 1, CN has a lower acceptance rate than the other two groups for Types 2 and 3.

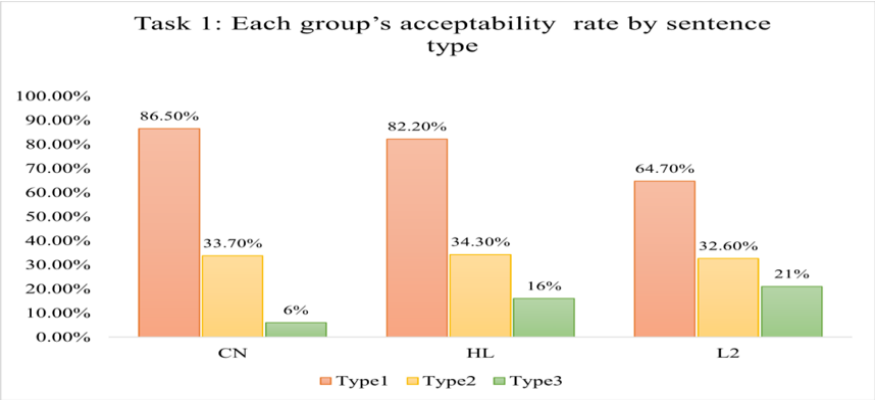

Figure 5. Each group's acceptability rate by sentence type (Task1).

3.1.2 Results of Model 1

A generalized linear model (Model 1) analysis was performed using the *glmer* function (Bates et al. 2015) to test the interaction between answer patterns and the group/age/environment/age of acquisition. The results are summarized in Table 1.

The independent variables were the speaker groups, predominant language, age, and age of acquisition. The dependent variable was response patterns (ACCEPTABLE and NOT ACCEPTABLE). The answers were coded as numerals, with 0 representing "A" (ACCEPTABLE) and 1 representing "NA" (NOT ACCEPTABLE). Therefore, the model treated "NA" as the default outcome. The output was the speaker groups, for example, the CN group, HL group, or L2 group. The model also included participants as a random effect. Model 1 compared CNHL (Mandarin Chinese native and Heritage speakers) with L2 speakers. CN was combined into the intercept with HL since there was only a small difference between HL and CN speakers. Thus, Model 1 compared CNHL with L2 speakers, which began with a large model and was gradually minimized by deleting insignificant independent variables. GroupL2 and Age_Post were significant variables. Pre_dominant language, Age_Pre, Age_Mid and Age_L2 were non-significant independent variables.

Pred English, Pred Chinese, Age_Pre, Age_Mid, and Age_L2Bef11 were non-significant independent variables. Therefore, these independent variables were deleted one by one based on the t-value closest to 0. The next step was to create a final model. This model was built with two independent variables with participants as a random effect: group L2 and Age_Post as shown in Table 1.

Table 1. Summary of the generalized linear model on Task 1.

	Estimate	Std. Error	t value	Pr (>|z|)	Confidence intervals	
					2.5 %	97.5 %
(Intercept)	-0.23	0.06	-3.82	< .001	-0.35	-0.11
GroupL2	0.38	0.12	3.12	< .001	0.14	0.62
Age_Post	-1.01	0.38	-2.71	< .006	-1.80	-0.29

Based on the model output, L2 speakers chose NOT ACCEPTABLE, significantly differently from CNHL speakers, with the L2 speaker having a higher NOT ACCEPTABLE (see Table 1 above). Moreover, age plays a significant role, with Age_Post of Group L2 having a substantially lower frequency when choosing NOT ACCEPTABLE than the same age group of CNHL.

Obviously, speakers' group and speakers' age post define the participants answers distribution. Moreover, it tells us about the interaction between age (applies to all speakers) and their answers.

3.1.3 Results of Model 2

The acceptability model shows each group's acceptability rate by sentence type (Table 2 below). The model used groups as independent variables and answering patterns (ACCEPTABLE and NOT ACCEPTABLE) as dependent variables. The answers were again coded as numerals (same as in Model 1), with 0 representing "A" (ACCEPTABLE) and 1 representing "NA" (NOT ACCEPTABLE). Therefore, the model treated "NA" as the default outcome. The output was the groups and sentence types. The model included participants as a random effect as well.

Table 2. Summary of the generalized linear mixed model on sentence type (Task 1).

	Estimate	Std. Error	t value	Pr (>|z|)	Confidence intervals	
					2.5 %	97.5 %
(Intercept)	-1.66	0.11	-14.55	< .001	-1.89	-1.44
GroupL2	0.41	0.17	2.37	0.02	0.07	0.75
Type 2	2.09	0.11	18.58	< .001	1.87	2.32
Type 3	3.59	0.21	16.91	< .001	3.19	4.02

The acceptability model compared the three groups' acceptability rate by sentence type. Based on the model results, CN has the lowest non-acceptance rate for sentence type 1 and sentence type 2 compared to L2. As for type 3, there were no significant differences between the two groups (see Table 2 above). In conclusion, speaker groups determine how participants' answers are distributed. It also provides information about the interactions of group responses with sentence types (see Table 2 above).

3.2 Task 2: Multiple choice (fill in the complement task)

3.2.1 Descriptive summary

Figure 6 below shows the distribution of the participants' response patterns. All groups presented similar responses in that A was chosen the most. The rest of the answer patterns B, C, D, E, and F are distributed differently as shown in Figure 6 below.

The overall pattern of responses for the L2 group is quite similar to the CN and HL groups. The responses for pattern F, NOT SURE, are much higher than in the other groups; almost twice as high as the HL group and nine times higher than in the CN group.

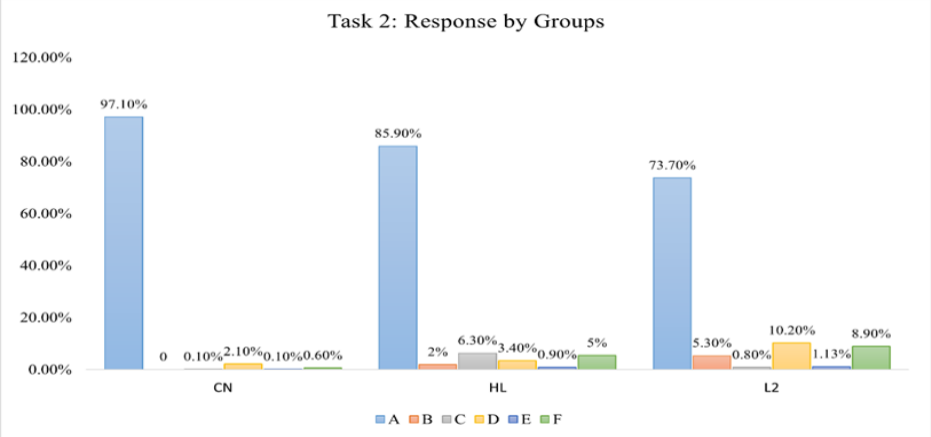

Figure 6. Overall responses by participant groups (Task 2).

3.2.2 Statistical Analysis for Task 2

As shown in Table 3 below, speaker group, dominant language environment, age post, and age of second language acquisition determine the participants' answer distribution. Furthermore, it indicates how language environment/age post/age of second language acquisition interact with their responses.

The intercept, the baseline, corresponds to the answer pattern A for the CN group. The coefficients, represented by log-odds values, indicate the change in the mean response associated with a change in one predictor (Group HL is 13.50). In contrast, the other predictors in the model are held constant. When looking at the results horizontally, the coefficients in the intercept column indicate an increasing trend from pattern B.

Moreover, L2 and HL speakers used simple complements significantly more than native speakers and showed differences in their use of various complements.

Table 3. Summary of the generalized linear mixed model 2 (Task 2).

Coefficients							
	(Intercept)	Group HL	GroupL2	Pred_Chinese	Pred_English	Age_Post	Age_L2Befl1
B	-16.91	13.50	14.15	-3.53	0.65	1.17	-0.89
C	-17.48	14.11	13.45	12.17	0.09	1.46	-1.79
D	-3.28	1.12	1.39	-0.16	0.46	0.61	-0.82
E	-5.03	2.03	1.09	-15.06	1.08	-12.82	-2.78
F	-5.62	2.78	2.97	1.39	1.23	0.76	-0.88
Std. Errors							
	(Intercept)	Group HL	GroupL2	Pred_Chinese	Pred_English	Age_Post	Age_L2Befl1
B	286.23	286.23	286.23	27.9	0.42	0.52	0.37
C	0.42	0.56	0.62	0.81	0.83	1.21	0.79
D	0.34	0.40	0.40	0.47	0.25	0.48	0.23
E	1.01	1.19	1.30	1.43e-06	0.84	1.08e-06	0.87
F	1.01	1.05	1.05	1.12	0.34	0.45	0.28
t value							
	(Intercept)	Group HL	GroupL2	Pred_Chinese	Pred_English	Age_Post	Age_L2Befl1
B	-0.06	0.05	0.05	-0.13	1.55	2.24	-2.41
C	-41.13	24.93	21.73	15.03	0.11	1.21	-2.25
D	-9.63	2.79	3.52	-0.35	1.81	1.31	-3.6
E	-4.98	1.71	0.84	-10527850	1.28	-1.19e+07	-3.21
F	-5.56	2.65	2.83	1.24	3.65	1.69	-3.17
P value							
	(Intercept)	Group HL	GroupL2	Pred_Chinese	Pred_English	Age_Post	Age_L2Befl1
B	0.929	0.962	0.961	0.899	0.120	0.024	0.015
C	0.000	0.000	0.000	0.000	0.915	0.227	0.024
D	0.000	< .005	0.000	0.727	0.070	0.187	0.000
E	< .001	0.087	0.403	0.000	0.199	0.000	< .001
F	< .001	< .008	< .004	0.215	0.000	0.091	< .002
Confidence intervals							

	(Intercept)		Group HL		GroupL2		Pred_Chinese		Pred_English		Age_Post		Age_L2Befl1	
	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %
B	-577.92	544.09	-547.51	574.50	-546.86	575.16	-58.25	51.18	-0.16	1.46	-1.62	-0.17	-1.62	-0.17
C	-18.31	-16.64	13.00	15.22	12.24	14.67	10.58	13.76	-1.53	1.71	-0.91	3.84	-3.34	-1.27
D	-3.95	-2.61	0.33	1.91	0.62	2.17	-1.07	0.75	-0.04	0.95	-0.30	1.53	-1.27	-0.37
E	-7.01	-3.05	-0.30	4.36	-1.46	3.64	-15.06	-15.06	-0.57	2.72	12.82	-12.82	-4.48	-1.08
F	-7.60	-3.64	0.73	4.83	0.92	5.02	-0.81	3.59	0.57	1.87	-0.12	1.64	-1.42	-0.33

3.3 Task 3: Multiple choice (word order task)

3.3.1 Descriptive summary

Figure 7 below shows the distribution of the participants' answer patterns. The answer pattern distributions of the HL and L2 groups were consistent with the CN answer distribution. Compared with the CN group, answer D has increased by almost five for the HL and was seven times higher for the L2 group.

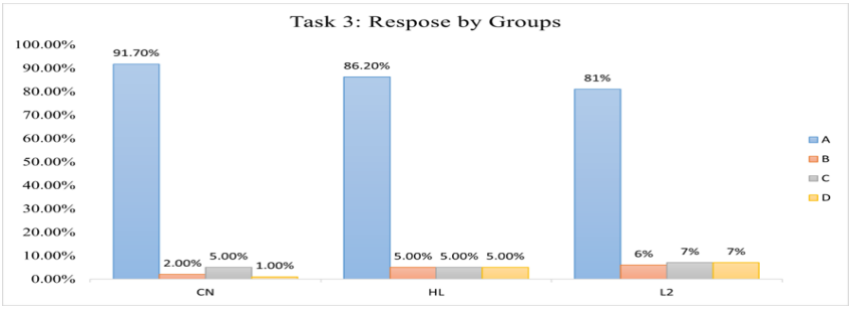

Figure 7. Overall responses by participant groups (Task 3).

3.3.2 Statistical Analysis for Task 3

As the following sections will show, the participants' answers are influenced by their group, predominant English, age post, and age of second language acquisition. Additionally, it indicates how participants' answers are influenced by English language environment, age post, and age of second language acquisition.

Results from Model 3 suggested that the participant group plays a significant role. As a baseline, the intercept represented the response pattern A for the CN group. The intercept column showed that the chance of selecting options was a declining trend in response distribution, and the log-odds value of selecting pattern B vs. pattern A will reduce, HL and L2 speakers utilize the word order incorrectly since there were instances of SVO order with the *ba*-sentences (see Table 4 below).

Table 4. Summary of the generalized linear mixed model 3 (Task 3).

Coefficients						
	(Intercept)	Group HL	GroupL2	Pred_English	Age_Post	Age_L2Bef11
B	-2.96	1.94	1.69	-0.95	0.85	-1.83
C	-2.21	0.43	0.45	-0.27	0.51	-1.15
D	-4.19	2.01	2.09	-0.35	1.54	-0.51
Std. Errors						
	(Intercept)	Group HL	GroupL2	Pred_English	Age_Post	Age_L2Bef11
B	0.28	0.39	0.40	0.35	0.69	0.30
C	0.20	0.34	0.34	0.32	0.67	0.24
D	0.42	0.48	0.48	0.32	0.50	0.32
t value						
	(Intercept)	Group HL	GroupL2	Pred_English	Age_Post	Age_L2Bef11
B	-10.38	4.94	4.18	-2.76	1.24	-5.98
C	-11.07	1.27	1.29	-0.85	0.76	-4.70
D	-9.97	4.17	4.35	-1.13	3.06	-1.60
P value						
	(Intercept)	Group HL	GroupL2	Pred_English	Age_Post	Age_L2Bef11
B	0	< .001	< .001	< .006	0.216	< .001
C	0	0.204	0.196	0.394	0.450	< .001
D	0	< .001	< .001	0.259	< .002	0.109

Confidence intervals												
	(Intercept)		Group HL		GroupL2		Pred_English		Age_Post		Age_L2Bef11	
	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %	2.5 %	97.5 %
B	-3.51	-2.40	1.17	2.71	0.90	2.46	-1.62	-0.28	-0.50	2.20	-2.42	-1.23
C	-2.61	-1.82	-0.24	1.11	-0.23	1.12	-0.89	0.35	-0.80	1.81	-1.63	-0.67
D	-5.02	-3.37	1.07	2.95	1.15	3.03	-0.98	0.26	0.55	1.81	-1.12	0.11

3.4 Summary

In Task 1, the participants were evaluated on their grammatical knowledge of *ba*-sentences. Model 1 demonstrates that L2 speakers judge *ba*-sentences in the same way as CN speakers. On the other hand, HL speakers judge them similarly to the CN group. Thus, HL speakers' pattern with CN speakers in judgment.

Task 2 consisted of multiple-choice questions that tested the participants' ability to use complements appropriately in *ba*-sentences. The results in Model 2 demonstrate that L2 and HL speakers use simple complements much more than CN speakers. In addition, HL and L2 speakers show differences in their use of various complements.

Task 3 asked the participants to construct a new *ba*-sentence to test the participants' word order production in *ba*-sentences. Model 3 illustrates that HL and L2 speakers utilize the word order incorrectly, accepting instances of SVO order with *ba*-sentences. A more detailed analysis of the data is available in my master's thesis (Wu 2023).

4. Conclusion

The investigation aimed to see whether and how the use of *ba* constructions differs between CN, HL, and L2 speaker groups. The hypothesis was that HL speakers would pattern with native speakers more often since both groups acquire Chinese as an L1 in early childhood.

The experimental findings revealed that the predominant English language environment for HL and L2 plays a role in distinguishing the speaker groups for production. Compared to the other age groups, the post-age of all speaker groups also affects HL and L2's choices.

The age of acquisition of the second language (Chinese for L2 and English for HL speakers) also impacts the use of *ba*-sentences distinguishing L2 from HL and CN speaker groups. The type of task also affected the performance. In judgment tasks (Task 1), HL and CN patterned more closely, while L2 and HL patterned more closely in production tasks.

Moreover, this study shows that SVO order with *ba* was present in both HL and L2 speakers, suggesting the influence of English on *ba*-sentence word order. Variations between HL and L2 in *ba*-sentences were thus very likely driven by f language transfer.

There are some limitations in the experimental design. The first sample size was small, and the second experimental question design was not sufficiently comprehensive.

For future research, it would be interesting to find if the production results also hold for free naturalistic data production, such as storytelling or sentence production.

This study bridges the gap between heritage speakers and L2 studies using Chinese *ba*-sentences. Moreover, it demonstrates the role of language dominance and language environment in the direction and strength of language transfer in speech production.

References

Bates, Douglas. 2005. Fitting linear mixed-effects models using lme4. *RNews* 5(1): 27–30.

Chang, Hui, and Lina Zheng. 2017. The development of *ba* construction in the Mandarin-speaking child: A longitudinal study. *Applied Linguistics* 2: 98–106.

Fu, Rong. 2013. *Foreign students' error analysis and teaching strategies in the acquisition of the complement of "ba" sentence.* Postgraduate thesis, Central Chian Normal University.

Lightbown, Patsy. M., and Nina Spada. 2006. *How Languages Are Learned* (3rd ed.). Oxford: Oxford University Press.

Liu, Peiyu, and Jinghua Zhao. 2005. Ba zi ju yu zhu dong bin ju de ju fa, yu yi he yu yong de cha yi. *Journal of Changsha University of Science and Technology* 20(3): 111–114.

Mandell, Paul B. 1999. On the reliability of grammaticality judgment tests in second language acquisition research. *Second Language Research*, 15(1): 73–99.

Polinsky Maria, Boyan Zhang, and Carlos Gomez Gallo. 2010. Heritage Chinese: A new view from production, In *Fourth Heritage Language Institute*. Manoa, HI: National Heritage Language Resource Center.

Montrul, Silvina. 2008. Second language acquisition welcomes the heritage language learner: Opportunities of a new field. *Second Language Research* 24(4): 487–506.

Polinsky, Maria, and Olga Kagan. 2007. Heritage languages: In the "wild" and in the classroom. *Language and Linguistics Compass* 1(5): 368–395.

Polinsky, Maria. 2018. *Heritage Languages and their Speakers*. Cambridge: Cambridge University Press.

R Core Team. 2019. *R: A language and environment for statistical computing*. In *R Foundation for Statistical Computing*, Vienna, Austria.

Snow, Catherine E., and Marian Hoefnagel-Höhle. 1978. The Critical Period for Language Acquisition: Evidence from Second Language Learning. *Child Development* 49(4): 1114–1128.

Wu, Xiaodan. 2023. *Use of the "ba" construction in Heritage Chinese speakers and L2 speakers of Chinese*. Master's thesis, University of Saskatchewan.