9-2020

Processing Coercion in a First, Non-Dominant Language: Mandarin-English Heritage Bilinguals

Christina N. Dadurian
The Graduate Center, City University of New York

PROCESSING COERCION IN A FIRST, NON-DOMINANT LANGUAGE:

MANDARIN-ENGLISH HERITAGE BILINGUALS

by

CHRISTINA DADURIAN

A thesis submitted to the Graduate Faculty in Linguistics in partial fulfillment of the

requirements for the Master of Arts, The City University of New York

2020

Processing coercion in a first, non-dominant language:
Mandarin-English Heritage Bilinguals

by

Christina Dadurian

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in
satisfaction of the thesis requirement for the degree of Master of Arts.

_____
Date

_____
Virginia Valian

Thesis Advisor

_____
Date

_____
Gita Martohardjono

Executive Officer

THE CITY UNIVERSITY OF NEW YORK

iii

ABSTRACT

Processing coercion in a first, non-dominant language: Mandarin-English heritage bilinguals

by

Christina Dadurian

Advisor: Virginia Valian

Recent work in heritage language grammars has shown variability in L1 competence, despite high proficiency in both languages. While sources of variation have been debated, little attention has been given to the role of language dominance. This thesis uses a self-paced listening task to explicitly investigate the roles of language dominance and pragmatic competence in how heritage speakers of Mandarin Chinese process aspectual coercion in their non-dominant home language, as compared to late bilinguals. Specifically, constructions that vary in acceptability and salience in input between Mandarin and English are tested: Iterative coercion, complement event coercion of entity NPs, and perfective states.

Stimuli are presented auditorily and participants are given two comprehension tasks: 1) Temporal interpretation, and 2) Judgment of acceptability. Answers are compared to Mandarin-dominant late bilinguals' judgments to derive accuracy and probability of accepting ungrammatical forms, while listening time is taken as a proxy for processing cost.

Temporal interpretation in Mandarin relies heavily on pragmatic inference while English marks tense and aspect mandatorily. In addition, resolution of aspectual mismatch has been theorized to be a function of pragmatic knowledge, which tends to be weakened in the heritage language. Thus far, no work has explicitly attempted to find psycholinguistic evidence for such claims. Participants are given a pragmatic competence task in Mandarin in order to test for a relationship with accuracy with correct choices on the temporal interpretation task.

ACKNOWLEDGMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

In this section, I will outline the motivation and aims of my thesis. This work is relevant to the fields of psycholinguistics, heritage languages, semantics, and bilingualism. To my knowledge, this is the first study to explicitly investigate the processing strategies of proficient Mandarin-English bilinguals in their *heritage* language Mandarin Chinese, particularly with regards to aspectual coercion, and would contribute to only recent work on modeling a heritage language grammar (Polinsky & Scontras, 2019; Montrul, 2012; Putnam & Sanchez, 2013). There is great variation in the status of the L1 in adult heritage speakers (Valian, 2019), and dominance can help account for this variation. I will contribute to the standardization of the *Bilingual Language Profile* (BLP: Birdsong, Gertken, & Amengual, 2012) as a measure of dominance, which will also be compared to an objective Relative Fluency dominance index (Stevens, 2019). I aim to answer the following research questions:

1. How does L2 dominance vary in heritage speakers, and what other linguistic variable(s) correlate with high L2 dominance?

2. Does English dominance contribute to the acceptability of Mandarin aspectual coercion? If so, is accuracy or processing affected?

3. Does language-specific salience of coercion constructions affect accuracy of temporal interpretation?

4. Is resolving aspectual mismatch and deriving temporal location in Mandarin in fact a function of pragmatic competence?

Results from this study could elucidate the role of input via structural salience in highly proficient heritage bilinguals, considering Mandarin and English share some aspectual coercion constructions but with varying salience. Pragmatic competence, though implicated in theoretical explanations of resolving aspectual mismatch (Brennan & Pylkkänen, 2008; Dölling, 2003), has not yet been tested despite being a weakness in some bilinguals. If pragmatic knowledge plays a role, it would also be beneficial for language teachers to focus on in bilingual students.

2. Heritage bilinguals and language dominance

Heritage speakers are a subgroup of bilinguals who are exposed to a minority home language from birth, either exclusively (*sequential bilinguals*) or in conjunction with the majority language (*simultaneous bilinguals*)[1]. Heritage speakers are children of immigrants or early immigrants (> age 5) themselves. The L1, or heritage language, is a minority spoken language reserved for home use and possibly the immediate community, while they are formally educated in the dominant language in which they become highly proficient. While the heritage language is typically learned orally and naturalistically, the majority language is learned formally. They thus show differential patterns of acquisition of their L1 and L2 from baseline monolinguals, given differential instruction and input of both.

By early adulthood, the heritage language tends to show non-native like competence, non-uniform proficiency, slower processing, and gaps in knowledge that resembles that of second language learners (Montrul, 2012; Polinsky & Scontras, 2020; Polinsky & Scontras, 2019). These gaps are attributed to incomplete acquisition, attrition, and transfer from the dominant L2. Yet, heritage grammars and their errors are systematic and cohesive: There is a preference for

---

[1] Despite the possibility for simultaneous bilingualism, I will refer to the majority language as the L2.

salient forms and resistance to irregularity, commitment to fully compositional expressions, and resistance to one-to-many mappings of form and meaning.

Despite similarities in heritage speakers, there is still great variability in behavior and competence. According to Montrul (2012), adult heritage linguistic ability varies "from minimal aural comprehension ability to full fluency in written and spoken registers, and everything in beween" (pp. 70). Variation can be attributed to features of the specific languages (Muysken, 2019; Puig-Mayenco et al., 2018), variety in input (Gollan et al., 2015; Valian, 2019), and degrees of dominance (Zyzik, 2019; Fernández, 2003). In this section, I will describe past work that has contributed to our knowledge of heritage speakers that will motivate the need for measures of processing and language dominance in this population.

2.1 Development of heritage speaker grammars

A key component to understanding variation in heritage speakers is balanced or unbalanced language dominance. For our purposes and consistent with Gertken, Amengual, & Birdsong (2014), dominance is a *relative* measure of bilingualism determined by current and past use and exposure, proficiency in different domains (reading, writing, speaking), and attitudes towards languages. While these constituting variables are certainly related to one another, they cannot alone explain phenomena such as attrition, code-switching, processing, or restructuring. Relativity is central to dominance and can be contrasted with traditional linguistic measures like proficiency or fluency, which apply to one language. Dominance is essentially the ratio of that variable between one language and the other, which also changes throughout life. Heritage speakers are often dominant in the home language until puberty, while late immigrants can be

dominant in the minority language well into adulthood, and foreign language learners retain dominance in the majority L1 throughout life.

Balanced dominance, common but underacknowledged in heritage speakers, refers to when neither language is particularly dominant. While heritage speakers are often assumed to be dominant in the L2, extents can vary given the particular background of the bilingual. For example, bilinguals who continue to reside in communities that use the heritage language are more likely to maintain use and exposure and thus more balanced dominance.

When a bilingual becomes dominant in a second language, the L1 will undergo attrition without consistent exposure and use. This is often the case in heritage speakers who are highly L2-dominant. L1 attrition begins to occur when the L2 becomes "functionally dominant" and begins to influence the L1. Most often, features that become attrited are not salient in input, like low-frequency lexical items. Heritage language features that are not apparent in the dominant language are also commonly attrited, like aspectual marking. It is important to note that attrition can be reversed, as Köpke & Genevska-Hanke (2018) found in a case study. Their L2-German dominant subject was able to restore correct use of null pronominal subjects in L1-Bulgarian after three weeks of re-exposure to the L1 environment. With consistent exposure and use, attrition can be avoided.

Typically for simultaneous bilingual heritage speakers, the issue that arises for the L1 is not attrition but *incomplete acquisition*. This refers to ceasing the development of native-like competence of a feature in the L1 when the heritage speaker becomes functionally dominant in the L2. So features that are typically acquired late in language development, like derivational morphology, might never develop in simultaneous bilinguals. Exposure to a *variety* of heritage

language speakers in childhood can often prevent incomplete acquisition; L1 proficiency is maintained later in life despite possible changes in dominance (Gollan et al., 2015).

Regardless of the reason, unbalanced dominance has also been found to *delay* acquisition of language-specific temporal morphology in simultaneous bilinguals learning two languages during childhood. Brebner, McCormack, and Liow (2016) found that language dominance facilitated development of English tense morphology in a large sample of English or Mandarin-dominant kindergarten-aged children living in multilingual Singapore. The authors grouped participants based on parent and teacher reports of proficiency and use. Three subgroups were additionally formed by age (3;9-4;8, 4;9-5;8, 5;9-6;8). Participants were given a picture naming task to elicit verb-tense morphology. English-dominant children used significantly more tense morphology in English than did Mandarin-dominant children. In fact, Mandarin-dominant children produced *no* verb-tense morphology in English by the end of kindergarten. These results support that bilingual children learn features of the non-dominant language more slowly, if at all.

2.2 Measuring bilingualism

Since adult heritage speakers are proficient in both languages, they show native-like accuracy with production and interpretation tasks. Often, differences from native speakers can be found in more sensitive measures of processing. This section will illustrate the distinction between these measures. I suggest that accuracy reflects proficiency while processing reflects dominance. The role of input in accuracy and processing will also be explored but remains ambiguous.

To parse these variables apart, O'Grady et al. (2009) tested English-dominant Korean heritage bilinguals with high proficiency in both languages. High, medium, and low-frequency body part names were elicited in Korean and English. While participants were highly accurate in

both languages, word frequency significantly affected accuracy and naming response time, suggesting that input affects proficiency and processing. Dominance also showed an effect for naming response time, but not accuracy. Specifically, response times were significantly slower in the non-dominant heritage language across all vocabulary subsets. So, while high proficiency predicted high accuracy, English dominance predicted relative processing ease.

When dominance is *not* considered, heritage speakers can show unexplained variation in processing. Keating, Jegerski, and Vanpatten (2014) compared resolution strategies of highly proficient Spanish-English heritage speakers to Spanish monolinguals in L1 Spanish. Consistent with the findings in O'Grady, there were group differences in processing, not accuracy. A self-paced reading task was used to measure preference for interpreting null pronouns in a second clause as referring to the antecedent in SpecIP, and overt pronouns as referring to the object antecedent. This asymmetry in anaphora resolution is called the Position of Antecedent Strategy (PAS) and has been found to be robust in Italian and Spanish, both languages that make use of null pronouns.

Monolinguals showed the expected advantage in reaction time (RT) with overt pronouns while heritage speakers showed no advantage, despite similar knowledge of Spanish rules of syntax (namely, proficiency). There was, however, great variation in heritage speaker RT to the extent that some did show native-like processing. RT advantage in these bilinguals correlated with exposure to Spanish reading. The self-paced reading task might thus be best suited for proficient readers, when heritage speakers primarily speak and hear their home language. Neither Spanish proficiency nor age of exposure to English predicted RT, further implicating dominance in measures of processing.

Results from acceptability judgment tasks without processing measures have yielded different results. Zyzik (2019) tested adult Spanish-English heritage speakers' knowledge of L1-Spanish morphologically complex words. Participants were grouped as English or Spanish dominant given the *Bilingual Language Profile* and compared to monolingual Spanish controls. Also note the absence of a *balanced* dominance group — if an experimenter uses 0 as the cutoff point, participants with scores of -2 and 2 would be grouped as having different dominance profiles when in fact they are likely both balanced.

An acceptability judgment task was given that tested morphologically well-formed (conventional) and creative words in Spanish. In English, this is exemplified by say, *length* instead of **longness.* For an example from Spanish:

(1)   a. Creative: **permanición, *profundez*

         b. Conventional: *permanencia, profundidad*

Across groups, accuracy with accepting conventional items was highest, again suggesting that input affects proficiency since conventional items are present in input. But with creative items, there were significant group differences. The English-dominant group was significantly more likely than the Spanish-dominant to accept ungrammatical forms, and both groups of bilinguals were more likely to accept ungrammatical forms than monolinguals. In contrast with results from O'Grady, Spanish suffix frequency did not affect accuracy with creative items (but maybe it would have affected processing), suggesting that acceptability of *ungrammatical* forms might not reflect input but dominance. Montrul (2012) found similar results when comparing acceptability of ungrammatical bare plural NPs in Spanish between Spanish heritage speakers, Spanish L2 learners, and monolingual native speakers. Despite differences in length of exposure to Spanish,

both groups of bilinguals were English-dominant with similar Spanish proficiency. These groups were more likely to accept ungrammatical forms rejected by monolinguals while simultaneously demonstrating knowledge of the grammatical constructions.

In sum, heritage speakers can be distinguished from native speakers of the L1 through measures of processing. Measures of processing reveal relative processing difficulty in the non-dominant language, particularly with less salient items. Salience becomes evident in input, which is often reduced in heritage speakers; but quantity and quality of input can vary. Extents of language dominance can also vary in heritage speakers. Therefore, when we study heritage speakers, it makes sense to measure processing and to group by dominance, with consideration for structural salience. My study will follow suit with a self-paced listening task to measure processing of constructions with varying frequency between English and Mandarin. I will also use the BLP to determine dominance and to contribute to standardized operationalization. Since proficiency seems to determine correct acceptability of grammatical forms while dominance seems to determine incorrect acceptability of ungrammatical forms, items will be followed by an acceptability judgment question. I expect that coercion constructions acceptable in Mandarin but unacceptable in English will be most difficult to process for English-dominant participants, but correctly accepted. I expect constructions unacceptable in Mandarin but acceptable in English will be easier to process for English-dominant participants, but incorrectly accepted.

3. Tense and aspect

3.1 Tense

Tense is grammaticalized location in time (Comrie, 1985), and can refer to past, present, or future relative to speech time. This section will discuss how tense is derived in English. In

English, the two main grammaticalized categories of tense are *past* and *non-past* (I do not include future because it does not have dedicated marking). Tense is mandatorily marked on the verb with past *-ed* and in the third person, present *-s*:

(2)　　PAST: She *walked* to school.

　　　　PRES: She *walks* to school.

Verbs marked for the present tense must agree with the subject. With first and second person subjects, verbs do not receive any marking, e.g. *I walk to school*. Some verbs favor use in present or past tense, which will be discussed later.

　　Tense is determined by underlying pragmatic principles: The Deictic Pattern and the Boundedness Event Constraint. The deictic pattern says speech time is the present tense, and past and future are relative to speech time. The Bounded Event Constraint says that bounded events cannot be in the present tense because completion must precede the time of speech. What follows is that by default, bounded situations (e.g., *run, break*) are past tense before speech time, and unbounded situations like ongoing events or states (e.g. *believe, running*) are continuing and present (Smith & Erbaugh, 2005).

3.2 Aspect

　　Aspect can be grammaticalized or lexical and represents the internal constituency of an event relative to *situation* time (Comrie, 1976). In English, aspectual information combined with tense serves to distinguish an event as Perfective (PFV) or Imperfective (IMP). Perfective aspect represents an event as a bounded whole, meaning it has terminated at the situation time. Imperfective represents an event as an unbounded action in progress, meaning it is continuing

and incomplete at the situation time. First, I will differentiate between the notions of lexical and grammatical aspect using examples from English. Then, I will briefly describe the Mandarin aspectual system.

3.2.1 Lexical aspect (Aktionsart, situation aspect)

Across languages, verbs can be distinguished into classes by their inherent aspect given three binary features: stativity (vs. dynamicity), durativity (vs. punctuality), and telicity (vs. atelicity). Verb types also prefer to be presented in a given tense. The examples in Table 3.1 below denote the default tense of each type.

Table 3.1 Lexical aspect

	Stative	Telic	Durative	
States	+	-	+	I like to read.
Activities	-	-	+	He wrote letters.
Accomplishments	-	+	+	He wrote two letters.
Achievements	-	+	-	I found my sweater.
Semelfactives	-	-	-	I blinked twice.

Of note are *activities* and *accomplishments*, which are *lexically underspecified* for telicity. As seen above, the dynamic verb (denotes a durative event) *to write* can be either [+telic] or [-telic] depending on the cardinality of the direct object of the verb: *To write two letters* implies an endpoint of the event of letter-writing, while *to write letters* does not. So, in English, a dynamic verb such as *to write* is a [+telic] accomplishment when accompanied by a quantized direct object, and a [-telic] activity when the direct object is non-quantized.

Accomplishments, achievements, and semelfactives are also inherently bounded. Recalling the boundedness event constraint, bounded events necessarily take place in the past since they are completed. Presenting these predicates in the simple present is typically not acceptable outside of a narrative, e.g. *I write two letters, I blink twice*.

3.2.2 Grammatical aspect (viewpoint aspect)

Grammatical aspect provides the final layer of information for how to view the event dictated by lexical aspect. While lexical aspect provides the possible situation types, grammatical aspect presents the situation as completed (PFV) or continuing (IMP) at situation time. For example, information about boundedness is integrated to determine whether a [+telic] event is [+bounded] Perfective or [-bounded] Imperfective. In English, boundedness is marked by verbal suffixes -*ed* or -*ing*. Simple past -*ed* regardless of aspect indicates that the telic situation did in fact come to an endpoint (terminated and completed), while progressive -*ing* regardless of tense indicates that the situation did not complete. A summary of English grammatical aspect is depicted below, where PROG is progressive and HAB is habitual.

Table 3.1 Grammatical aspect

	PST	**PRES**	**Exceptions**
PFV (*bounded*):	Had + -ed	Have + -ed	
IMP (*unbounded*):			ACH, SEM, STA
• **PROG**	Was + -ing	Am + -ing	
• **HAB**	Used to + INF	Unmarked	

Since achievements and semelfactives are [-durative], or instantaneous, they cannot naturally occur in the progressive, e.g. *I am finding my sweater*. States are also unnatural in the progressive, e.g. *I am loving you.*

3.3 Mandarin tense and aspect

Mandarin does not explicitly mark tense. Instead, temporal interpretation relies on *deictic patterning* and the *bounded event constraint*. Without additional context or temporal adverbials,

default interpretation is speech time. Thus, only stative predicates can be used felicitously without additional information, though bare activities can also be felicitous when describing habitual action.

(3)   Yīchén   hĕn   cōngmíng.

          Yichen   very   smart

          'Yichen is very smart.'

States can also appear felicitously with present tense adverbials and receive a past interpretation with past tense adverbials. However, to indicate an eventive past reading, perfective marker *-le* must be used:

(4)   a.   *Kèrén gāngcái   dào.

               visitor just-now   arrive

        b.   Kèrén gāngcái   dào le.

               Visitor just-now   arrive PFV

               'The visitor arrived just now.'

This rule is not present in English because English past *-ed* implies the perfective. Since studies investigating Mandarin coercion constructions chiefly involve the perfective marker *le*, I will focus on its uses for the remainder of this section. A more comprehensive summary of all Mandarin aspectual markers comes from Slabakova (2015), where RVC means resultative verb complement:

        le: bounded event, terminated but not necessarily completed, tends to be interpreted as past;

RVC: bounded event, complete, tends to be interpreted as past;

guò: bounded prior situation, tends to be interpreted as past, or resultant state;

zài: unbounded event in progress, tends to be interpreted as present;

zhe: unbounded situation, tends to be interpreted as present;

[null]: telic events tend to be interpreted as past; atelic eventualities tend to be interpreted as present. *(pp. 288)*

Thus, Mandarin tense is constructed with a diverse system of aspectual markers combined with adverbials. Since English expresses tense overtly, e.g. past *-ed* or *was* x-*ing*, co-occurring temporal adverbials are only required for coercive operations (discussed later) and some habitual actions.

3.3.1 Perfective marker -le

In this section, I will present environments where verbal suffix *le* is acceptable or unacceptable and draw comparisons with English. Verbal *le* is generally associated with termination, boundedness, and perfectivity. According to Li & Thompson (1989: pp. 201):

-*le* is used when the event described by a sentence is perfective, which

means that the event is bounded, and an event is bounded if (1) if its

temporal or spatial limits are specified, (2) if it signals a specific event and

its direct object is definite, (3) if boundedness is inherent in the meaning

of the verb of the sentence, or (4) if it is followed by another event.

Use of -*le* depends on whether or not a situation is bounded. Examples of (1), (3), and (4) are shown below, respectively.

13

(5)   qián-shang   guà-*le*   yifu   huà

        wall-on   hang-PFV   one-CL painting

        "A painting was/had been hung on the wall."

(6)   Gàizi   diào-*le*

        lid   fall-off-PFV

        "The lid fell off."

(7)   Wǒ   chī-wán-*le*   nǐ chī

        I   eat-finish-PFV you eat

        "After I have finished eating, then you eat."

It is important to note that boundedness relies heavily on context, and native Mandarin speakers vary in their judgment of whether or not *le* is obligatory in many sentences given intended emphasis of the action (e.g. *sneezed*) vs. the object (e.g. *sneezed once*). Telic events can convey perfectivity without *le*. Locative phrases (e.g. *in the drawer*), directional phrases (e.g. *to his house*), and indirect objects (e.g. *to him*) can also indicate perfectivity without *le*.

Next, I will distinguish a *completed* event from a *terminated* event. In the perfective, telic situations are interpreted as completed while atelic situations are interpreted as terminated (Xiao & McEnery 2004). To illustrate:

(8)   wo zuotian xie-*le*   xin,   keshi   mei xie-wan

        I yesterday write-PFV letter, but   not write-RVC

        "I did some letter writing yesterday, but I didn't finish."

(9)   *wo mai-*le* san-ben shu,   keshi   mei mai-dao

        I buy-PFV three-CL book,   but   not buy-RVC

*"I bought three books, but I didn't succeed in buying them"

The difference in acceptability between (6) and (7) depends on the situation type. Since non-states must be bounded, we can look to the direct object to determine situation type. *Xin* is a bare noun Sentences like (6) are not contradictory in Mandarin. The perfective functions similarly to past progressive in English: *Wrote* is interpreted as a terminated event of writing that is no longer occurring at the time of speech. Felicitous coordination with RVC *wan* (finish) signifies that completion is not inherent to *le* with atelic predicates in Mandarin. Conversely in English, past perfective implies both completion and termination, as in *I had written a letter*.

While the default meaning of *-le* is perfectivity, with stative verbs, it derives an inchoative reading (Wang, 2008) signifying change of state.

> (10)   tianqi   leng-le,   nimen duo   chuan dian yifu
>   weather cold-PFV,   you   more wear little clothes
>   "The weather is becoming cold; you should wear more clothes."

> (11)   wo xihuan-le   ni   liang nian
>   I like-PFV   you   two year
>   "I liked you for two years (and no longer do)."

It is important to note that without the bounded adverbial *for two years* in (9), the sentence would be ungrammatical. Given the Bounded Event Constraint, states are necessarily unbounded in order to apply to the time of utterance, and to be bounded requires an adverbial. Like in English, where it is anomalous to use a stative as a perfective without an additional adverbial (e.g. *I have loved music*). *Le* will convert the state to a dynamic change of state. Note that the tense of the predicate *is-cold* in (5) is present while (6) is past, so while *-le* is typically given a past interpretation, it can also be present with overriding information.

3.4 Cross-linguistic influence of tense and aspect

Languages vary by the means that telicity and boundedness are expressed. In contrast to English, Russian morphologically marks the verb with prefix *PRO-/DO-* to indicate that it is [+telic]. For example, from Mikhaylova (2011):

(12)	Kolja čital	(eti) pisma.	[-telic], activity
	Kolja Ø.read.PST	(these) letters	
	'Kolja would read/was reading (these) letters.'		
(13)	Kolja PROčital	(eti) pisma.	[+telic], accomplishment
	Kolja PRO.read.PST	(these) letters	
	'Kolja read (these) letters.'		

The direct object of *to read* is quantized in both (1a) and (1b), but the predicate is [-telic] unless accompanied by *PRO*. Boundedness in Russian is also indicated by verb morphology: Imperfective suffix –(*y*)*va* is used like English *-ing* to mark [-bound] events that did not reach their endpoint. But unlike in English, –(*y*)*va* cannot be used with [-telic] predicates. Thus, every instance of Russian Imperfective necessitates a predicate that must be either marked with *PRO* or inherently telic.

Bilinguals who are dominant in a second language have been found to show reduced sensitivity to aspectual violations specific to the first language, in the following case Russian. Mikhaylova (2011) compared groups of Russian heritage speakers and Russian foreign language learners (L2s) to Russian monolinguals. Bilinguals scored within native speaker range of proficiency given a Cloze test, though L2s scored significantly lower than heritage speakers. Mikhaylova used a "Stop-Making-Sense" word-by-word self-paced reading task (SPRT) in

Russian. Sentences began with a disambiguating adverbial, e.g., *For two hours*, which either matched or mismatched the aspectual features of the predicate, which encoded telicity either lexically or morphologically (for underspecified predicates). Participants were told to push a button when the sentence stopped making sense (mismatched prior information).

Results showed significant between-groups and significant within-groups differences, except in the native speakers who were equally sensitive to all mismatches. L2s had significantly lower accuracy than heritage speakers and native speakers at detecting boundedness violations, particularly when morphologically telic verbs required Imperfective/[-bounded] *-yva* to match the durative adverbial. They also performed significantly worse than native speakers at detecting telicity violations, though were equal to heritage speakers. Within heritage speakers, mismatches with morphologically telic verbs were significantly easier to detect than with lexically telic verbs. They ultimately performed significantly worse than native speakers with telicity violations but performed the same with boundedness. These findings suggest that compared to native speakers, heritage speakers rely more heavily on overt morphology to interpret aspect in their non-dominant L1, whether due to incomplete acquisition or attrition of less salient features.

Mikhaylova's study also supports the proposal that heritage speakers have L1 grammars that are simultaneously distinct from and similar to that of L2 learners and native speakers. I think an interesting expansion would be to measure RT to decide ungrammaticality at the presentation of the verb because we could find more differences between heritage speakers and native speakers, who otherwise had comparable acceptability judgments. In addition, while the author considers the L2s to be "late bilinguals," and assumes English dominance of all bilinguals in the study, they lack objective measures of dominance and the operationalization of late bilinguals is not consistent with other studies, since late bilinguals in this study are formal Russian learners living

in English-dominant U.S.. A more nuanced categorization of the diverse backgrounds of these participants could lead to discovery of further differences.

3.4.1 Mandarin-English bilinguals

There is ample research based on comprehension and production of tense by Mandarin-English bilinguals in their dominant L2-English, but a dearth focusing on the heritage language. We might look to studies using L2-Mandarin in native English speakers assuming this population shows effects of prolonged English exposure on comprehension of Mandarin tense. Slabakova (2015) compared intermediate and advanced L2-Mandarin native English speakers residing in the U.S. and L2-English native Mandarin speakers residing in mainland China. Participants were given in Mandarin a Temporal Interpretation Choice task (no outside context), a Stories task (with context), and a Translation task from Mandarin to English.

The choice task stimuli consisted of simple sentences with bare states or activities, or aspectual morphology. Sample stimuli from two conditions include:

(14)   *Bare activity*:   Wǒ chī běijīng kǎo-yā

I eat   Beijing roast duck

'I eat Beijing roast duck.'   => *Present, generic*

- *le*:   Lǐsì chī-*le*   wǎnfàn

Lisi eat-PERF supper

'Lisi ate supper.'   => *Past*

Participants chose one of four possible options for temporal interpretation: Present, past, both, or neither. Native speakers and advanced learners did not statistically differ in accuracy, while

intermediate learners performed significantly worse than both other groups with RVC accomplishments and past-resultant *-guo*. Proficiency again predicted comprehension accuracy.

In the Stories Task, participants read five clauses constructing a story. An adverbial in the first clause indicated habitual past or present ongoing, and participants were asked to interpret the time frame of the last clause in each passage as past, present, both, or neither. Results from this task showed a significant effect of condition (past vs. present adverbial), group, and condition by group interaction. Neither native speakers nor advanced learners conformed to predictions that the past temporal adverbial in the first clause would hold scope over the aspectual information in the last clause. Instead, they were more likely to select a present interpretation with past adverbials and did not significantly differ. Intermediate learners significantly differed from native speakers and were more likely to select a past interpretation. In English, adverbials are commonly used in narratives to indicate the time of events as past or present. It is possible intermediate learners are less sensitive to aspectual morphology in Mandarin that follows a temporal adverbial because in English this adverbial holds scope over the following information.

The Translation task required participants to translate Mandarin sentences with conflicting lexical, aspectual, and adverbial information into English in order to determine what feature takes precedence in deciding temporal location. Accuracy was scored given use of past, present, or future tense in English. Again, intermediate learners were significantly less accurate than native speakers and advanced learners at translating sentences with past adverbials:

(15)　　Past adv + state:　　　　　　　Wǒ shàng ge xīngqī hěn máng
　　　　　　　　　　　　　　　　　　　I last-CL week very busy
　　　　　　　　　　　　　　　　　　　'I was very busy last week.'

Past adv + prog activity: Lǐsì zuótiān sì-diǎn zài děng gōngchē
Lisi yesterday four-o'clock Asp wait-bus
'Lisi was waiting for bus at four o'clock yesterday.'

This finding contradicts that of the stories task, on which intermediate learners favored an interpretation consistent with the past adverbial. Taken together, results from these experiments indicate that despite group differences in accuracy, participants were all highly accurate at inferring temporal location, which I would also expect in heritage speakers. Variability in native speaker choices that deviated from predictions was echoed in learners' judgments, though to a lesser extent only significant in intermediate learners. Slabakova attributes this variability to pragmatic inferencing, in which non-native learners seem to have some degree of competence.

4. Aspectual coercion

Aspectual coercion is a combinatorial semantic operation by which the default temporal reference for an event is overridden by extraverbal information to derive an interpretable reading. In English, aspectual coercion is chiefly employed to derive habitual or iterative readings from eventive predicates. While an eventive predicate is [+punctual], meaning it is completed, adding a temporal adverbial that is semantically incompatible with the aspectual features of the verb will lead to a necessarily iterative or habitual reading. Take the following example from Piñango, Zurif, & Jackendoff (1999):

(15)   a. The girl slept until dawn. [+durative, +durative]

         b. The girl jumped until dawn. [+punctual, +durative]

While (15a) is interpreted through *transparent* syntactic composition since *sleep* is durative and thus aspectually compatible with the adverbial, (1b) necessitates an iterative reading because

of the semantic incompatibility of the event *jump* and the durative adverbial *until dawn.* Re-analysis is not dependent on any one piece of temporal information nor any morphosyntactic marking, but instead relies on integration of semantics and syntax. While the means of aspectual coercion are debated, it is clearly not an exclusively syntactic phenomenon.

4.1 Proposed explanations

Brennan & Pylkkänen (2008) summarized the proposed hypotheses about the representation of aspectual mismatch:
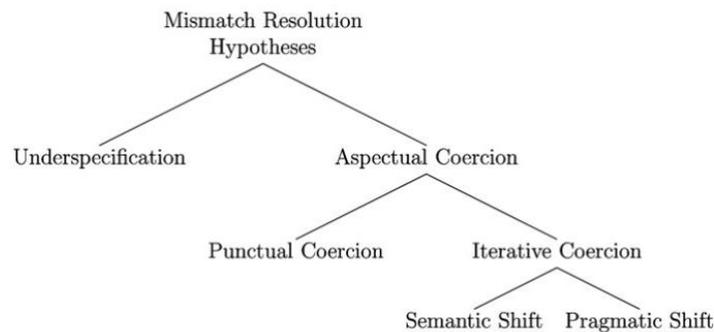
Figure 4.1 Mismatch resolution hypotheses

Of note are the two explanations for iterative coercion: Semantic vs. pragmatic shift. Semantic shift, originally proposed by Pustejovsky (1995), relies on the inherent properties of the verb and aspectual mismatch with adverbials is resolved through a semantic coercion operation. According to de Swart (2000), this coercion operator is only introduced if an event description (a state, event, or process) does not meet the input requirements of an aspectual operator or an aspectually sensitive tense operator. Thus, the coercion operator is used to modify the eventuality description in order to satisfy the constraints set by the aspectual operator. To illustrate without adverbials:

(16)   a.   Eve reads poetry.

         b.   Todd calls the cops.

In English, the Simple Present tense (*not* including the historical present or reporters'

present) requires a stative predicate as input because of the Boundedness Event Constraint.

However, the eventuality description of (2a) is a process, while that of (2b) is an event. This

mismatch triggers the coercion operator, which in turn triggers an aspectual transition of the

event into a state to fulfill the requirements of the Simple Present. Iteration and habituality are

two possible transitions, and so the event of *calling the cops* is understood as repeating in order

to express the event as a state.

Conversely, pragmatic shift (Dölling, 2003) is a two-step process: 1) The mismatching VP

and time-span adverbial semantically compose but result in an anomalous reading, and 2)

Pragmatic competence resolves the anomalous reading by coercing the VP. De Swart's

explanation also relies on both semantic and pragmatic knowledge: while the coercion operator

is triggered by semantic features, the aspectual transition is pragmatically determined, and thus

allows for variation in coerced meaning. According to de Swart, the reinterpretation itself is

felicitous depending on "linguistic context and knowledge of the world" (de Swart, 1998, p.

360), which may necessitate a second stage of processing. Brennan & Pylkkänen (2008) found

evidence supporting pragmatic shift given measures of magnetoencephalography (MEG). Prior

work had implicated the anterior midline field (AMF) in complement coercion (discussed in

section 4.3.1) and the right anterior temporal lobe in semantic anomaly detection. Brennan &

Pylkkänen gave participants 30 coerced and 30 control sentences all *beginning* with a temporal

adverbial, and 360 filler sentences all presented word-by-word. Following each sentence,

participants were asked to judge whether the sentence made sense or not within 4s. The authors

hypothesized that the semantic anomaly detection should precede the coercion in order to support the pragmatic shift; results indeed showed increased activity in the right anterior temporal lobe 340 ms after verb onset followed by increased activity in the AMF 100 ms later. Detection of meaning anomaly appears to be followed by meaning shift.

4.2 Aspectual coercion in bilinguals

There is thus evidence for an additional computational load associated with re-analysis. In monolinguals across languages, the processing cost of aspectual coercion has been attested in *offline* self-paced reading tasks (SPRT) (Todorova et al., 2000; Sampaio & França, 2018) and lexical decision tasks (Piñango, Zurif, & Jackendoff, 1999; Piñango, Winnick, Ullah, & Zurif, 2006), and *online* eye-tracking paradigms (Townsend, 2013) and ERP studies (Yano, 2017).

Whether resolving aspectual mismatch involves the syntax-semantics interface or pragmatic competence, bilinguals have been found to show difficulty in both areas. Of interest are bilinguals who speak languages with different resolution strategies for aspectual mismatch. In the first study to investigate processing of aspectual coercion in L2 learners, Chan (2012) tested comprehension of iterative and *grammatical* coercion constructions in L2-English learners of various L1 backgrounds, including Mandarin. According to the author, grammatical coercion is exemplified by the English semelfactive progressive, as in *Tom was kicking the ball*. Semelfactives alone are [+punctual, -telic], while semelfactive progressives are [-punctual, -telic]. Specifically, these constructions denote iterative action-in-progress— not compatible with punctual adverbials. In contrast to simple past semelfactive *kicked,* which encodes a single completed event iteratively coerced by a durative adverbial, the semelfactive progressive

encodes iteration by default, without requiring a durative adverbial. To illustrate with the

conditions of the study:

(17)   a. At noon the kid jumped into the pool. <= baseline
         b. All day the kid jumped into the pool. <= *Iterative coercion*
         c. At noon the kid was jumping into the pool. <= *Grammatical coercion*
         d. All day the kid was jumping into the pool. <= baseline

The grammatical coercion operation is triggered by the mismatch between the punctual

adverbial *at noon* and the durative nature of the iteration inherent to the progressive *jumping*. To

satisfy the input requirements of the punctual adverbial, the repeated events implied by *jumping*

are coerced to a single iteration in a series.

Participants were given a word-by-word SPRT to measure processing of the four conditions

given in (3). When controlling for English proficiency, English native speakers conformed to

predictions that both iterative and grammatical coercion conditions yielded significantly slower

RTs than control conditions, particularly in the V+2 (the second word after the verb) region.

Only the L1-Mandarin group showed significantly *faster* RTs in the iterative coercion condition

than in the respective baseline. Chan attributes this RT advantage to influence from the L1. In

Mandarin, past is denoted by perfective aspect marker *-le*, which favors bounded, telic contexts.

Since semelfactives are inherently atelic, co-occurrence with *-le* is only felicitous when bounded

by a *for X time*-adverbial or verbal classifier phrase like *once*. This result is echoed in Chen

(2009), who found Chinese English L2-learners showed no sensitivity to the ungrammatical lack

of tense morphology following an auxiliary verb but demonstrated knowledge of them on

explicit English proficiency tests. Prior work additionally found *-le* + semelfactive constructions

appeared more frequently in native Chinese corpus data with boundedness adverbials than

without (Xiao & McEnery, 2004), and thus durative adverbial + semelfactive sentences are more common in Mandarin than in English. Because of this frequency effect, we might expect to find correlations between RT advantage in iterative coercion and Mandarin exposure, which I will test. Ultimately, these results provide evidence of transfer from the L1, possibly facilitated by heightened exposure to salient constructions. While language dominance is not noted, Chinese bilingual participants were attending college in the U.S., averaged 23.81 years old, and started learning English at average age 10.48. This profile suggests these participants were late bilinguals, and the processing advantage may disappear in heritage speakers and early bilinguals due to attrition or incomplete acquisition (Keating, Jegerski, & Vanpatten, 2014; Montrul, 2002).

Montrul (2002) tested comprehension of inchoative coercion in L1-Spanish bilinguals grouped by age of exposure to English (simultaneous, early child, late child). We can assume the simultaneous and early bilinguals are heritage speakers. In Spanish, Preterite and Imperfect markers *-ó* and *-a* encode perfectivity and imperfectivity respectively. Verbs marked with Spanish Preterite *-ó* coerce a perfective, punctual reading from an imperfective, stative event:

(18) Juan sabia   la verdad.
Juan know-IMP the truth
"Juan knew the truth"

(19) Juan supo   la verdad.
Juan know-PRET the truth
"Juan found out the truth"

According to prior work, the Preterite/Imperfect distinction with stative verbs is often eroded or not present in L2-dominant Spanish bilinguals. To test the robustness of this finding, Montrul (2002) gave participants a morphological choice task, an elicited storytelling task, and two

comprehension tasks: sentence conjunction judgment and truth value judgment. I will focus on the results of the latter tasks, though the production tasks showed that simultaneous bilinguals had significantly lower accuracy in proper use of both the Imperfect and Preterite, and struggled most with coercing a perfective reading from a stative. The sentence conjunction task gave participants two clauses that were either logical or illogical together given the use of either the Preterite or Imperfect in the first clause, where the IMP made it logical and PRET illogical. For an example with a stative, where (20) uses IMP and (21) PRET:

(20)   La clase <u>era</u> a las 10 pero empezó a las 10:30.

The class was at 10 but started at 10:30

(21)   *La clase <u>fue </u>a las 10 pero empezó a las 10:30.

The class was at 10 but started at 10:30

Simultaneous bilinguals and early L2 learners were significantly more likely to accept the illogical uses of the Preterite as logical, especially with statives. The truth value judgment task used stories that supported either a stative or eventive interpretation, which was evaluated in a true or false comprehension question which used either the Preterite or Imperfect. The IMP was true in stative and habitual stories and false in eventive stories, while the opposite was the case for the PRET (e.g. Marcos conoció/conocía a Susana, where the verb means either *met* as in PRET or *knew* as in IMP, respectively). This task yielded similar results: Simultaneous and early L2 bilinguals were significantly less accurate than monolinguals and late bilinguals with habitual stories with both PRET and IMP, and with eventive interpretations. While there were apparent group effects, the authors admit a large amount of individual variation within groups, particularly in proficiency, as can be seen by 3 participants across groups who scored within NS competence,

26

and 4 across groups who showed *no* indication of the PRET/IMP distinction. Results from this study support differential acquisition of the L1 with English contact, and lead Montrul to conclude that with regard to aspectual semantic interpretations, bilinguals may never converge on NS grammar regardless of age of exposure.

4.3 Mismatch resolution in Mandarin

In this section, I will outline triggers and resolution strategies for two types of coercion in Mandarin and draw comparisons with English. I will also summarize studies that test processing, interpretation, and acceptability of these constructions in bilinguals.

4.3.1 Complement event coercion

Prior work (Xue, 2016) has shown cross-linguistic influence in comprehension of a different type of coercion acceptable in English but not always in Mandarin: *complement event coercion* (CEC). In English, this construction looks like the following, from Xue, where X = verb that requires an event-denoting complement:

(22) a. X [_ VP]
    He *began* (=X) [reading/to read the book]

  b. X [_ NP$_{eventive}$]
    He *began* [his reading]

  c. X [_ NP$_{entity\text{-}type}$]
    He *began* [the book]

NP objects as in (22b,c) are coerced into events, so English verbs that require event-denoting complements (also including *try, finish, attempt, regret*) can not only take eventive VPs as objects, but also coerced NPs. Furthermore, results from English language corpus analysis

(Lowder & Gordon, 2016) showed that for nine event-selecting verbs embedded in relative clauses, entity-type coercion was significantly more likely to occur in ORCs than in SRCs. For example, with the verb *finish*:

(23)   SRC: *Rose, who on Monday* **finished** *a five-month prison term for tax felonies, was banned from baseball.*

ORC: *Emily asked as she threw the comic book she had just* **finished** *toward the stack.*

Lowder & Gordon (2016) found this pattern reflected in results from a separate eye-tracking task, in which native English speakers read CEC coercion in ORCs significantly faster than in SRCs. ORCs showed no differences from controls in RT at the embedded verb, which supports that salient constructions in input facilitate processing, even with coercion constructions. CEC in SRCs, which occur less in natural input, were more difficult to process.

In Mandarin, events cannot be coerced from an entity-type NP such as *the book*, so sentences like (3c) are not grammatical. To test L1 Mandarin comprehension of CEC and whether length of residence (mean = 11 years), exposure (to both languages before and after immigration to the L2 environment), or self-reported proficiency would play a role, Xue (2016) compared Mandarin-English bilinguals to English and Mandarin monolingual controls. Bilinguals were given an acceptability judgment task in both languages. Items included entity-type coercion sentences acceptable in English but not Mandarin, non-coercion sentences acceptable in both English and Mandarin, and coercion exception sentences also acceptable in both English and Mandarin.

(24)   Coercion: *ta liji jiu houhui-*le *naxie pinglun*

He regretted the comments immediately.

28

(25)   Non-coercion: gongsi-de jingli houhui fan-*le* na-ge cuowu

                               He regretted making the comments immediately.

(26)   CEC exception: wo zhongyu wancheng-*le* wo-de lunwen

                               I finally completed my thesis.

Results suggested influence from English that correlated with L2 proficiency: Highly L2-proficient bilinguals accepted both the English entity-type NP coercion constructions *and* their ungrammatical Mandarin equivalents, regardless of exposure and length of residence. Recall from section 2 that English dominance also determined likelihood of accepting ungrammatical forms in the non-dominant language, while proficiency in the non-dominant language only determined correct acceptability of conventional forms. While dominance is not explicitly measured in Xue's study, results suggest that proficiency in the *dominant language* could play a role in accepting ungrammatical forms in the non-dominant language, rather than proficiency in the non-dominant language. The bilinguals in this study were not heritage speakers but immigrated to the U.S. from China after age 11. It is possible that for late bilinguals, relative proficiency can be used as a proxy for dominance; but again, the notion of relativity invokes dominance as a more comprehensive measure.

4.3.2 Inchoative coercion

Mandarin makes prominent use of *inchoative* coercion. Inchoative coercion derives a change-of-state (COS) predicate from a stative predicate, in contrast with iterative or habitual coercion, which derive repeated states from punctual predicates. Some languages, like Quechuan, can morphologically mark state-denoting words to derive a COS-denoting meaning. Other languages,

like English and Mandarin, must rely on context. In English, stative predicates can be shifted into COS predicates when combined with a telic adverbial *within/in x time*:

(27)   a. Within two minutes, she was asleep.

         b. Within a half hour, they distrusted the politician.

Likely because inchoative coercion is not salient in English, the only study I could find that tests this construction in English is Brennan & Pylkkänen (2010), which claims that it was the first to do so. MEG data was again collected from a SPRT, and the authors found increased neural activity in the ventromedial PFC previously implicated in coercion. This activity appeared 300 ms post-stimulus onset— 100 ms sooner than the iterative coercion constructions previously tested in Brennan & Pylkkänen (2008), discussed earlier. Consistent with earlier results, another spike in activity was found 100 ms later. While inchoative coercion is triggered and resolved sooner than iterative, both support a two-step pragmatic shift.

Inchoative coercion in Mandarin is triggered by context, rather than adverbials. Recall that the perfective marker *le* by default encodes perfectivity in dynamic situations and inchoativity in stative situations. In dynamic situations such as (5) below, mismatch between the default perfective interpretation in the first clause and the context in the second clause is resolved by coercing an inchoative reading of *le,* which by default requires stative situations. To satisfy this requirement, the situation type of the *le*-clause is then coerced from dynamic to stative. For example, from Wang (2008):

(28)   *Wo xiu zhe liang   che le, Yiqian yizhi   bu yuanyi xiu li*

            I   repair this   car LE, before always not want repair

            "I want to fix the car, (but) I did not want to before."

While the default interpretation of *le* in the dynamic first clause is perfective "I fixed the car", the context provided by the second clause mismatches default *le* and triggers inchoative coercion as a resolution strategy. In turn, "I fix the car" (now [-perfective]) is coerced from an accomplishment predicate into a stative for *le* to encode inchoativity by adding "shift operator" *want*, yielding "I want to fix the car". According to Wang, shift operators are implicit and necessary for well-formedness. Once the clause is a stative, inchoative coercion works to derive a COS meaning, namely that the speaker has *started* to want to fix the car but did not always.

Studies of processing of inchoative coercion in bilinguals are seemingly non-existent, let alone in Mandarin bilinguals and heritage speakers. Wang (2008) tested processing of inchoative coercion constructions in Mandarin native speakers from China who were living in the U.S. for a mean of 4.1 years and whose average age was 31.1. While no other language background is provided for these participants, it seems as though they are on average late bilinguals likely in the process of restructuring their L1 due to recent immigration. Participants were given a self-paced reading task with four constructions: coercion, in which context coerced dynamic *le* into a stative reading, two concord constructions, in which dynamic and stative *le* matched the following context, and controls.

5. The current study

This experiment seeks to compare the processing cost for aspectually coerced sentences in L1 Mandarin between groups of English bilinguals varying by dominance. Specifically, I seek to answer the following research questions, reiterated from the introduction:

1. How does L2 dominance vary in heritage speakers, and what other linguistic variable(s) correlate with high L2 dominance? This addresses a source of variation in

heritage speakers and would help to develop a more comprehensive profile of these bilinguals.

2. Does English dominance contribute to the acceptability of Mandarin aspectual coercion? If so, is accuracy or processing affected? This elucidates the effects of unbalanced dominance on linguistic behavior in adulthood.

3. Does language-specific salience of coercion constructions affect accuracy of temporal interpretation? This serves to test effects of input, often reduced in heritage speakers.

4. Is resolving aspectual mismatch in fact a function of pragmatic competence?

I make the following predictions with regard to these questions:

1. Heritage speakers can be balanced or English-dominant. No prior studies have accounted for balanced heritage bilinguals and this would serve to distinguish their profiles. High English dominance will be tied with low Mandarin proficiency and discontinued use of the heritage language.

2. English dominance will not affect accuracy of judgments of acceptable constructions in Mandarin (Conditions A and B) but will affect accuracy with unacceptable constructions. Specifically, English-dominant heritage speakers will be more likely to accept unacceptable coercion constructions in Mandarin than balanced bilinguals and Mandarin-dominant bilinguals. English dominance will lead to higher processing cost in Mandarin-specific coercion constructions.

3. Because of attrition and incomplete acquisition, coercion constructions that are not salient in Mandarin (Condition A) will yield low accuracy with temporal interpretation in heritage speakers.

4. Higher pragmatic competence in Mandarin will yield higher accuracy with temporal interpretation of coerced sentences. Acceptability will not be affected.

To address these questions, I will conduct a timed listening task with three types of Mandarin coercion constructions varying by language-specific acceptability and salience.

5.1 Participants

Since a primary goal of this study is to attempt to account for variation in heritage grammars, and since the processing cost of coercion constructions in monolinguals has already been widely documented, monolingual controls are not necessary. Instead, Mandarin-dominant late bilinguals serve as baseline controls. Bilinguals will be recruited through virtual mailers from CUNY campuses and EFL schools in NYC to account for late bilinguals more likely to be Mandarin-dominant. All participants will be screened through email in English to ensure enough fluency to be able to understand the comprehension questions, which will be in written English. The screening will gather biographical data such as when they arrived in the US and age and length of exposure to English and Mandarin. Heritage speakers are defined as having been born in the US or having arrived by the age of 5, and late bilinguals are defined as having moved to the US after age 17. Upon completion, participants are asked to email the experimenter with a randomly generated code presented once the results are submitted to the server, along with a means for virtual payment of $10 either through Venmo, Zelle, CashApp, or PayPal.

In order to assess dominance, participants will be given the *Bilingual Language Profile* (see Appendix A). The BLP was explicitly designed to assign a composite dominance score, which falls between -200 and 200; the further away from 0, the more dominant in one language. This score will be used to group participants as Mandarin-dominant, English-dominant, or balanced. The BLP takes less than 10 minutes to complete. In order to develop a more comprehensive profile of heritage speakers, they are grouped by dominance given the BLP. Scores over 75 are considered English-dominant, under -75 are considered Mandarin-dominant, and between -75 and 75 are considered balanced. Below is a possible distribution of dominance across groups in an ideal yet realistic sample of 75 bilinguals, so there could be a comparable numbers across groups.

Table 5.1 Groups of bilinguals

	Heritage speakers (35)	Late child bilinguals (15)	Adult bilinguals (25)
English-dominant (30)	15	7	8
Balanced (30)	15	5	10
Mandarin-dominant (15)	5	3	7

More details about these groups can be given with self-reports of proficiency in both languages given in the BLP, particularly comprehension proficiency. Length of residence and weekly language use are also available. For proficiency and length of residence, range, mean, and standard deviation should be calculated. For weekly language use, range should be sufficient. In addition, basic biographical information for age of all participants should be averaged with range and SD taken. Counts for gender should also be included.

Since self-reports of proficiency have been found to correlate with objective measures (O'Grady et al., 2009; Keating et al., 2015; Montrul & Ionin, 2012; Treffers-Daller 2010; Dunn, 2009; Mikhaylova, 2011) and are reflected in accuracy of temporal interpretation tasks (Slabakova, 2015), independent tests of English and Mandarin proficiency will not be given. This study also does not seek to measure effects of proficiency, but self-reports of proficiency in speaking, understanding, reading, and writing on a 1-6 Likert scale are included in the BLP and will be considered in analysis.

Participants will also take a pragmatic competence assessment in Mandarin (see supplemental Appendix B). If successfully deriving coerced sentences necessitates pragmatic knowledge, we might expect to find higher pragmatic scores correspond to better accuracy, particularly since Mandarin temporal interpretation also relies on pragmatic knowledge. In ESL/EFL learners of various L1 backgrounds, Roever (2006) found that knowledge of English *speech acts* and *implicature* increased with proficiency, while knowledge of English *routines* increased with exposure. This finding contradicts an earlier study, which found that knowledge of ESL implicatures was instead tied to exposure (Bouton, 1999). Roever attributes this difference to different populations: The Bouton study used ESL learners all living in the L2 environment, while Roever also included EFL learners who had no exposure to the L2 environment. This test could thus also elucidate the relationship between pragmatic competence and dominance, which includes both exposure and proficiency. There is a surprising lack of work that has attempted to explore this relationship, let alone in the L1 of heritage speakers.

I will use the assessment developed in Li (2018), originally created to measure L2-Chinese pragmatic comprehension. This test consists of 39 aurally presented short dialogs followed by four multiple-choice options for implied meaning. The aural presentation of stimuli, simple

35

answer selection (hitting 1, 2, 3, or 4 on the keyboard), and thus no need for writing make this

test ideal for heritage speakers who tend to maintain their L1 orally rather than through text. The

multiple choice also eliminates the requirement for a keyboard with Chinese characters. To

observe distribution of test scores across groups, raw score means, ranges, and SD will be

calculated per group. To evaluate the relationship of dominance with pragmatic competence, a

Pearson correlation between pragmatic test scores (between 0 and 39) and continuous dominance

scores (between -200 and 200) can be taken. In order to derive this correlation, pragmatic test

scores must be normalized while dominance scores must be standardized since they are on

different scales. A strong correlation would support a relationship between pragmatic

competence in the L1 and stronger L1 dominance.

5.2 Materials

All materials are presented with the online experiment platform pcIBEX (Zehr & Schwartz,

2018). The full test should take about one hour to complete.

5.2.1 Relative fluency

In addition to the BLP, an objective dominance measure will be given in the form of an

elicited narration task to derive a relative fluency (RF) score from both languages, based on

Stevens (2019). Participants are asked to describe two short silent films in randomized order:

Charlie Chaplin in "the Lion's Cage" and Dolby's "Silent." Both films are less than 3 minutes

long. One will be presented before the experiment and the other one after the experiment, but the

first one will include a prompt to describe the events in Mandarin and the second in English. This

order will be maintained to keep participants' Mandarin knowledge more accessible during the

experiment. Participants will be able to start a five-minute timer when they are ready to begin

recording. The RF score will be calculated with the following formula: (Mandarin WPM –

English WPM)/(Mandarin WPM + English WPM) * 100, excluding false starts and code-

switches. A score of 0 indicates balanced fluency, while a negative score indicates faster

Mandarin speech rate and a positive score indicates faster English speech rate. According to

Stevens (2019), this score highly correlated with the continuous scores from the BLP.

5.3.2 Modified self-paced listening task

The three coercion constructions in Mandarin outlined in Table 5.2 will be tested through a

self-paced listening task. Self-paced listening tasks typically present stimuli word-by-word or

phrase-by-phrase while the listener pushes a button to advance to the next part. Button presses

are recorded to provide time taken to process each part. Longer listening times reflect greater

processing cost. Because listening tasks tend to be unnatural—particularly in Mandarin where

tonal boundaries are also meaningful but vary by sentence—sentences will be presented as a

whole. Thus, this experiment is a modified version. Instead of collecting listening time per word,

time to make a decision on the choice task is measured: Response time is collected from the

timepoint that the audio finishes its first play to the timepoint that temporal interpretation choice

is made.

Before beginning the experiment, instructions are presented on-screen. Participants are told

that they will listen to sentences in Mandarin and determine 1) When the event took place, and 2)

How grammatically acceptable the sentences seem. They are told that the audio may sound

unnatural and to focus on the content. Instructions also indicate that some items will be

accompanied with an audio control widget that will allow them to replay the sentence, while

other items will not allow replay. This is to ensure adequate attention to all stimuli. They are also

told that once a choice is selected, they will not be able to change their response or replay the audio. When participants are ready to begin, they click Continue.

Each page of the experiment displays the same rules to help decide choices. Participants are told to select an interpretation of BOTH if the event took place in the past and still holds in the present, EITHER for events that could have taken place in the PAST or PRESENT, and NEITHER for sentences that are unacceptable. They are told to press the spacebar when they are ready to hear the sentence. Once the audio has played once, the five choices for temporal interpretation are displayed on the screen underneath the audio player for target items, and underneath the instructions for fillers. An example is shown below.

Figure 5.1 Sample of display during trial

progress

Select **BOTH** for events that took place in the past and still hold in the present. Select **EITHER** for events that could have taken place in the PAST or PRESENT.

For sentences that are unacceptable, select 'NEITHER'.

When you are ready to begin, **press the spacebar.**

▶ ━━━● 0:03 / 0:03 🔊 ━━━━●

PAST PRESENT BOTH EITHER NEITHER
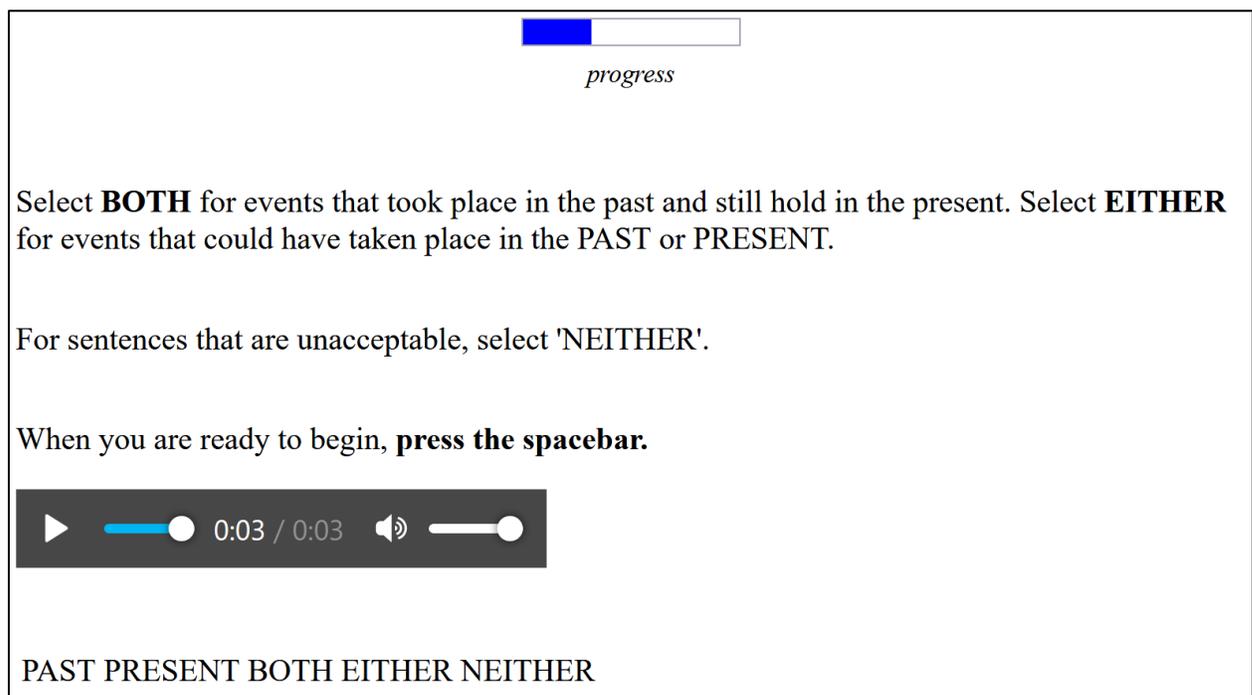
Before selecting a choice, participants will still be allowed to replay the sentence to their satisfaction. Once they select a choice, the audio widget and the choices will disappear, and the acceptability judgment is shown. The acceptability judgment is elicited on a 6-point Likert scale

ranging from 1 (completely unacceptable) to 6 (completely acceptable). Once a choice of acceptability is selected, a button labeled "Next" appears and participants can click it to move on to the next item.

All test items (10 in each condition, 30 altogether) and paired controls (30) but none of the non-relevant fillers (60) display the audio control (see supplemental Appendix C). So, half of total presented stimuli will be able to be replayed before making a temporal interpretation choice. Target items make up 1/3 of total stimuli. Fillers are acceptable Mandarin sentences with various aspectual marking. Controls will be non-coerced equivalents, after Xue (2016). For condition A, this means that the state is bounded with an adverbial. For B, this means that the durative adverbial is replaced by a quantized object. For C, this means that the NP_{entity} is replaced by a VP or NP_{event}. Conditions are summarized in Table 5.2, where MC is Mandarin Chinese and ENG is English.

Table 5.2 Sentence types with language features

Condition	Salience	Acceptability	Example
A. State$_{[-bounded]}$ + *-le*	MC > ENG	MC: Yes ENG: No	Zhangsan yang-le yitiao jinyu Zhangsan keeps a goldfish.
B. VP$_{[-durative]}$ + *-le* + Adv$_{DUR}$	MC > ENG	MC: Yes ENG: Yes	Ta kesou-le liang fenzhong He coughed for two minutes.
C. VP$_{eventive}$ + NP$_{entity}$	MC < ENG	MC: No ENG: Yes	*Ta liji jiu houhui-le naxie pinglun He regretted the comments immediately.

Condition A reflects a construction that is not salient in Mandarin. According to Xiao & McEnery (2004), *-le* appears significantly less frequently in stative situations than in dynamic

situations in natural Mandarin speech, which serves as the baseline input for naturalistic learners such as heritage speakers. When states with -*le* are bounded by a PP or past adverbial, they *tend to* no longer hold at time of utterance, like in English, e.g. *I had liked you for two years*. When states with -*le* are unbounded, they *tend to* still hold in the present. In English, these sentences would require a temporal adverbial, e.g. # *I have liked you*. Both of these types will be included. Therefore, despite the infrequency of both of these constructions in Mandarin, I expect unbounded states with -*le* to cause more processing difficulty and lower accuracy in both temporal interpretation and probability of acceptability in highly English-dominant bilinguals than controls.

Condition B reflects iterative coercion, which is infrequent and difficult to process in English. While not necessarily common in Mandarin, events must be bounded by some means, like an adverbial. According to Xiao & McEnery (2004), semelfactives occurring with a durative adverbial are quantitatively more common in Chinese than in English, which Chan (2012) found to give Chinese bilinguals an advantage in processing iterative coercion in English. I expect English-dominant bilinguals to lose this advantage, but for accuracy at inferring temporal location to not be affected.

Finally, condition C reflects Complement Event Coercion (CEC), which is unacceptable in Mandarin with entity-type NPs but acceptable in English. I expect English-dominant bilinguals to show a processing advantage relative to Mandarin-dominant bilinguals. I also expect that English dominance will lead to accepting both unacceptable experimental items and their respective non-coerced controls.

Predictions within the English-dominant group and between groups are summarized below with our dependent variables of interest. Recall that accuracy of interpretation is derived from the temporal interpretation task and probability of acceptability is derived from the acceptability judgment task. MD denotes the Mandarin-dominant group and ED the English-dominant group.

Table 2.3 Predictions for sentence type by group

Condition		Response time	Accuracy of interpretation	Probability of acceptability
A.	ED vs. MD:	MD < ED	MD > ED	MD > ED
	ED: Targ vs. Controls:	T > C	T > C	T < C
B.	ED vs. MD:	MD < ED	MD = ED	MD = ED
	ED: Targ vs. Controls:	T < C	T = C	MD = ED
C.	ED vs. MD:	MD > ED	MD = ED	MD < ED
	ED: Targ vs. Controls:	T = C	T = C	T = C

6. Discussion of possible results

In this section, I will outline methods for statistical analysis for each test in order to review possible results from the experiment had it been run in a sample of 75 participants as noted in section 5.1. Implications will be considered in case my hypotheses are rejected or borne out.

6.1 Predictions of accuracy

Here I will consider how dominance and pragmatic competence could affect accuracy on the temporal interpretation choice task. Accuracy data is chiefly collected to measure responses to Condition A, since temporal location is explicit in B and C. Since the Mandarin-dominant late bilinguals' responses are used as baseline controls, responses on the temporal interpretation

41

choice task will be compared to theirs in order to derive an accuracy score. After Slabakova (2015), counts of interpretation by controls will be calculated per condition to derive a percentage for correct choice. So, if the 7 controls choose a Present interpretation for Condition A 40 times out of a possible 70 (58% of the time) and Both 20 times (29%), responses of Present or Both would be considered correct, while only Past would be incorrect.

Since we cannot assume a normal distribution or equal sample sizes, data will be fit to a generalized linear mixed effect model (GLMM) to test effects of dominance and condition. A GLMM is also appropriate because an accuracy score is binomial: either correct or incorrect. Independent variables include fixed effects of Group (English-dominant, Mandarin-dominant, and Balanced), Condition; we can also test for an interaction between Group and Condition. Random effects of Item and Subject are also accounted for with a GLMM. If the full model yields non-significant results, I will exclude Group and the interaction of Group and Condition from the model to test for an effect of Condition alone. Significant results in Group would necessitate a post-hoc test to detect specific differences between groups. Specifically, the Tukey-Kramer post-hoc test allows for comparisons between groups with unequal sample sizes.

I have predicted significant Group by Condition interactions: English-dominant bilinguals will be significantly less accurate at inferring temporal location in Condition A than Mandarin-dominant bilinguals. I do not expect significant group differences between the Balanced group and the English or Mandarin-dominant groups across conditions. If my predictions are borne out, then the inclusion of a Balanced group is important in heritage language research; combining the Balanced group with the other groups would lead to more variability in responses. In addition, English dominance would show an effect in Mandarin accuracy despite high proficiency, further necessitating a use for dominance as a grouping measure for heritage speakers. If null results are

found, it is possible the model is underpowered and more stimuli or participants should be included, as in the case of a Type II error. Null results could also indicate that dominance simply does not affect accuracy of temporal interpretation, which is also possible since I expect all participants to show high accuracy despite differences in Group and Condition. In this case, traditional linguistic variables such as proficiency or age of exposure should be investigated. Since I have access to these data, I can regroup participants and run more GLMMs. Significant results would indicate that these variables are better proxies for accuracy than dominance.

To test for effects of pragmatic competence on accuracy of temporal choice, I will simply run a Pearson correlation with normalized raw scores. Correlations of pragmatics scores with global accuracy and accuracy per group will be measured. Since pragmatic competence and accuracy in part reflect proficiency, we could expect a linear relationship between the two. I am particularly interested in Condition A, in which correct interpretation relies on pragmatic inference. To decide whether the difference between r_A and global r is significant, I will use Fisher's r-to-z transformation. If there is a significant effect of pragmatic competence on global accuracy, then it should certainly be considered in future work and emphasized in language learning contexts with heritage speakers. If accuracy with Condition A is significantly more correlated than global accuracy, then we have psycholinguistic evidence in support of pragmatic knowledge contributing to Mandarin temporal interpretation. If no correlations are found, then it is possible that the test is not valid at measuring the pragmatic competence necessary in interpretation. It is also possible that there is simply no effect of pragmatic knowledge. Nonetheless, if it is shown that heritage speakers score lower than other groups on the pragmatic test, then we can offer evidence that pragmatic knowledge is weakened in the home language and contribute to research on the heritage grammar.

6.2 Predictors of acceptability

For acceptability judgments, mean ratings and standard deviations are calculated per group and condition. Differences by group and condition can be analyzed using the Kruskal-Wallis H test for ordinal data, which includes acceptability judgments. Use of this test does not rely on normal distributions, and if the shape is different by group or condition then mean ranks are compared instead of median ranks. First, I will compare judgments of test items by condition within groups. If $p(H)$ is found to be significant, then at least one condition yielded significantly different acceptability judgments. Since conditions have the same number of items, post-hoc Pairwise Mann-Whitney U tests can be used to determine specific differences. Across groups and within conditions, a significant Kruskal-Wallis result would indicate that at least one group judged the condition differently. Since we are not assuming a normal distribution in our sample, post-hoc Nemenyi tests must be used to determine specific group differences.

I have predicted that English-dominant participants are significantly more likely to accept unacceptable Mandarin coercion constructions than Balanced and Mandarin-dominant participants, in addition to the non-coerced controls. Thus, I expect no significant results within the English-dominant group in Condition C, in which both unacceptable CEC coercion and acceptable controls should be as likely to be accepted. I expect Mandarin-dominant bilinguals to show a significant difference between accepting CEC and controls. Across groups, I expect a significant difference between groups in Condition C, particularly because the Mandarin-dominant group will be more likely to correctly reject the construction than both other groups. If my predictions are supported by results, then there is more psycholinguistic evidence for transfer of acceptable structures from the dominant L2, already shown in the English-dominant Spanish heritage speakers in Zyzik (2019) and the English-dominant L2 Mandarin learners in Xue

(2016). The effect of dominance on acceptability judgments is thus also shown in Mandarin-English bilinguals. If my predictions are not confirmed, then it is still interesting. Type II error is still possible since I used the same stimuli for Condition C as Xue (2016) but with less items. However, a larger sample size used in this study should make that unlikely. Instead, it is possible that heritage speakers retain awareness acceptable constructions from their home language, which would provide evidence against the results in Zyzik (2019) and merit further work.

To test an influence of pragmatic knowledge on acceptability judgments, test scores will again be normalized while judgments will be converted to a dichotomous variable 1 or 0 to allow for a point-biserial correlation. I have opted for this correlation method instead of Spearman's rank-order correlation because I believe it could be more informative: 1 would indicate agreement with the Mandarin-dominant judgments while 0 would indicate disagreement. This way, higher pragmatic competence in Mandarin would predict higher agreement with baseline controls. However, I do not expect a relationship between pragmatic competence and correct judgments. If there is in fact a correlation, then determining acceptability could also be a function of pragmatic knowledge, which has not yet been revealed in any work.

6.3 Predictions of response time

Response time (RT) will be calculated for the subset of items with correct responses on the temporal choice task. Time is taken from the end of the audio presentation to the selection of the temporal interpretation, which includes replays. Formulaically, it is (time of choice – end of audio), all included in the results file provided by pcIBEX. Mean, median, and SD for raw RT will be reported per group and condition. However, since length of stimuli varies, we must again control for Item as a random effect, in addition to Subject. But unlike accuracy, RT is a

continuous variable so we can use a linear mixed effect model (LMM) instead of a GLMM. In order to make RT linear enough for an LMM analysis, it must be log-transformed. Therefore, the DV is logRT; fixed effects are Group, Condition and their interaction, and random effects are Item and Subject. Like with accuracy, if this model yields non-significant results, Group in addition to the interaction of Group and Condition will be removed from the model to observe if Condition alone is a better fit. If instead the model yields significant results, the Tukey-Kramer post-hoc test can again be used to detect specific differences.

I have predicted significant interactions between Group and Condition: The English-dominant group will show significantly longer RTs than the Mandarin-dominant group with Conditions A and B, but significantly shorter RTs than the Mandarin-dominant group with Condition C, which occurs frequently in English. If these hypotheses are supported by results, then transfer from the dominant L2 can both facilitate and weaken processing in the L1 given particular salience and acceptability of the structure in English. If there are only significant differences in Condition A, then constructions not salient in Mandarin input may have never been acquired in the case of heritage speakers or had attrited with English-dominance. These constructions would then become harder to process automatically later in life—especially if there were no differences in the temporal interpretation task, which reflects knowledge of the rules. However, I would not think it possible for Condition A to yield significant differences without B since the underlying reasons are the same. If C was the only condition to give the English-dominant group a processing advantage and they had accepted the Mandarin constructions, it could reflect a lack of knowledge of the rules of CEC in Mandarin. An objective proficiency test or elicited narration task could offer elucidation.

6.4 Limitations

The most obvious limitation is that this study was not able to be run due to the COVID-19 pandemic. Without access to proper recording equipment and pay for a native speaker, audio was artificially synthesized using https://soundoftext.com. Future work should use more natural sounding speech recorded by a native speaker who could incorporate appropriate tones. A pilot study with Mandarin-dominant late bilinguals would also be important to test stimuli. In addition, the variables included in the BLP should be considered independent of dominance. To reduce the chance of a Type II error, more stimuli should be included per condition.

While I did not seek to test the effect of pragmatic competence on processing, it could be done by correlating normalized test scores with RT. But raw RT would need to be transformed depending on the shape of the correlation with normalized test scores. Log-transformation could be warranted, but so could square-root transformation, Box-Cox transformation, and square root transformation. The obvious solution to this problem is to collect data.

Lastly, future work that I hope to pursue would objectively measure *relative salience* in input given corpus data. For example, what percent of perfectives co-occur with states in English as opposed to accomplishments, and to what extent is that greater or less than in Mandarin? Computational methods can quickly gather such occurrences from, say, forums on the internet to reflect natural use.

6.5 Conclusions

Ultimately, this thesis has laid out the argument for using dominance as a grouping measure for bilinguals. English dominance can lead to deviations in both linguistic behavior and competence in the minority home language. In addition, this has been a first step in understanding the role pragmatic competence plays in Mandarin temporal construal and in

resolving aspectual mismatch. Heritage speakers make up a heterogeneous majority of bilinguals in New York City and much work remains to be done to untangle what exactly makes them unique from other bilinguals and monolinguals, let alone how to facilitate maintenance of the minority language. Hopefully, the experiment outlined in this thesis can be run and results can contribute to our limited knowledge of this population and their relationship with pragmatic knowledge and input.

Appendix A: Bilingual Language Profile.

We would like to ask you to help us by answering the following questions concerning your language history, use, attitudes, and proficiency. This survey was created with support from the Center for Open Educational Resources and Language Learning at the University of Texas at Austin to better understand the profiles of bilingual speakers in diverse settings with diverse backgrounds. The survey consists of 19 questions and will take less than 10 minutes to complete. This is not a test, so there are no right or wrong answers. Please answer every question and give your answers sincerely. Thank you very much for your help.

I. Biographical Information

Name _____ Today's Date _____/_____/_____

Age_____ ☐ Male / ☐ Female / ☐ Other Current place of residence: city/state_____ country_____

Highest level of formal education: ☐ Less than high school	☐ High school	☐
Some college		
	☐ College (B.A., B.S.)	☐ Some graduate school ☐
Masters		
	☐ PhD/MD/JD	☐ Other: _____

II. Language history
In this section, we would like you to answer some factual questions about your language history by placing a check in the appropriate box.

1. At what age did you **start learning** the following languages?

English
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ☐ 20+☐

Chinese
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ☐ 20+☐

2. At what age did you **start to feel comfortable** using the following languages?

English
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
As early as I 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ☐ 20+☐ not yet
can remember

Chinese
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
As early as I 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 ☐ 20+☐ not yet
can remember

3. How many years of **classes (grammar, history, math, etc.)** have you had in the following languages (primary school through university)?

English

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

Chinese

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

4. How many years have you spent in a **country/region** where the following languages are spoken?

English

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

Chinese

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

5. How many years have you spent in a **family** where the following languages are spoken?

English

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 20+

Chinese

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

6. How many years have you spent in a **work environment** where the following languages are spoken?

English

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

Chinese

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

19 ☐ 20+ ☐

III. Language use
In this section, we would like you to answer some questions about your language use by placing a check in the appropriate box. Total use for all languages in a given question should equal 100%.

7. In an average week, what percentage of the time do you use the following languages **with friends**?

English

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Chinese

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

50

Other languages ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. In an average week, what percentage of the time do you use the following languages **with family**?

English ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Chinese ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. In an average week, what percentage of the time do you use the following languages **at school/work**?

English ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Chinese ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

10. When you talk to yourself, how often do you **talk to yourself** in the following languages?

English ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Chinese ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

11. When you count, how often do you **count** in the following languages?

English ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Chinese ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Other languages ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

IV. Language proficiency
In this section, we would like you to rate your language proficiency by giving marks from 0 to 6.

0=not well at all

6=very well

12. a. How well do you speak **English**?
0 1 2 3 4 5 6
☐ ☐ ☐ ☐ ☐ ☐ ☐

b. How well do you speak **Chinese**?
0 1 2 3 4 5 6
☐ ☐ ☐ ☐ ☐ ☐ ☐

13. a. How well do you understand **English**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. How well do you understand **Chinese**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

14. a. How well do you read **English**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. How well do you read **Chinese**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

15. a. How well do you write **English**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. How well do you write **Chinese**?   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

V. Language attitudes

In this section, we would like you to respond to statements about language attitudes by giving marks from 0-6.

0=disagree

6=agree

16. a. I feel like myself when I speak **English**.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. I feel like myself when I speak **Chinese**.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

17. a. I identify with an **English-speaking** culture.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. I identify with a **Chinese -speaking** culture.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

18. a. It is important to me to use (or eventually use) **English** like a native speaker.   ☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

   b. It is important to me to use (or eventually use) **Chinese** like a native speaker.   ☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

19. a. I want others to think I am a native speaker of **English**.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

   b. I want others to think I am a native speaker of **Chinese**.   0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

Appendix B: Pragmatic competence test

Audio stimuli as .wav files are located in PCT folder of supplemental material.

Multiple choice questions are included in OCT folder as a .docx file.

Appendix C: Stimuli by sentence type

Samples of audio stimuli .wav files are located in stimuli folder of supplemental material.

Target *States with -le* are in u.state/target. Non-coerced controls are in u.state/control.

Target *Semelfactives with -le and durative adverbial* are IC/target. Non-coerced controls are in IC/control.

Target *Complement event coercions* are in CEC/target. Non-coerced controls are in CEC/control.

All fillers are in fillers.

References

Bott, O., & Hamm, F. (2014). Cross-Linguistic Variation in the Processing of Aspect. In B. Hemforth, B. Mertins, & C. Fabricius-Hansen (Eds.), *Psycholinguistic Approaches to Meaning and Understanding across Languages* (Vol. 44, pp. 83–109). Springer International Publishing. https://doi.org/10.1007/978-3-319-05675-3_4

Bouton, L. (1999). *The amenability of implicature to focused classroom insruction*. TESOL, New York.

Brebner, C., McCormack, P., & Rickard Liow, S. (2016). Marking of verb tense in the English of preschool English-Mandarin bilingual children: Evidence from language development profiles within subgroups on the Singapore English Action Picture Test: Tense marking in English-Mandarin bilinguals. *International Journal of Language & Communication Disorders*, *51*(1), 31–43. https://doi.org/10.1111/1460-6984.12181

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of Aspectual Coercion. *Brain and Language*, *106*(2), 132–143. https://doi.org/10.1016/j.bandl.2008.04.003

Brennan, J., & Pylkkänen, L. (2010). Processing psych verbs: Behavioural and MEG measures of two different types of semantic complexity. *Language and Cognitive Processes*, *25*(6), 777–807. https://doi.org/10.1080/01690961003616840

Chan, H. L. (2012). *Tense-Aspect Processing in Second Language Learners*. University of Pittsburgh.

Chen, C.-C. (2009). *The representation and processing of past tense in Chinese English-Language learners*. University of Arizona.

Comrie, B. (1985). *Tense*. Cambridge.

Comrie, B. (1998). *Aspect: An introduction to the study of verbal aspect and related problems* (reprinted). Cambridge Univ. Press.

de Swart, H. (1998). Aspect Shift and Coercion. *Natural Language & Linguistic Theory*, *16*(2), 347–385.

de Swart, H. (2000). *Tense, aspect and coercion in a cross-linguistic perspective* (M. Butt & T. H. King, Eds.; p. 20). CSLI Publications.

Depraetere, I. (1995). On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy*, *18*(1), 1–19. https://doi.org/10.1007/BF00984959

Dölling, J. (2003). Aspectual (re-)interpretation: Structural representation and processing. In H. Härtl & H. Tappe (Eds.), *Mediating between Concepts and Grammar*. DE GRUYTER MOUTON. https://doi.org/10.1515/9783110919585.303

Fernández, E. M. (2003). *Bilingual sentence processing relative clause attachment in English and Spanish*. Benjamins.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing Language Dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Proficiency* (pp. 208–225). Multilingual Matters. https://doi.org/10.21832/9781783092291-014

Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number-of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*, *22*(1), 147–155. https://doi.org/10.3758/s13423-014-0649-7

Hsu, C.-C., & Hsieh, S.-K. (2013). To Coerce or Not to Coerce: A Corpus-based Exploration of Some Complement Coercion Verbs in Chinese. *Proceedings of the 6th International Conference on Generative Approaches to the Lexicon*, 13–20.

Keating, G. D., Jegerski, J., & Vanpatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, *19*(1), 36–49. https://doi.org/10.1017/S1366728914000418

Köpke, B., & Genevska-Hanke, D. (2018). First Language Attrition and Dominance: Same Same or Different? *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01963

Li, S. (2018). Developing a test of L2 Chinese pragmatic comprehension ability. *Language Testing in Asia*, *8*(1), 3. https://doi.org/10.1186/s40468-018-0054-7

Lowder, M. W., & Gordon, P. C. (2016). Eye-tracking and corpus-based analyses of syntax-semantics interactions in complement coercion. *Language, Cognition and Neuroscience*, *31*(7), 921–939. https://doi.org/10.1080/23273798.2016.1183798

Mikhaylova, A. (2011). Interaction of Aspectual Morphology in L2 and Heritage Russian. In G. Granena (Ed.), *Selected Proceedings of the 2010 Second Language Research Forum* (pp. 63–77). Cascadilla Proceedings Project.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition*, *5*(01). https://doi.org/10.1017/S1366728902000135

Montrul, S. A. (2012). Is the heritage language like a second language? *EUROSLA Yearbook*, *12*, 1–29. https://doi.org/10.1075/eurosla.12.03mon

Montrul, S., & Ionin, T. (2012). Dominant Language Transfer in Spanish Heritage Speakers and Second Language Learners in the Interpretation of Definite Articles. *The Modern Language Journal*, *96*(1), 70–94. https://doi.org/10.1111/j.1540-4781.2012.01278.x

Muysken, P. (2020). The case for contact induced-change in Heritage Languages. *Bilingualism: Language and Cognition*, *23*(1), 37–38. https://doi.org/10.1017/S1366728919000373

O'Grady, W. (2009). A Psycholinguistic Tool for the Assessment of Language Loss: The

HALA Project. *Language Documentation*, *3*(1), 13.

Piñango, Maria M., Zurif, E., & Jackendoff, R. (1999). Real-Time Processing Implications of

Enriched Composition at the Syntax-Semantics Interface. *Journal of Psycholinguistic*

Research, *28*(4), 395–414. https://doi.org/10.1023/A:1023241115818

Piñango, Maria Mercedes, Winnick, A., Ullah, R., & Zurif, E. (2006). Time-Course of Semantic

Composition: The Case of Aspectual Coercion. *Journal of Psycholinguistic Research*,

35(3), 233–244. https://doi.org/10.1007/s10936-006-9013-z

Polinsky, M., & Scontras, G. (2020a). A roadmap for heritage language research. *Bilingualism:*

Language and Cognition, *23*(1), 50–55. https://doi.org/10.1017/S1366728919000555

Polinsky, M., & Scontras, G. (2020b). Understanding heritage languages. *Bilingualism:*

Language and Cognition, *23*(1), 4–20. https://doi.org/10.1017/S1366728919000245

Puig-Mayenco, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., & Rothman, J. (2018).

Language Dominance Affects Bilingual Performance and Processing Outcomes in

Adulthood. *Frontiers in Psychology*, *9*, 1199. https://doi.org/10.3389/fpsyg.2018.01199

Pustejovsky, J. (1997). The Generative Lexicon. *Language*, *73*(3), 597.

https://doi.org/10.2307/415891

Putnam, M. T., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition?: A

prolegomenon to modeling heritage language grammars. *Linguistic Approaches to*

Bilingualism, *3*(4), 478–508. https://doi.org/10.1075/lab.3.4.04put

Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*,

23(2), 229–256. https://doi.org/10.1191/0265532206lt329oa

Sampaio, T. O. da M., & França, A. I. (2018). Event-duration semantics in online sentence processing. *Letras de Hoje*, *53*(1), 59. https://doi.org/10.15448/1984-7726.2018.1.28695

Schwarz, F., & Zehr, J. (2018). *PennController for Internet Based Experiments (IBEX)*. https://doi.org/10.17605/OSF.IO/MD832

Slabakova, R. (2015). Acquiring Temporal Meanings Without Tense Morphology: The Case of L2 Mandarin Chinese. *The Modern Language Journal*, *99*(2), 283–307. https://doi.org/10.1111/modl.12216

Smith, C. S., & Erbaugh, M. S. (2005). Temporal interpretation in Mandarin Chinese. *Linguistics*, *43*(4). https://doi.org/10.1515/ling.2005.43.4.713

Stevens, L. (2019). *Heritage Speaker and Late Bilingual L2 Relative Clause Processing and Language Dominance Effects*. CUNY Graduate Center.

Sun, H. (2014). *Temporal construals of bare predicates in Mandarin Chinese*. LOT.

Sun, Y., Díaz, L., & Taulé, M. (2019). The development of dynamicity in the acquisition of Spanish by Chinese learners. *ITL - International Journal of Applied Linguistics*, *170*(1), 79–110. https://doi.org/10.1075/itl.18001.sun

Todorova, M., Straub, K., Badecker, W., & Robert, F. (2000). Aspectual Coercion and the Online Computation of Sentential Aspect. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *22*(22).

Townsend, D. J. (2013). Aspectual Coercion in Eye Movements. *Journal of Psycholinguistic Research*, *42*(3), 281–306. https://doi.org/10.1007/s10936-012-9216-4

Valian, V. (2019). Variability: Definitions of language and language learning. *Bilingualism: Language and Cognition*, 1–2.

Wang, Z. (2008). *Context Coercion in Sentence Processing: Evidence from Chinese* (M. K. M. Chan & H. Kang, Eds.; Vol. 2, pp. 959–974).

Xiao, R., & McEnery, T. (2004). *Aspect in Mandarin Chinese: A corpus-based study*. J. Benjamins Pub.

Xue, W. (2016). *The effect of L2 English on L1 Mandarin Chinese from the perspective of complement event coercion—An empirical study of advanced Mandarin Chinese-English adulthood bilinguals*. Norwegian University of Science and Technology.

Yano, M. (2018). Predictive processing of aspectual information: Evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, *33*(6), 718–733. https://doi.org/10.1080/23273798.2017.1416150

Zyzik, E. (2020). Creativity and Conventionality in Heritage Speaker Bilingualism. *Language Learning*, *70*(S1), 157–187. https://doi.org/10.1111/lang.12349