Crossing barriers: The influence of linguistic and cultural background on [I + verb] belief constructions in expressions of opinion

Lucy Zhao, Jesus College

Doctor of Philosophy in Education

Word count: 63,349

TABLE OF CONTENTS

TABLE OF CONTENTS 2

ABSTRACT 8

DEFINITION OF TERMS 10

1. INTRODUCTION 12

1.1. Research Rationale 12

1.1. Research questions 15

1.2. Outline of the thesis 17

2. LITERATURE REVIEW: PRAGMATICS OF COMMUNICATION AND DISCOURSE 24

2.1. Effect of culture on communication and discourse 25

2.1.2. Effect of shift in cultural environment 27

2.1.3 A cross-cultural comparison of communication style 29

2.1.3.1. American culture and communication style 32

2.1.3.2. Chinese culture and communication style 38

2.2. Pragmatic realization via focus-marking 48

2.2.1. Cross-linguistic realizations of focus-marking 51

2.2.1.1. Focus-marking via pitch in English 54

2.2.1.2. Focus-marking via pitch range in Mandarin 55

2.3. Effect of language knowledge on communication and discourse 59

2.3.1. Influence of L1 Mandarin transfer on EFL discourse 59

2.3.2. Influence of L2 proficiency on EFL discourse 62

3. CONCEPTUAL FRAMEWORK AND RATIONALE: PRAGMATIC FUNCTION OF [I + VERB] BELIEF CONSTRUCTIONS 66

3.1. Syntactic variation within [I + verb] belief constructions 67

3.1.1. Syntactic categorization of [I + verb] forms 69

3.1.2. Alternative semantics revisited: Prosodic cues for [I + verb] forms 71

3.1.3. The DM conundrum: Contextual clues for [I + verb] forms 73

3.2. Developed framework for determining pragmatic function of [I + verb] 74

4. METHODOLOGY 77

4.1. Research design 77

4.1.1. Research questions 79

4.1.2. Hypotheses 82

4.2. PHASE 1: Proof of Concept 85

4.2.1. Participants 85

4.2.2. Materials 86

4.2.3. Procedure 91

4.2.4. Analysis 92

4.2.4.1. Examination of timing 94

4.2.4.2. Coding scheme for open-ended speaker confidence definitions 95

4.2.5. Pilot data 98

4.3. PHASE 2 99

4.3.1. Participants 100

4.3.2. Materials 104

4.3.3. Procedure 108

4.3.4. Analysis 110

4.3.4.1. Length of audio recording 116

4.3.4.2. Orthogonal and prosodic transcription 117

4.3.4.3. Coding scheme for [I + verb] constructs 120

4.3.4.4. Coding scheme for Expression of negative appraisal 126

4.3.5. Pilot data 127

4.4. Limitations of studies 131

4.5. Ethical considerations 136

5. PHASE 1 RESULTS 137

5.1. Examination of timing 137

5.2. Perception task 137

5.3. Coding of speaker confidence definitions 160

6. PHASE 1 DISCUSSION 166

6.1. Perception task 166

6.2 Speaker confidence definitions 172

7. PHASE 2 RESULTS 184

7.1. Length of audio recording 184

7.2. Frequency of functional [I + verb] usage 187

7.2.1. ANALYSIS 1 - Percentage of pragmatic function of [I + verb] to total utterances produced 187

7.2.3. ANALYSIS 3 - Calculated CR of speaker confidence 209

7.2.4. Summary of trends and significance 214

7.3. Comparison of Self-opinion vs. Opinion of others prompt set 216

7.4. Effect of L2 proficiency on frequency of [I + verb]: 222

7.4.1. Descriptives for EFL-specific between-subject factors 223

7.4.2. ANALYSIS 1a - Percentage of pragmatic function of [I + verb] to total utterances produced for N = 50 225

7.5. Effect of Time immersed in new sociolinguistic environment on frequency of [I + verb] 230

7.5.1. Descriptives for EFL US-specific between-subject factors 231

7.5.2. ANALYSIS 1b - Percentage of pragmatic function of [I + verb] to total utterances produced for N = 25 231

7.6. Expressions of negative appraisal in the Opinion of others prompt set 243

8. PHASE 2 DISCUSSION 250

8.1. Frequency of functional [I + verb] usage 250

8.2. Expressions of negative appraisal 262

8.3. Effect of Time immersed in new sociolinguistic environment 266

8.4. Effect of L2 Proficiency 268

9. CONCLUSION 272

9.1. Summary of key findings 272

9.2. Limitations and recommendations for future research 273

9.3. Implications 280

9.4. Concluding remarks 284

REFERENCES 286

APPENDICES 302

Appendix 1: Information sheet (Phase 1) 302

Appendix 2: Consent form (All Phases) 305

Appendix 3: Language questionnaires (All phases) 307

Appendix 4: Perception task example trial page (Phase 1) 312

Appendix 5: Perception task condition distribution (Phase 1) 314

Appendix 6: Phase 1 Pilot data 316

Appendix 7: Sample selection of transcriptions (10%) 328

Appendix 8: Coding for reliability 364

Appendix 9: Information sheet (Phase 2) 385

Appendix 10: Production task (Phase 2) 389

Appendix 11: Semantic annotation instructions and example 393

Appendix 12: Phase 2 Pilot data 398

Appendix 13: ANCOVA results of Confidence Rating of [I + verb] prosodic variations with potential outliers removed 417

Appendix 14: Results from repeated measures ANOVA with M = 100 419

Appendix 15: Non-parametric Kruskal-Wallis test for Phase 1 calculated CR 421

Appendix 16: Effect of L2 proficiency on frequency of [I + verb] 422

ABSTRACT

How does cultural and linguistic background influence communication style? This topic is examined through the [I + verb] belief construct before the expression of an opinion. Since opinions carry inherent notions of speaker belief, these constructions may at first appear superfluous. However, [I + verb] forms may actually fulfill various pragmatic functions depending on prosodic variation. Unfortunately, there is little congruent data on universality vs. cross-linguistic variability of pragmatic-prosodic mappings (prosodic variation as a cue to pragmatic interpretation) of [I + verb] belief constructs before an opinion. Thus, a Proof of Concept perception test was first implemented, followed by a production task investigating the effect of sociolinguistic background on a speaker's frequency of usage for various [I + verb] forms in expressing opinions, and how this relates to perceived speaker confidence.

Usage of various forms and functions of this construct was analyzed and compared between native Mandarin (CHI) and English (US) speakers, as well as EFL Mandarin speakers. The Proof of Concept task supported hypotheses overall, suggesting a possible universal pragmatic-prosodic mapping for [I + verb]. In addition, while as predicted sociolinguistic background did not have a significant effect on universality of pragmatic-prosodic mapping in terms of confidence rating, it did have an observable effect on semantic interpretation of 'speaker confidence', thus indicating that sociolinguistic background may play a role in influencing these interpretations.

Results from the production task supported predictions that frequency of functional [I + verb] usage corresponded to culturally specific attitudes of each culture. Based on

confidence rating calculations for each [I + verb] variation from pragmatic-prosodic mapping of the perception task, it was determined that Native US individuals were most confident in expressing self-opinions but least confident in expressing opinions of others whilst Native CHI individuals were most confident in expressing opinions of others and least confident in expressing self-opinion, with the EFL group in the US more closely mirroring the Native US group and the EFL group in China more closely mirroring the Native CHI group. Additionally, going against theories of previous research, Time immersed in a new L2 environment and L2 proficiency did not significantly influence performance.

Through investigating pragmatic-prosodic mappings of [I + verb] forms vs. functions, this study aimed to demonstrate the bi-directional link between language, thought and culture. By understanding and familiarizing oneself with the root of pragmatic differences, there is hope to better understand the cause of cross-cultural miscommunications between native and foreign speakers in conversation and to minimize any such discrepancies in pragmatic knowledge and sociocultural norms.

DEFINITION OF TERMS

Discourse = Spoken conversational communication in a social context[1].

Culture = "The collective programming of the mind which distinguishes the members of one group or category of people from another" (as best as may be homogenized) (Hofstede, Hofstede, & Minkov, 2010:1).

Cultural background = An individual's knowledge of culture based on their past and present experience in their cultural environment, including social norms, expectations, and behavior.

Sociolinguistic background: An individual's cognitive mindset based on their linguistic and cultural background, both past and present, with a focus on how language relates to communication (Gumperz & Cook-Gumperz, 2008). Assuming that language and culture are inevitably intertwined (as linguistic experience varies according to culture), this thesis hereafter adopts this term to refer to such influences.

"I presupposition" = **[I + verb] belief construct** = First person pronoun-verb combinations in which the verb is classified as a mental, or "cognitive" verb (Quirk et al. 1985:1112).

ACRONYM GUIDE:

US = American English

CHI = Mandarin Chinese

EFL = English as a foreign language

[1] While 'discourse' can refer to both spoken and written entities of various types of communication (e.g. academic; conversational), for the purpose of this thesis, we will focus on natural conversational speech.

IDV vs. (Rel-)COL = Individualist vs. (Relational-)Collectivist cultures

IndSC vs. RelSC = Independent vs. Relational self-construals

DIR = Direct sentence (no [I + verb] construct)

MC = "Main clause" variation of [I + verb] with focus-marking on the pronoun **(PRN)**

EM = "Epistemic marker" variation of [I + verb] with focus-marking on the verb **(V)**

DM = "Discourse marker" variation of [I + verb] with no focus-marking

DMF = DM with sentence-initial syntactic position **(FRONT)**

DMM = DM with sentence-medial syntactic position **(MID)**

DME = DM with sentence-final syntactic position **(END)**

1. INTRODUCTION

This thesis begins by providing an introduction to the current research, its relevance, aims, and significance. Following this, research questions will be set forth and an outline of the thesis presented.

1.1. Research Rationale

What effect does cultural and linguistic background have on the cognitive process behind language use? Given the disparate views on this subject over the past century, the answer may vary greatly depending on whom you ask. Whilst some, under the lens of a Universalist perspective, tend to treat language as a dependent notion of cognition without regard to culture (Alford, 1978), others believe that this view gives no evidence as to why culture should not play a role in determining thought (Deuschter, 2010). Indeed, past pioneers in the field of social science have repeatedly stressed the interrelatedness of culture, thought, and language, specifying that culture may actually have an effect on cognition and thereby language, in which case cultural experience should not be retracted from consideration (Whorf, 1956; Gumperz & Levinson, 1991). Recent scientists have also adopted a similar perspective: Segalowitz (2010) asserts that real time language use reflects not only the current cognitive state of the speaker, but also the developmental history of the speaker. Such sociolinguistic development in the speaker's past may inevitably differ from the development and experiences of speakers from other cultures, thereby indicating the influence of sociolinguistic background on the cognitive process behind language use.

It follows then that pragmatics, the realm of linguistics dealing with communicative language use, is an area where culture may affect language and cognition, for pragmatic knowledge is derived from social and philosophical contexts within a language community, not from anything inherent to the arbitrary lexicon itself (Glucksberg & McGlone 2001:8). That is, pragmatic knowledge may be an area of linguistic knowledge shaped by social environmental input. This input would consequently differ based on the culture in which the speaker is immersed, resulting in potential cross-linguistic variations in pragmatic knowledge and inference during communication.

This thesis aims to investigate the intricate web between language, thought and culture, and how such factors may influence communication and speaker style. Whilst interdisciplinary in nature, this research remains primarily in the field of applied linguistics—particularly that of cross-cultural pragmatics (Phase 1) and interlanguage pragmatics (Phase 2), with a focus on the pragmatics-prosody interface in Phase 1 and discourse analysis in Phase 2. The specific topic of this research is the effect of sociolinguistic background[2] on the communication style of conversational discourse. Specifically, this thesis examines a speaker's use of pragmatic tools in expressing an opinion, and how this may relate to pragmatic inference of perceived confidence and directness of a statement. These pragmatic-prosodic variables will be examined through a particular pragmatically ambiguous construct that many of us may frequently use without conscious effort: the [I + verb] belief construct preceding an opinion (e.g. “I think”, “I feel,” “I believe,” etc.).

[2] Refer to Definition of terms for “Sociolinguistic background.”

Because opinions inherently express speaker belief, the addition of [I + verb] belief constructs may seem redundant (Langacker, 1987). For example, a speaker who asserts, "Learning languages is fun," does not need to preface this utterance with 'I think/believe/suppose,' since the statement itself already serves to illustrate the speaker's belief. Pragmatically, however, these constructs act on modality to convey assertive or reluctant attitudes of speaker stance. Whilst frequency of [I + verb] has been analyzed for verbs like 'think' and 'believe,' these studies have primarily focused on their function as discourse markers (DM). Additionally, viewpoints differ in what constitutes DM vs. non-DM (Epistemic marker, Main clause) use, resulting in inconsistent analyses of [I + verb]'s performative functions. Furthermore, studies have not always addressed the variable of L2 proficiency on usage frequency, resulting in a tangle of misinterpreted results.

This thesis' investigation is conducted through a two-part, cross-cultural and cross-linguistic study comparing communication style between speakers of American (US) and Chinese (CHI) sociolinguistic background. This research aspires to better understand the cause of communicative difficulties and cognitive disparities that inevitably arise when individuals immigrate to a new culture, and does so through contributing a novel empirical Proof of Concept experiment examining universality of pragmatic-prosodic mappings via the effect of focus-marking in [I + verb] on perceived speaker confidence, as well as a novel mixed-methods experiment investigating how frequency and form of pragmatic [I + verb] usage may vary cross-culturally and cross-linguistically. For part two, consideration is also given to how L2 proficiency and Time immersed in a new L2 culture may affect pragmatic knowledge and speaker style in expressing an opinion.

The significance of this research lies both in its Proof of Concept phase and its original corpora of English and Mandarin expression of opinion prompts, from which interesting linguistic insights may be drawn. No research has yet been conducted regarding whether there exists a confidence-mapping trend of various pragmatic-prosodic [I + verb] forms, let alone whether this trend of confidence rating is universal or differs cross-culturally; therefore, the Proof of Concept phase in this thesis seeks to offer insight regarding both these uncertainties. Furthermore, given the lack of congruent research in the field regarding DM vs. non-DM functional use of [I + verb], this thesis' investigation of cross-linguistic, cross-cultural frequency of functional [I + verb] usage aims to make a contribution to this area of discrepancy. Moreover, to address incongruous interpretations concerning L2 proficiency's effect on filler vs. interactional/interpersonal use of [I + verb], the effect of L2 proficiency on frequency of functional [I + verb] usage was examined. Thus, through investigating the pragmatics of speaker intent vs. listener interpretation in expressing an opinion, this thesis hopes to contribute to a better understanding of how sociolinguistic background and language proficiency may affect the pragmatic use of language.

1.1. Research questions

The general aim of this research is presented in the research questions (RQ) below for Phase 1 and Phase 2. Specific hypotheses for these RQs are formed and presented in Chapter 4.

Phase 1:

1) Are specific prosodic variations of [I + verb] belief constructions linked to different pragmatic interpretations in discourse?

1a) For Native vs. Mandarin speakers of English, do these pragmatic-prosodic mappings vary as a function of sociolinguistic background?

2) Does sociolinguistic background influence pragmatic attitudes toward speaker confidence?

Phase 2:

3) Does sociolinguistic background influence the frequency of functional [I + verb] usage when expressing Self-opinion vs. Opinion of others?

3a) For the "Opinion of others" prompt set, does sociolinguistic background influence whether participants express a negative appraisal when asked for their impression of the character in the story?

4) Does this vary as a function of Time immersed in a new L2 environment?

5) For Mandarin EFL speakers, does L1 pragmatic knowledge transfer[3] influence the functional use of [I + verb] belief constructions when expressing an opinion in English?

6) For Mandarin EFL speakers, does frequency of functional [I + verb] usage when expressing an opinion in English vary as a function of L2 proficiency?

[3] "Pragmatic knowledge transfer" refers to the subset of language transfer in which an individual applies knowledge developed from an L1 language (in this case, pragmatic knowledge developed from sociolinguistic background) to the learning of another language (Spencer-Oatey, 2000:184). This pragmatic transfer may be done in an attempt to "grasp and understand the [L2] language" in a linguistic and cultural context (Lado, 1957:2).

1.2. Outline of the thesis

This thesis opens with a literature review on the pragmatics of communication and discourse, in reference to how sociolinguistic background may affect second language acquisition (SLA) via pragmatics. The overall discussion investigates how pragmatic inference may universally affect linguistic communication, specifically through cross-linguistic manifestation of pragmatic function via prosodic focus-marking.

The review begins by exploring communication from an interactional sociolinguistics (IS) perspective, wherein language and language use "are conceived of as social and cognitive activities" (Fetzer, 2011:258) and involve both speaker intent and listener comprehension. Speakers use pragmatic tools, or "inference triggers," to express speaker attitude toward their statement as well as to communicate their intent to listeners; in turn, listeners draw upon their own pragmatic knowledge and contextual clues to try and deduce the speaker's intent.

The subsequent section discusses how sociolinguistic background may affect the style in which a speaker utilizes pragmatic knowledge to convey speaker attitude toward an utterance. A neuro-cultural interaction model (Kitayama & Uskul, 2011) is examined in which culture is a collective process derived from behavioral norms in the social environment. Given this model, it follows that pragmatic knowledge, being a sociocognitive process, is also derived from social environment and may thereby differ cross-culturally.

Next, this thesis explores how a shift in cultural environment may accordingly cause a shift in cognitive perspective via new environmental input (Hofstede, 2011; Zou et al., 2010). Priming and immersion studies regarding cultural shifts between US vs. CHI

cultural mindsets are presented and analyzed. The discussion then investigates Hofstede's (1980/2011) model of cultural dimensions, with particular focus on the individualism vs. collectivism scale and Kitayama's (1991) coinciding concepts of independent vs. interdependent self-construal. This explores the culture and communication style of what Hofstede deems the prototypical individualistic society of the United States, with its direct and assertive self-focused perspective and independent self-construal, and the prototypical collectivist society of China, with its face-saving other-relational perspective and relational self-construal. While this scale may appear to craft certain generalizations surrounding culture and the potential homogeneity of culture, it acts to create not necessarily a steadfast dichotomy between US vs. CHI speakers but more so the conceptual basis from which analyses may be drawn. Consequently, using mixed-methods analyses, the perspectives surrounding these sociocultural theories are compared and contrasted within each culture as data is collected, resulting in an evolving concept of US vs. CHI sociolinguistic background as data exploration progresses.

The following section explores how speakers use prosody as a pragmatic tool to convey meaning by placing stress ("focus-marking") within an utterance. Focus-marking is examined via interactional sociolinguistics (IS) and alternative semantics, in which focus-marking is utilized by a speaker to imply a certain meaning over all other possible meanings. Studies are examined in which prosody is altered via pitch (F0) to signify meaning, both in Mandarin and English, with analyses concluding that F0 manipulation may be used cross-linguistically as prosodic cues for pragmatic inference.

Subsequently, the effect of language knowledge on communication and discourse is explored, both in terms of L2 proficiency and L1 Mandarin transfer, on EFL discourse.

L2 proficiency may correlate with verbal routine, or "filler" forms of phrases that may otherwise be used as focus-marked prosodic cues. Furthermore, because speakers learn to associate pragmatic intent with particular prosodic cues (i.e. mapping) based on cultural input, L1 transfer may be possible regarding pragmatic knowledge. Additionally, culturally variable interpretations of universal cues may affect how pragmatic cues are perceived as discourse style between Mandarin/English speakers. Accordingly, cross-linguistic pragmatic knowledge is investigated in Phase 1 of this study, and any variation found between pragmatic-prosodic mappings is taken into account for Phase 2.

Before discussing methodology, this thesis first presents a conceptual framework and rationale for the pragmatic function of [I + verb] belief constructs. The pragmatic-prosodic influences surveyed above will be investigated cross-linguistically through this specific syntactically ambiguous construct, chosen for its prosodic variation-supporting multiple lexical item composition, multiple cognitive forms within the same surface structure, and ambiguity-induced variation in pragmatic function.

The framework begins by explaining the lexical composition of [I + verb], and how similarities in English and Mandarin allow the construct to appear freely and frequently within each language. Variation in focus-marking may lead to three separate prosodic forms of [I + verb], each of which fulfills a different pragmatic function. A syntactic categorization is presented for each prosodic form with determining variables. These syntactic forms are then linked to assertiveness of language via the modality of involvement vs. politeness (Nikula, 1996), with more assertive forms of [I + verb] being linked to more confident, direct communication. The discussion then examines how this link to modality of assertiveness is further supported by revisiting alternative semantics,

and goes on to correlate alternative semantics with Dehe and Wichmann's (2010) prosodic determination of how focus-marking signifies [I + verb] form. Next, the framework examines contextual cues for pragmatic [I + verb] forms more closely, exploring how a listener may distinguish between two pragmatic forms (interactional/interpersonal discourse marker vs. gap filler) that share the same prosodic pattern (no focus). Noting a lack of research in the current field, this thesis concludes that although differentiation may not be possible, these forms may still be treated similarly as they both result in decreased speaker confidence through hedging.

Following this, a developed framework for determining pragmatic function of [I + verb] is put forth, with both prosodic and syntactic cues utilized. From this framework, a measure of communication style was developed and labeled as speaker confidence. Since [I + verb] forms function pragmatically to convey speaker stance, each prosodic-syntactic [I + verb] variation in this measure (Main clause, Epistemic marker, Discourse marker) is assigned a confidence value of pragmatic attitude (Deliberate, Uncertain, Reluctant) based on the semantics implied by its focus-marking. In summary, [I + verb] seems an opportune construct for observing deviations in pragmatic inference and cross-cultural speaker perception.

The next chapter presents methodology, subsequently outlining a two-part investigative research design to explore variations in the cognitive and pragmatic function behind linguistic communication across cultures and languages. The chapter commences with an overview of the various pragmatic-prosodic forms of the [I + verb] construct and sets a context in which they may be observed: Expression of opinion. Incongruous classifications and methodological shortcomings in past literature are

addressed, particularly regarding lack of consideration for L2 proficiency in usage frequency and differing viewpoints in what constitutes Discourse marker (DM) vs. Main clause (MC) and Epistemic marker (EM) use of [I + verb]. Thus, this sets the stage for a Proof of Concept task (Phase 1) to be implemented. Variables affecting frequency and form of [I + verb] usage are discussed, including sociolinguistic background, L1 pragmatic transfer, and L2 proficiency. Subsequently, research questions and hypotheses are set forth, and each task in the study is relayed to specific RQ(s).

Next, Phase 1 is presented, wherein an interpretive, mixed-methods task was designed to determine Proof of Concept for pragmatic interpretation of prosodic variations of [I + verb]. Participants and recruitment details are relayed, followed by a description of materials (language questionnaire, perception task) and procedure for data collection. Methods for analyzing data are described in terms of quantitative analysis for the perception task and examination of timing, and coding scheme for the open-ended definition of speaker confidence at the end of the task. Pilot data is then presented, with subsequent changes from pilot phase to main phase relayed.

Following Phase 1, the methodology for Phase 2 is presented, in which a mixed-method, open-ended production task was created to investigate the influence of sociolinguistic background on communication style. Similarly, participants and recruitment details are put forth, followed by a description of materials (language questionnaire, production task, Lex30 vocabulary test) and procedure for data collection. Analysis of data is then described regarding length of recording, transcription, coding scheme for [I + verb] constructs, and quantitative analysis for the production task

(frequency of functional [I + verb] constructs). Finally, Phase 2 pilot data is presented, noting the subsequent changes made to methodology for main phase testing.

To conclude the methodology, potential limitations of the study are addressed, and ethical considerations are examined.

The next chapter presents results from Phase 1: Proof of Concept. Examination of timing is inspected for proper completion of task and to disqualify improperly completed surveys. Then, perception task results are presented, followed by results for speaker confidence definition. Given that Phase 1 is a Proof of Concept phase, these results are explored in the next chapter's Phase 1 discussion, prior to presentation of Phase 2 results.

Following Phase 1 discussion, Phase 2 results are explored for both Self-opinion prompts and opinion of other prompts. Length of audio recording is examined to ensure no mean difference of utterances per prompt between groups and thus avoid this potentially confounding variable. Next, results are presented for frequency of functional [I + verb] usage, taking into account the variables of L2 proficiency and Time immersed in new sociolinguistic environment within in-depth sub-analyses, followed by an examination of expression of negative appraisal in Opinion of others prompts in relation to sociolinguistic background. Discussion of Phase 2 results follows in the penultimate chapter.

The concluding chapter presents a key summary of findings, followed by potential limitations, recommendations for future research, and implications of findings. The chapter closes with concluding remarks regarding the study as a whole. It is hoped that through understanding the root of pragmatic differences, there is potential to better understand why cross-cultural miscommunications occur between native and foreign

speakers in conversation, and to better explore the relationship between speaker perception and speaker intent.

2. LITERATURE REVIEW: PRAGMATICS OF COMMUNICATION AND DISCOURSE

Based on past literature, a conceptual framework will be established regarding how pragmatic knowledge directs language use. Speaker and listener interactions, or "speech activity," are viewed through interactional sociolinguistics, which provides a framework for discourse analysis where “language and language use are intrinsically connected, and both are conceived of as social and cognitive activities" (Fetzer, 2011:258). Thus, language and language use are treated as interchangeable, and will hereby be identified as "language use" to avoid confusion with "language knowledge" (defined in previous section). Interactional sociolinguistics coincides with pragmatics in both its view of language use as a sociocognitive act, and speakers as "intentional agents" who use language to communicate speaker intent (Fetzer, 2011:259; Grice, 1989).

As pragmatic communication is considered a cognitive process, various theories of how this process functions have been posited in the past (See Grice, 1975/1989; Sperber & Wilson, 1995). Such theories on pure pragmatic inference explore processes from the stance of both speaker intent and listener interpretation: a speaker intends for a listener to understand X, and the listener interprets X based on what they infer the speaker to intend. This inference is based on perceived common ground, or socio-cultural and linguistic knowledge shared between speaker and listener (Grice, 1989).

Grice's (1975/1989) inferential model of communication focuses not just on what is said but also on how information is interpreted by the listener (Sperber & Wilson, 1995). That is, communication consists less of what is said but more of what is *meant*. A communicative act constitutes a speaker saying something with the intention for the addressee to recognize what

is said; the speaker's communicative goal is not satisfied unless the addressee recognizes this intent (Grice, 1989)[4].

A speaker draws upon his/her pragmatic knowledge to create pragmatic markers, or tools, called "inference triggers" (Fetzer, 2011:258), which indicate that the speaker intends to create a conversational implicature; these conversational implicatures are used to guide the listener toward the correct inference for an utterance (Grice, 1989). Such pragmatic tools are used by speakers not only to direct listener interpretation, but also to express speaker attitude toward one's statement. How these pragmatic tools are used and vary cross-culturally is explored in the next section.

2.1. Effect of culture on communication and discourse

This section will discuss how culture may affect the style in which a speaker utilizes pragmatic knowledge to convey speaker attitude and stance toward an utterance. Hofstede (1980) stresses that culture is not inherited, but rather is derived from social environment. The sociocognitive nature of pragmatics (see Section 2) suggests that discourse is influenced by a speaker's mental state, which is constructed largely by environmental input from experience and socio-cultural behavioural norms (Lakoff & Johnson, 1980). These norms are transmitted cross-generationally on a community's macro-level, in an incremental process that builds upon itself to strengthen existing patterns and systematically shape perception

[4] Although *how* listeners come to infer an utterance's meaning is interesting, the aim of this thesis is communication style; thus, the focus will remain mainly on speaker intent and consequential production of an utterance[4]. That is, while theoretical constructs are tangentially relevant, the crucial elements of communication style reside in the applied use of a speaker's pragmatic knowledge to produce an implied meaning, or conversational implicature (Grice, 1970/1989).

(Kitayama & Uskul, 2011).[5] In other words, culturally specific patterns in thought are passed down historically and over time through perpetuated ideas and values, which are produced through behavioral norms prevalent in that culture; such behavioral norms act to strengthen these ideas and values, creating a perceived cultural consensus on customary ways to think and act.

This collective view of culture is aptly expressed through Kitayama and Uskul's (2011) neuro-cultural interaction model. This model illustrates how individuals within a society choose certain "cultural tasks" (p. 421) based on macro-level factors such as cultural interdependence and independence[6] and repeatedly engage in these tasks; this repetition of input shapes the brain's neuroplasticity[7] through forming culturally-shaped neural patterns that subconsciously become automatic behaviours and subsequently contribute to cultural identity formation. Consequently, it appears that pragmatic knowledge is not exacted by anything inherent to one's own cognitive representations, but is instead shaped by shared social representations, and cultural variation may produce sociocognitive variations in communicative language use (see section 2; e.g. Zhu, Zhang, Fan, & Han, 2007). A speaker's development of pragmatic knowledge may therefore differ across sociolinguistic environments depending on which communicative behaviors are salient in each culture. The potential effects from such a shift in environment are discussed in the following section.

[5] While how distinctive cultures come to *exist* is of interest, this thesis' focus is on how individuals come to *adapt to a cultural mindset*. See Kitayama & Uskul (2011) and Markus & Kitayama (1991) for review on the former.
[6] More on interdependence vs. independence of a society later in this section.
[7] According to Kitayama & Uskul (2011:423), neuroscientists have widely agreed on the theory that neurons which fire together become wired together in patterns based on sustained patterns of engagement.

2.1.2. Effect of shift in cultural environment

A shift in cultural environment may accordingly cause a shift in cognitive perspective via new environmental input (Hofstede, 2011; Zou et al., 2010). This can be seen in studies implementing cultural priming in order to alter participant perspective and response (e.g. Oyserman & Lee, 2008; Hong, Chiu, & Kung, 1997). For instance, Wong and Hong (2005) found that priming individuals with Chinese cultural cues caused individuals to cooperate more with friends in a prisoner's dilemma experiment than when primed with American cultural cues, coinciding with Chinese culture's focus on interdependence and social helping.

Additionally, Heine and Lehman's (2004) study investigating East Asian vs. American self-esteem ratings showed that while American self-esteem ratings were higher than Asian participants', Asian ratings increased with more time lapsed, or "acculturated," residing in Western culture. That is, Asians immersed in American culture began to coincide with Americans in neuro-cultural activation patterns overtime: A shift in cultural environment led to a shift in self-perception. In other words, notions of identity, including how identity is defined and by whom it is defined (i.e. by the individual vs. by others), may shift as a function of how long one has been immersed in a new socio-cultural environment.

Another conduit in which a shift in cultural mindset may occur is through the language in which one is speaking. For instance, Ross, Xun and Wilson (2002) found that Bi-cultural Chinese individuals described themselves in terms of internal attributes when using English as opposed to in terms of external attributes when using Chinese, a shift that mirrors cultural variation between the self-defined identity of US culture and the relationally-defined identity of Chinese culture (to be further discussed in the next section). Briley, Morris and Simonson (2005) further support this notion: Results showed that proverbs written in English or

Cantonese primed Chinese participants to respond differently in terms of endorsement, with the same proverbs that advocated extremity being endorsed more strongly in English than in Cantonese versions. Given that proverbs were the same across both languages and participants remained constant, it appears that cultural variability was induced through language priming, thereby causing a shift in cultural mindset when responding to the task. Thus, these results suggest that language may act as a cultural prime, with a shift in spoken language potentially activating a shift in cultural mindset.

Such results suggest that increased cultural influence, whether through priming or immersion, augments that culture's influential input to alter participant's socio-cognitive patterns. Subsequently, given that cultural perception is continually constructed and reconstructed through linguistic input and communication (Zou et al., 2010:586), the influence of a new cultural environment may strengthen with time and increased societal discourse with others within that culture.

Furthermore, as previously mentioned, given that pragmatic knowledge is derived from behavioral norms within socio-cultural input, it then follows that a shift in cultural environment may cause a shift in pragmatic use of language and inference, with similar potential of varying as a function of Time immersed in a new L2 sociolinguistic environment. This notion of pragmatic knowledge shifting due to a shift in cultural environment, or in other words pragmatic transfer, is further discussed in section 2.4.1.

2.1.3 A cross-cultural comparison of communication style

According to Kashima (2008) and Maass, Arcuri and Suitner (2014), cultural norms and differences are created, transmitted and perpetuated through communication. Thus, a cross-cultural comparison of communication style seems a relevant and appropriate method for measuring and testing culture. Such cross-linguistic research in the past has often taken an etic approach, in order to examine how cultural variability may explain differences or similarities in communicative behavior across cultures (Gudykunst, 1997).

The choice to investigate US vs. CHI communication style derives from what Miyamoto and Eggen (2013:20) deem as two of the most widely examined themes of cultural patterns: Hofstede's (1980/2011) model of cultural dimensions is accordingly compared with Markus and Kitayama's (1991) theory of self-construal, along with more recent empirical results[8] in research to support these concepts (Oyserman & Lee, 2008). Hofstede's (1980) model assigns values to countries on interconnected spatial scales, including the bi-modal spectrum from individualism to collectivism, which refers to one's degree of interdependence within a society and which Triandis (2001) deems as "the most significant cultural difference among cultures" (p. 907). Whilst not mutually exclusive, the amount of interdependence/independence in a society differs based on perceived priorities shaped by aforementioned collective culture and psychological patterns, and have been linked to different forms of self-construal in a manner that Cross, Hardin and Gercek-Swing (2011) deem as nearly indistinguishable, with the former individualism-collectivism dimension describing overall cultures and the latter self-construal dimension for individuals within that

[8] See Oyserman, Coon, & Kemmelmeister (2002) for review.

culture (p. 143). Research on self-construal refer to theories of how individuals make meaning of, and perceive themselves, in relation to others. Specifically, research appears to distinguish between independently defined vs. interdependently defined self-construals, as discussed below.

Although the terms individualism and collectivism were originally developed by Hofstede, their meanings have grown and evolved throughout cross-cultural studies of the past two decades: Individualism in research denotes an emotional independence of the individual to the group or society, leading to an emphasis on individual survival and self-expression, and to the salience of what Markus and Kitayama's (1991) theory of self-construal deems an independently-constructed identity; in contrast, collectivism denotes emotional dependence of the individual on the group or society, with such a cultural context resulting in an emphasis on relational attachment, social harmony, and to the salience of an interdependently-constructed identity (Hofstede, 1980, Markus & Kitayama, 1991; Kitayama & Uskul, 2011; Singelis 1994). Moreover, the notion of "collectivism" has since been refined, segmenting into two separate notions of what Triandis and colleagues (see Hui and Triandis, 1986; Triandis, 1994) deem "relational collectivism," wherein only in-group members are considered, versus collectivism on the whole, that is without exclusion of out-group members (see Cross, Hardin, & Gercek-Swing, 2011 for an overview). Brewer and Chen's (2007) study further supports this distinction: Through comparing US participants' results on a relational vs. collectivist version of a self-construal scale, results indicated little correlation between relational and collectivist notions of the self; thus, these findings provide evidence for a differentiation within the umbrella concept of collectivism.

Both Hofstede (1980) and Triandis (1994) identify the United States as the prototypically individualistic society, with China as the juxtaposing prototypically collectivistic society (The distinction of China as a relational collectivism will be further discussed in section 2.1.3.2 below). This coincides with Western societies apparently placing prominence on independence, while Eastern societies traditionally embody interdependence[9]. These labels and their cognitive consequences at the individual level will be examined below. Whilst survey-based studies like Hofstede (1980) have provided qualitative evidence, newer experimental studies provide empirical support for culture's influence on cognition, many focusing specifically on cognitive differences from cross-cultural variation of interdependence vs. independence, particularly in terms of self-construal. Indeed, through their meta-analysis of studies examining various effects of collectivism and individualism, Oyserman, Coon, and Kemmelmeier (2002) concluded, with the caveat that different aspects of the scale are salient in different aspects of life, that not only can individualism and collectivism be successfully assessed, but that countries *do* appear to differ systematically on these measures, thereby setting the stage for cross-cultural investigations in this area. Furthermore, Hofstede's scale of independence/interdependence may manifest itself in different ways in specific aspects such as personality, motivation, or communication, and thus particular consequences of individual vs. relational vs. collectivist self-construal should be viewed in context-specific situations (i.e. genre analysis; see Paltridge & Phakiti, 2009:209 for reference), such as the specific genre of self-expressive communication examined in this thesis, as opposed to broadly generalized across all aspects of life (Brewer & Chen, 2007; Gudykunst, 1997). Additionally, other factors outside of the boundaries of

[9] While still an assumption, this divide has been widely adopted in the field of cultural psychology (Kitayama & Uskul, 2011:423) and proves successful in integrating diverse evidence from empirical studies into a working conceptual theory. See Markus & Kitayama (1991) and Trinadis (1989) for review.

this scale may inevitably play a role in shaping the cultures of America vs. China, such as the influence of Confucianism in Chinese culture that coincides with relational-collectivist notions (Yum, 1988).

In the following sections, relevant quantitative studies will be presented alongside the aforementioned qualitative theories regarding the concepts of individualism-collectivism, independent-interdependent self-construal, and their relation to US and Chinese communication style. Note that because this cross-cultural comparison is being made between only one country from each side of the scale, this thesis makes no claims about generalizing individualist vs. collectivist cultures overall, but merely aims to compare and contrast how cross-cultural variation in self-construal may influence cognitive processing of communication style within the specific cultures of the US and China.

2.1.3.1. American culture and communication style

It may first be noted that the United States (US) ranks highest of all countries on the individualism (IDV) scale, with a score of 91 out of 100 indicating its position as a self-oriented culture focusing on the "I" rather than "we", and wherein one's identity is defined by the individual, or independently constructed (Hofstede, 1980, Markus & Kitayama, 1991). This autonomous self may derive from such a culture's economic affluence and social and geographic mobility, which allows an individual to easily leave their in-groups and join other groups as they please, thereby resulting in looser relational ties and more focus on the individual (Triandis, 1994). Indeed, Yum (1988) states that Americans appear to be committed to individualism in a way that drives them to distinguish themselves from the past and define one's own self as unique (p. 376). That is, it seems that a high level of

individualism corresponds with the singular individual as the primary unit of society, therefore placing greatest value on the unique and individual self. This in turn leads to an agency to express one's internal needs and to withstand social pressure in order to maintain the valuable uniqueness of the individual (Markus & Kitayama, 1991).

In line with this thinking, Sedikides' (2013) study found that Americans ranked themselves higher in self-esteem ratings than Chinese individuals; given Sedikedes' (2013) notion that high self-esteem equates to value placed on self (p. 237), US individuals thus may value the individual self more than their collective counterparts. That is, in relation to the consensual opinion, one appears to consider the personal self, including one's own opinions and desires, as possessing a high level of significance on the conversational floor. Katriel and Philipsen's (1981) case studies on self-assessment of American communication style corroborate this notion: One participant expressed that her view of the function of parent-child communication is to help the child develop a "positive self-image" (p. 304). A positive self-image should in turn lead to an elevated worth of the self and thus increased confidence in the self. Accordingly, a low score (40) on the Power Distance Index (PDI) scale (level of general equality) indicates a conversational ground wherein each speaker's opinion is valued (Hofstede, 1980). In turn, this high self-worth and value placed on the individual's opinion may foster high speaker confidence and thereby result in more comfort and conviction in self-expression.

The combined high IDV/low PDI score of the US suggests that within the society, discourse is freely offered with direct, informal style. Ingelhart and Welzel (2010) and Hofstede (1981/2011) note that cultures ranking high on self-expression values tend to rank high on interpersonal trust, in turn placing significant value on freedom of self-expression

and creating an atmosphere in which no opinion is automatically deemed as improper as long as a speaker presents it as one's own. Heine et al. (1999:780) support this notion, asserting that Americans dispel negative information about the self[10] in order to maintain positive self-regard, for negative self-esteem contradicts the individualistic notion of independence and self-valuation. This notion is further supported by findings of Katriel and Philipsen's (1981) study mentioned above, wherein another participant expressed that because she is cheerful and pleasant with everybody, the assumption is therefore that everyone will be cheerful and pleasant to her in return (p. 304). This communicative ritual of symmetric and reciprocal pleasantness coincides with the previously stated notion of a positive and comfortable speaking environment. Indeed, following their conclusion that the polar opposite of positive evaluation is not negative evaluation but instead is the *absence* of positive evaluation (p. 313), it appears that in the US sharing one's own opinion may not result in anything but neutral or positive regard. This is not to say that the positive response is necessarily directed at what has been expressed, but appears to be more an unconditional approval of the "unique and precious individual" giving the statement (p. 308). Thus, while this positive regard may be more of a communicative ritual than a sincere response to an utterance, the result is nonetheless the absence of fear of negative evaluation. Given such a forgiving conversational floor in the US, no wonder, then, that individuals may have such a high positive regard for oneself and comfortableness to share one's opinions.

As a result, US individuals may incorporate a high degree of modality (i.e. commitment) to express self-assuredness, causing a speaker to appear more assertive or direct (e.g, "We *should* go to the store [and so we will]" vs. "We *could* go to the store [but maybe not]")

[10] See Heine et al. (1999:780) for a review of studies regarding Americans' maintenance of positive self-regard, including Steele's (1988) Self-Affirmation Theory.

(Markus & Kitayama, 1991). Such assertive modal forms are pragmatic markers relating to the speaker's attitude toward the social context (i.e. speaker stance toward the discourse) (Fox, 2001). Indeed, the independent self-construal of individuals in an individualistic society appears to be associated with concern for clarity and thereby directness in conversations (Kim, Sharkey, & Singelis, 1994). The use of such explicit, direct communication style by US speakers is further supported by Chai and Fontaine's (2007) study on Chinese and American students at an American university, with results indicating that US individuals preferred low-context communication (i.e. favoring a direct verbal communication style with the need to satisfy individual concerns) whilst Chinese individuals preferred high-context communication (i.e. favoring an indirect verbal communication style with a focus on satisfying other's concerns and face-saving). US speakers' tendency toward high modality constructs establish authority and responsibility toward an utterance (Kärkkäinen, 2003). In turn, this high modality explicitly indicates personal stance toward discourse.

Such high modality may stem from the individualistic vs. collectivistic influence on perceived dispositionism, or how much one's behaviour is shaped externally vs. internally. In a study where participants listened to an individual read an essay based on pre-specified, given arguments, Choi, Nisbett and Norenzayan (1999) found that East Asians were less likely than Americans to attribute the essay's views to an individual's internal disposition vs. external motivations. From this finding, it seems that in the US, one's opinion is thereby most likely governed by one's own perception, whereas East Asian's expressions of opinion may be more likely to be formed by external influences. Further support is provided by Choi and Markus (1998), whose study found that when asked to describe a murder case, American

participants explained the murder's motivation in relation to the murderer's internal attributes, whereas Asian participants explained the murderer's actions through how social contexts leading up to the murder shaped the murderer's disposition; thus, it appears that US individuals view disposition as inherently self-determined and fixed, while Easterners view disposition as externally motivated and socially context-conditional. Norenzayan et al.'s (1998) study provides more support: When participants were questioned concerning the ability to change one's personality, East Asians view disposition (personality) as flexible and malleable, with expectations to change as circumstances unfold. The US perception of disposition as steadfastly linked to one's nature thus naturally leads to increased speaker confidence and commitment: One's opinion is consistently accurate to one's disposition, so one may speak with certainty and assuredness. An individual within this environment may therefore speak without worry of having chosen an unsuitable viewpoint, for their viewpoint is derived from knowledge of the self, which, given such a high value placed on knowing oneself, is much less variable than the ebb and pull of external opinions and motivations. This naturally leads to what Yum (1988) deems a "speaker-centeredness" perspective in communication (p. 384). Given that one's opinion is attributed to the self and not others, it follows then that the individual's goal in communication is for one's own self interest as opposed to being determined by external motivations. Additionally, this should lead to the same sort of communication style across all conversational floors (Markus & Kityama, 1991), as the internally-defined self should be the same regardless of whom they are speaking with. This view is further supported by Cross, Hardin, and Gercek-Swing (2011), whose review of cross-cultural self-construal research concluded that whilst the ability to change one's behavior depending on situational context is a sign of "maturity" for the

interdependent self, it is demonstrating a consistent disposition across all situations, particularly whilst communicating assertively, that is a sign of maturity for the independent self (p. 143).

However, it appears that in certain conversational contexts even speakers from an individualistic culture like the US may tend to use indirect speech, such as when requesting a favor without the guarantee of help. In such a scenario, Yu's (1999) investigation of Chinese speakers vs. American speakers in their L1 found that Americans used direct request strategies in only 1% of instances, as opposed to 33% for Chinese speakers. While both cultures appear to prefer indirect request strategies overall, perhaps the reason Americans made so many less direct requests is rooted in the independent nature of its culture: Triandis (1994:172) claims that the most undesirable outcome for an individualist is to have to depend on other group members. Subsequently, given that the individual is the autonomous unit of society, it follows that one would tend to place less dependence on others and more on priding the distinctive self and self-reliance, in contrast to an interdependent society in which individual humility may lead to ease of cooperation for social harmony.

In summary, a cultural environment such as the US appears to value the unique self above all, thereby allowing for behavior as a function of individual attitudes and internal disposition. Such a culture subsequently strives through communication to nurture a positive image of the self as a distinctive entity, with self-enhancement and one's own needs being the goal of communicative exchanges. This positive self-valuation cultivates confidence in speakers' expression of opinions, allowing individuals to voice their ideas with conviction and without fear of consequence or judgment. Consequently, such a positive self-valuation, combined with an open and forgiving conversational floor, seems to naturally allow for the

development of unabashed self-expression and direct, deliberate communication style in discourse that is expressed by US speakers. However, a speaker's communication style may nonetheless differ across communicative situations based on variables external to the speaker's relationship to the listener, thus highlighting the significance of specifying a conversational context when examining cross-cultural communication style. That is, more insight may be offered through genre analysis of a particular social situation.

The next section explores Chinese culture and communication style, compared and contrasted to the American culture and communication style discussed above.

2.1.3.2. Chinese culture and communication style

Whilst individualist cultures emphasize the unit of the self, collectivist cultures such as China (20 on IDV scale) may place more significance on the group or society (Hofstede, 1980). Such a value placed on others may stem from the way in which one cognitively processes other- vs. self-perception. Indeed, Zhu et al.'s (2007) fMRI study on self-reference provides evidence for cross-cultural variation in self-perception: Whilst US participants processed judgments about their mothers similarly to control judgments about unrelated others (public figures), Chinese participants processed mother-judgments in the same neural region as self-judgment. In other words, Chinese participants appeared to think of the self and a close in-group member as the same, whereas US individuals made a distinction between themselves and even their closest others. Thus, it seem that Chinese participants process in-group members similarly to the self, although the same cannot be said for out-group members. In contrast, individuals in a culture such as the US do not appear to think of

in-group members as comparably to oneself, for their distinction lies between the self vs. the other entirely, regardless of group membership.

Due to such a focus on relationships with close in-group members but not out-group members, Chinese culture may thereby be considered a relational-collectivist culture: Markus & Kitayama (1991) further distinguish that the focus of a relational-collectivist culture is on the relation to a selective in-group, thereby fostering empathy and reciprocity within the in-group at but at the expense of the out-group (p. 227). Thus, given the significance placed on other-relational perspectives, Chinese individuals may form their own identity through what they perceive in-group others to think and believe, as opposed to a US individualist perspective (refer to section 2.1.3.1) that leads to internally-defined identity through its lack of significant influence from any particular in-group (Gudykunst, 1997). That is, it appears that whilst US speakers' communicative behavior appears to be universalistic toward all other individuals, Chinese speakers are strongly influenced by a few general in-groups across situations, and may behave differently toward in-group vs. out-group members. This notion of selective empathy and therefore selective relation of Chinese culture may stem from the pervasive influence of Confucianism in China's history, which imparts humanism (*jen*, or natural human feelings for others) and reciprocity (*shu*) as part of its core teachings; together, these principles imply reciprocal feeling, or empathy, between the self and another (Yum, 1988:377). However, within Confucian thought this notion of empathy was not necessarily universalistic but more particularistic depending on relationships and situations. Thus, from these teachings Chinese individuals may have developed different social and communicative conventions when associating with the in-group vs. out-group.

Given this, China's dependence on close group bonds may cause individuals to define the self relationally, or in other words through relational self-construal (RelSC), wherein others become a source of definition for the self[11] (Sedikides, 2013; Markus & Kitayama, 1991; Cross, Hardin & Gercek-Swing, 2011). That is, Chinese individuals may allow the self to be defined through other's perceptions of how they appear and are subsequently appraised. Young's (1994) study on dialogues between Chinese EFL and American speakers found that Chinese speakers constantly referred to *other's* opinions and views whilst being unwilling to commit themselves to an opinion, stating that the self-defined, autonomous individual is "almost insignificant" within a Chinese mindset (p. 41). Thus, such evidence reinforces the idea that the external perspective of others, as opposed to the internal self, appears to be the desired view to adopt and express.

Providing additional support, Zou et al. (2010:580) posit that such an interdependent society's relational focus on others increases salience (compliance) of adherence to cultural norms (p. 584), meaning that one may strive to encapsulate the ideas of the general group over personal opinions. This notion is further corroborated by Kim and Markus' (1999) study, which found that when asked to choose a preferred figure or a pen from a group, Americans chose unique figures and different coloured pens whilst Asians preferred ones that looked less divergent. Additionally, coders' dissemination of 193 advertisements found themes bolstering uniqueness as the "right way" to be in American advertisements and conformity in Asian advertisements (Kim & Markus, 1999:793). Thus, whether simply choosing an object

[11] Recall that base on consensus results from questionnaires regarding the meaning of the individualism-collectivism dimension, Hui & Triandis (1986) proposed two different modes of collectivist thought: one being the concern for a certain subset or group of people, and the other being a concern for the general collective whole. These correspond to notions of the relational self and collectivist self, of which China appears to be the former. Thus, from hereon out "other-relational" will denote "direct relation to another in-group member," with exclusion to out-group members.

or adopting a mentality, results demonstrate an American drive for unique self-expression in contrast to an East Asian desire for similarity. This corresponds with Cross, Hardin and Gercek-Swing's (2011) view that self-esteem for RelSC cultures is based on fitting in, whereas self-esteem for IndSC cultures is based on uniqueness. Indeed, this furthermore coincides with Triandis' (1994:172) claim that the "worst thing" for a collectivist is to be excluded from a group, whereas the worst for an individualist is to have to *conform* to the group.

Chinese culture's desire for conformity coincides with the collectivist goal that communication is not utilized to affirm self-identity or achieve individual goals, but instead to preserve communal relations (Gao, 1998:168). That is, the goal of communication is to create bonds instead of boundaries (Young, 1994:41), striving to encapsulate the group perspective whilst restraining one's own inner needs or desires (Markus & Kitayama, 1991:240). Indeed, China's large PDI value (80), double that of US', indicates high discipline, or wariness, in self-conduct (Hofstede, Hofstede, & Minkov, 2010). That is, one may perceive less tolerance for presenting an opinion differing from social norms, in turn adopting a careful communication style to reach "epistemically sound judgments," or opinions congruent with the norm (Zou et al., 2010:580). This wariness coincides with an urge to conform; if not careful to present the "correct" thoughts of mass culture (Triandis, 1994:171), one might accidentally promote a self-view that diverges from the norm and thus acts detrimentally. Such a concept visibly coincides with Triandis' (1994) notion that behavior is motivated by different agencies in individualist vs. collectivist cultures, with the former focusing on the individual's internal attitudes with a concern for how others may "think," and the latter focusing not on how the individual may feel but instead on what the

individual chooses to do (p. 172). Accordingly, this focus on external action instead of innate attitude follows the logic of a relational-collectivist view, with the individual's own feelings being put aside in order to promote in-group harmony.

As a result of such a RelSC's desire for conformity and harmony, the content of one's self-expression may be more focused on listener judgment. Indeed, Yum (1988) states that in contrast to US culture's self-centered speaker perspective of communication, Chinese culture emphasizes listening and interpretation of the listener above the intent of the speaker, with the ability of the listener to then cast negative evaluation on the speaker's statement (p. 384). Furthermore, a relational-perspective means a speaker must take into consideration whether what they say will offend or upset the listener. Such a high dependence on listener judgment may in turn lead to more tentative, indirect discourse, as one cannot know the thoughts or reactions of another as confidently as one's own thoughts. In Lu's (2002) study evaluating first impressions between American/Chinese participants conversing in English, Americans commented that Chinese speakers tended to be vague or indirect, "talking in a roundabout way" which juxtaposes the "directness of Americans" (p. 217). This linguistic imprecision distances Chinese speakers from their ideas (Fox Tree & Schrock, 2002), allowing them to "manoeuvre" the conversation in the future (Lu, 2002:215). Such cautious conversational style coincides with Chinese speakers' trepidation of presenting a personal view that disrupts mass perception: By avoiding speaker commitment, speakers may retract or modify previous statements by claiming it was not their own opinion, but merely an objectively-presented idea. This non-committal mentality is explicated by Wu (1996:146), who says that Chinese culture teaches a speaker to never reveal one's thoughts or feelings, for fear that they will be used against oneself later. In such a situation, caution therefore clouds the speaker's decision

of what or what not to say. This notion of indirectness and face-saving is further expressed through Ma's (1992) study on intermediaries in interpersonal conflicts: In contrast with the assertive approach of direct and open communication to resolve conflicts in the US, in Chinese culture it is more common to use close in-group members to mediate conflicts without addressing the other party directly. The use of a mediator thus facilitates mutual- and other- face-saving, whilst all the while avoiding any direct communication and potential rejection from the listener. Thus, through aversion to self-expression, one may guard oneself from potential retribution.

Such tentativeness to express one's own opinion may also be rooted in the cultural differences of how positively or negatively a statement is evaluated by the listener. In a self-report study evaluating Chinese immigrants' perspectives on complimenting style in their native Chinese culture vs. newly residing American culture, Fong (1998) found that participants perceived Chinese individuals to give compliments less frequently, with such selective complimenting being a manifestation of Chinese culture's indirect communication style. Furthermore, Chinese compliments were perceived to be less strong in nature (i.e. lower modality), and less often encouraging, than US compliments, the latter of which Chinese immigrants perceived Americans to expect regardless of what was expressed. Additionally, one participant expressed that their classmates only offered compliments to close friends whereas US individuals complimented everyone, further supporting the notion of the Chinese in-group vs. out-group relational distinction mentioned above. Finally, the sentiment that Chinese individuals actually *preferred* receiving the more explicit and encouraging compliments offered in a US sociolinguistic environment further lends itself to the notion that despite cultural differences, Chinese students may feel more comfortable and

would benefit more from a more positive and rewarding conversational ground. From these findings, it thereby appears that in a culture such as China, there is far less likely of a chance that offering one's own opinion will be rewarded positively with encouragement or praise.

This cross-cultural distinction in response type may be rooted in differences of what motivates the facilitation of self-esteem or face. Triandis (1994) posits that the different bases of what comprises self-esteem inevitably leads to differences in whether certain motives are perceived as positive or negative: Chinese individuals may desire a "need for debasement" in order to correct for weaknesses and further the goal of self-improvement (i.e. in order to better assimilate to the views of others), whereas US individuals on the other hand may react poorly to such debasement, given its threat to their goal of self-enhancement and uniqueness (Cross, Hardin & Gercek-Swing, 2011). In other words, a culture like China, wherein identity is construed by others, may promote greater levels of criticism and social debasement in order to facilitate self-correction and improvement, which may not be readily accepted in a culture like the US where positive self-image and unique self-enhancement are comprehensively boasted. In sum, becoming a better person in the US seems to be driven by motivations of self-enhancement via encouragement and praise, whilst becoming a better person in China seems to be driven by motivations of self-improvement via criticism and appraisal.

This corresponds with Markus and Kitayama's (1991) notion of ego-focused vs. other-focused emotions: Whilst those with an independent self-construal are more likely to express and experience ego-focused emotions like pride, those with an interdependent self-construal are more likely to express and experience other-focused emotions like shame. That is, when Chinese individuals offer one's personal opinion, the underlying motivations are focused less

on the pride and high positive regard of individual expression, but perhaps more on the potential shame from presenting an inappropriate or offensive statement. In short, for fear of one's idea being received negatively, speakers may therefore be even more reluctant to express an opinion.

Consequently, Chinese speakers may tend to employ pragmatic tools to distance themselves from their statements, such as indirectly implying an idea or communicating with low modality. Low modality markers display a *perceived* degree of uncertainty (regardless of speaker's *actual* certainty), acting as impersonal indicators that make no explicit claim of personal stance (Kärkkäinen, 2003). Rather, they function as face-saving modal forms (Liu, 2013), serving an interpersonal functions of politeness (hedging) as opposed to high modality forms of involvement (emphatics) (Nikula, 1996). Indeed, this modal distinction appears to relate to face (Brown & Levinson, 1987): Whilst Americans' tendency toward high modality orients toward negative face (need to be independent and unimpeded in one's actions), Chinese speakers' tendency toward low modality orients toward positive face, or need to be liked or approved as a member of a group.

However, it must be noted that whilst past research in general has claimed that Chinese communication style is indirect in nature, this indirectness may be context-dependent. Given that collectivism is linked to a tendency for indirect communication derived from a concern for what the listener may think as well as a concern for how one may appear to the other, Brewer and Chen (2007:146) state that the possibility of offending another individual or being negatively appraised by others is a relational, and not collective whole, notion. That is, such concern and care may derive only from an interpersonal connection to an in-group

member, and thus the related indirect orientation of communication should be associated with in-group communication.

What, then, about communication with an out-group member within a relational-collectivist society? This thesis posits that a Chinese speaker may not necessarily approach communication with an out-group member with such a cautious and indirect communication style. Whilst different aspects of cultural variability may affect different spheres of life at the individual level, Gudykunst (1997) claims that independent vs. interdependent self-construal predominates the way in which one interacts with a stranger: Recall Gudykunst's (1997) notion that interdependent self-construal in a relational-collectivist society leads to different value standards in regard to in-group vs. out-group members, combined with the notion of particularistic empathy from Confucianism teachings. That is, concern for the welfare of in-group members may not extend to the out-group. Triandis (1994:173) claims that this behavioral differentiation for RelSC cultures like China is rooted in the level of closeness and interactive interconnectedness felt with in-group members, citing also that IndSC like the US, who have many loose relationships with a wider circle of people, do not have such in-group intimacy and thus do not favor one group over another (a non-bias which he claims IndSC individuals are proud of, whereas RelSC individuals think it natural to favor the in-group). Such a view is supported by Bond and Wang's (1983) meta-analysis of qualitative surveys and psychological experiments regarding out-group aggression in Chinese culture: Whilst reciprocity leads to harmonious relations with peers and superiors, there appears to exist extreme forms of physical aggression toward out-group others that is virtually non-existent with the in-group; this divide in behavior may be due to a lack of reciprocity with the out-group, so that the relationship to the out-group is temporary and anonymous, with personal

goals and needs taking precedence and less concern given to normative consideration. Accordingly, if a Chinese individual does not perceive any relational ties to an out-group member, then it follows that he or she may not feel the same sense of empathy or concern for the other's thoughts or feelings, thereby reducing concern for maintaining harmony and reducing fear of negative appraisal. Consequently, in the presence of conflict with an out-group member, the need to maintain mutual face is no longer present, and any calmness or decorum then falls away, allowing individuals to "express their anger freely" (Triandis, 1994:171). In turn, this hypothetically leads to more direct and assertive communication style when speaking with or about strangers.

This performance by Chinese participants accordingly parallels behavior of persons primed with high RelSC in Chen, Chen, and Portnoy's (2009) study: Responses to inequitable treatment were dependent on the participant's RelSC and relationship with the other person, with those primed with high RelSC being more lenient toward inequitable treatment from a close other but producing far more negative responses if the other was not a close person. Thus, research appears to indicate that individuals in a Chinese culture exhibit discriminating behavioral tendencies toward in-group vs. out-group members coinciding with a RelSC influence, with out-group members such as strangers receiving significantly less consideration than given to the in-group. In sum, it appears that for those with a primarily RelSC culture such as China, a speaker's social behavior toward the other may be determined by whether the other is considered to be interdependent of or independent of the speaker, with low social distance signifying an interdependent relationship and high social distance signifying the lack of such a relationship.

In sum, a RelSC culture such as China appears to focus less on the individual self and more on the in-group, subsequently restraining one's personal needs and desires and instead behaving as a function of social norms in order to fit in and maintain in-group harmony. Although modern-day China is rapidly transforming via westernization and international influence (Zou et al., 2010; Yum, 1988), the prevalence of a perceived cultural consensus (see section 2) may still exist. A relational-collectivist focus on a "correct" general opinion may continue to be perpetuated, resulting in more cautious communication style and concern for the thoughts and appraisals of in-group members, although perhaps at the expense of the out-group. The next section will explore how pragmatic knowledge is realized cross-linguistically to indicate speaker attitude and intent.

2.2. Pragmatic realization via focus-marking

Recall that an utterance's meaning lies not only in *what* is said, but also in *how* it is said (Grice, 1989). Speakers alter how they utter statements in order to direct a listener to an intended interpretation. The manner in which an utterance, or speech signal, is presented relies heavily on prosody (Wilson & Wharton, 2006; Pierrehumbert, 1980/1990). Prosody refers to how elements, or lexical items, of the speech signal are grouped (syntactically parsed based on pragmatic interpretation) and given prominence (Pierrehumbert, 1980). Wilson & Wharton (2006) purport the notion of prosody as the "packaging" rather than "content" of an utterance; this analogy parallels with Grice's (1989) comparison of *how* it is said vs. *what* is said, stressing the importance of *how* an utterance's "packaging" is prosodically presented.

Placing prominence within a prosodic grouping of syllables, or "focus-marking," allows speakers to pragmatically guide the inferred meaning of their utterance. This section discusses the conceptual framework behind focus-marking[12]. Viewing pragmatic discourse as a cognitive tool instead of mere text, Wichmann and Blakemore (2005:1538) assert that prosody's role involves adjusting or maintaining the accessibility of referents. That is, prosody is used to create conversational implicatures and direct speaker interpretation (Calhoun, 2005). Hence, focus-marking functions pragmatically to reduce the number of possible interpretations, whilst simultaneously making the speaker's intended interpretation more psychologically salient (Wilson & Wharton, 2006; Wichmann & Blakemore, 2005). This manipulation may be viewed through information structure (IS)[13], which explores how each lexical item refers back to, alters, or updates the existing discourse (Calhoun, 2007:54). New, and thus more semantically significant, information is accented as the focus of an utterance, while non-accented items in spoken language are considered given, or background information (Tench, 1990/1996).

Within IS, focus-marking may be explicated via alternative semantics (see Rooth, 1985/1992/1996). Alternative semantics states that all non-focal lexical items maintain their original semantic values, whereas focus-marked items take on a different semantic value than their original lexico-semantic identities (Rooth, 1985/1992). That is, whilst some lexical items do not function pragmatically to transform discourse, focus-marked items under the lens of alternative semantics redirect or update the conversational implicature.

Focus-marked items possess a set containing all semantic objects of the same semantic type, known as the "focus semantic value" (Zimmermann & Onea, 2011; Rooth, 1985). For

[12] More on how prosodic prominence is phonetically realized will be discussed in the following section.
[13] See Steedman, 2000; Jackendoff, 1977/2002.

example, a speaker utters a proposition (*p*): "I think cats are nice." If the speaker had placed prosodic prominence on the word, "*I*," then an alternative set would be created:

X think(s) cats are nice,

where *X* is any object of the same semantic type, including:

[I, you, Mary, John, etc....]

By placing prosodic prominence on "I," the speaker contrasts "I" with all other possible exemplars. Alternatively, if the speaker had focus-marked the word, "think," then the following alternative set would be created:

I *Y* cats are nice,

where *Y* is any object of the same semantic type, including:

[think, feel, believe, guess, know, etc...]

This set contains all objects known to the listener in the semantic class of mental belief verbs (i.e. the semantic type of the focus-marked word, "think").

Focus-marking directs implicature by indicating speaker intent to choose the focus-marked object over any other object. Essentially, focus-marking brings up all alternative

possible semantic objects of the same semantic type in order to negate them (Zimmermann & Onea, 2011). In the first case with pronoun emphasis on "*I*", the conversational implicature created is "I, but not you or Mary or John etc., think cats are nice." In the second case with verbal emphasis on "*think*", a different conversational implicature is derived from the same utterance: "I think, but do not necessarily know, cats are nice." Note that if the cognitive verb uttered was the strong verdical verb of "know" instead of a weaker verdical verb such as "think, believe, feel, etc.," focus-marking the "I" would result in the same meaning as with other cognitive verbs of belief; however, focus-marking the verb for "know" would create the same semantic set but invoke a different meaning: I know, as opposed to merely thinking, cats are nice (Cappelli, 2009:155-6). Thus, by varying prosodic prominence, one utterance may provide completely different semantic interpretations, whose implied meanings are derived from the nature of the cognitive verb chosen.

Yet, while the universality of prosodic focus-marking as a pragmatic tool has been presented in past research, there appears to be little empirical evidence specifically regarding universality vs. cross-linguistic variation of this pragmatic-prosodic link. Thus, this thesis proposes a Proof of Concept phase to be implemented as Phase 1 of this study to establish an empirical link.

2.2.1. Cross-linguistic realizations of focus-marking

Research shows that prosodically, *kontrast* (i.e. "contextually contrastive" item, behaving akin to focus-marking to signify prominence; Wichmann et al., 2009:10.) appears to be realized universally as a pragmatic tool for presenting information, but how it is realized

cross-linguistically may vary (Zimmermann & Onea, 2011). Kontrasts may be signified through various prosodic cues, including pitch, syllable duration, and loudness (intensity). Because the phonetic correlate of pitch is fundamental frequency (F0) of voice (Pierrehumbert, 1980/1990), pitch is often equated with F0 by many studies involving prosody[14]. Hence, because F0 is the production result of pitch, this thesis will refer to pitch as F0.

There is congruent evidence that use of pitch to signify kontrasts is interpreted similarly by both American English and Mandarin speakers. In a study incorporating F0 variation into Mandarin speech synthesis of a native speaker reading a corpus, Zou, Bao, and Luo (2010) found that F0, often dubbed "intonation" (Pierrehumbert, 1980/1990), appears to be the most important feature for prosodic prominence and pragmatic interpretation in Mandarin. Lehiste's (1970) theoretical framework on suprasegmentals (F0, intonation and duration of speech) expresses a similar sentiment regarding English, with speakers relying more on F0 and less on duration to signify focus. Although English speakers use other prosodic cues (intensity, duration) to signify focus (see Breen, Fedorenko, Wagner, & Gibson, 2010), multiple studies express incongruity on whether Mandarin speakers interpret these variables as prosodic cues for prominence.

Chen and Gussenhoven (2008) found that while both Mandarin and English rely on F0 and duration to signify prominence, how these cues are interpreted appear to vary cross-linguistically. In English, F0 movement seems to increase gradually with increased focus, whereas in Mandarin, it seems more of a binary distinction between focal/non-focal words, with no significant difference between different levels of emphasis. Conversely, duration was

[14] For examples, see Ouyang & Kaiser, 2013; Cao, 2004; Chen, Chen, & He, 2012; Chen & Gussenhoven, 2008.

found to increase gradually with increased focus in Mandarin, but appeared binary between focal/non-focal words in English, with little perceptible increase between varied levels of emphasis (Arvaniti & Garding, 2007; Lehiste, 1970). However, Chen and Gussenhoven (2008) make this assertion based on the contrast between new information focus and corrective (contrastive) information focus, claiming that corrective focus is more emphatic than new focus. Subsequently, given that no real pragmatic difference exists between new vs. contrastive information for kontrasts, results merely show that both English and Mandarin speakers use pitch and durational cues similarly to mark kontrasts.

Despite these claims that Mandarin speakers use duration as a prosodic cue, results have been inconsistent. In an audio-perception study asking speakers to identify kontrasts via prosodic cues, Wang (2008) found that although Mandarin speakers were able to similarly detect F0 variations in English spoken discourse, they were significantly less effective at detecting focus via intensity and duration. Perhaps durational cues are not as consistently accurate across studies as F0 cues are because of Mandarin's tonal nature vs. English's accent nature. As a stress-based language, English increases syllable duration to increase stress; more relative stress corresponds to more prominence. However, the different tones in Mandarin possess inherently different durations; for instance, the third tone produced in isolation is 1.5 times longer than other tones. Hence, higher syllabic duration in Mandarin may derive from inherent tonal properties instead of degrees of prominence.

Nonetheless, Mandarin speakers' acuity to F0 detection--perhaps due to the tonal nature of Chinese as a language--allows F0 to serve as accurate a cue of kontrast for Mandarin speakers as for English speakers. Given F0's significance, combined with its relevant cross-linguistic similarities for signifying focus, this thesis will build a conceptual framework for

analyzing communication style based on how prosodic prominence, marked through F0, functions as a pragmatic tool in conversational discourse.

Although focal intonation may be seen as universally driven by cognitive attitude (Pierrehumbert, 1980), different languages vary in their manner of manipulating F0 to indicate lexical identities (Zimmermann & Onea, 2011). The next section discusses cross-linguistic variation of F0 in signaling prominence.

2.2.1.1. Focus-marking via pitch in English

In an intonation, or stress, language like English, focus is realized via increased F0. This thesis adopts the current predominant method of prosodic analysis: ToBi (Tones and Break Indices) (Beckman & Hirschberg, 1999). ToBi is a systematic framework for analyzing prosodic intonation and structure in spoken language (see Pierrehumbert, 1980).[15]
In English, a kontrast within a syntactic grouping is marked with a pitch accent, which functions as the grammatical equivalent of focus in IS (Zimmermann & Onea, 2011:1657). ToBi marks pitch accents and prominence via stress patterns, where the relative maximum point of F0 indicates placement of kontrast. ToBi's linear style model of prosodic analysis measures these F0 accents in each syntactic grouping, which ToBi calls an "intonational unit." From those F0 points, ToBi determines a trend of pitch movement. From this global

[15] For practical purposes, only ToBi indexes directly relevant to F0 for focus-marking (see section 2.2) will be discussed.

pitch movement, a nuclear accent (focus of entire utterance) is determined in intonation languages (Pierrehumbert, 1980/1990; Zimmermann & Onea, 2011; Ladd, 1996)[16]. Zimmermann and Onea (2011) note that while pitch accent and focus have a one-to-one correlation in English, such mapping may not hold true for other languages. Flemming (2008) asserts that Mandarin lacks F0 accent completely, but nevertheless it uses F0 variation in a different manner to mark focus. The mapping of F0 realization to focus-marking in Mandarin is discussed in the next section.

2.2.1.2. Focus-marking via pitch range in Mandarin

Mandarin focus is realized via similar variables as in English. However, while English focus is realized through F0 raising, this changes in a tonal language like Mandarin. Before discussing the difference between prosodic variation between English and Mandarin, it is important to first explicate the nature of tonal languages.

Mandarin employs four distinct tones for lexical distinction. Each lexical item is a single syllable, and contains one tonal marking[17], which establishes a syllable's local pitch contour. Table 2.1 lists English equivalents for Mandarin tones (Duanmu, 2000).

[16] Because this thesis is investigating a particular syntactic construct that functions pragmatically over the whole utterance, this discussion will center on prosodic realization of kontrasts within each syntactic grouping and not on global intonation of the entire utterance

[17] In rare instances, certain syllables do not possess a tone marking; however, these syllables are lexical items that do not possess any lexico-semantic identity. Instead, they are textual pragmatic markers, in that they alter the global meaning of an utterance. For example, 'ma' is inserted at the end of an utterance to transform it into a question. Because these textual markers have no semantic content, they cannot be prosodically altered to create different conversational implicatures, and thus will not be included further in this literature review.

Table 2.1: Mandarin tones and ToBi values

	Qualitative description	ToBi equivalent
Tone 1	High tone	H
Tone 2	Rising tone	LH
Tone 3	Low tone	L
Tone 4	Falling tone	HL

Thus, while English relies on F0 raising to alter pitch contour and mark prominence with little F0 limitation from lexical identity, Mandarin relies more heavily on F0 contour to determine a syllable's lexical identity. Cutler, Dahan and Donselaar (1997) address this cross-linguistic issue by stating that when considering the role of prosody in pragmatic interpretation, it is necessary to distinguish between prosodic structure above vs. within the word level. Prosodic structure/variation above word level (independent of the word's lexical specifications) qualify as contributing to pragmatic interpretation. However, prosodic structure within a lexical item works to facilitate lexical access and determine lexical identity instead of pragmatic interpretation of an utterance.

Wu (1996b) recognized the importance of this distinction in Mandarin, asserting the necessity of separating lexical F0 contour of tonal structure from F0 variation resulting from pragmatic motivations. F0 contour, limited by four tones, was therefore teased apart from "F0 register," or the average F0 value calculated according to the high point (maxF0) and low point (minF0) of each syllable (Wu, 1996b; Cao, 2004:33). Based on Wu's (1996b) work, Cao (1998) attempts to address this issue by analyzing recordings of Mandarin news announcements and attempting to break down the audio into F0 contour and register. Cao found similar results that while Mandarin F0 contour remained relatively stable to signify

lexical identities, the F0 register varied between different syllables; this variation in F0 register differed within- and between-tones.

Building on his work, Cao (2004) used his method for calculating F0 contour vs. register to differentiate between local (lexical) and global realizations of intonation. When F0 registers were connected through lines drawn through them to identify a trend of F0 movement for a standard (declarative) utterance, this revealed a downward global intonation akin to English's. Schack's (2000) comparison of cross-linguistic intonation patterns in native Mandarin-English speakers demonstrates this cross-linguistic similarity of falling global F0 in declarative sentences. Thus, regardless of local F0 contour differences arising from tonal vs. accent languages, Mandarin and English appear to be cross-linguistically comparable in utilizing F0 movement as a prosodic cue of pragmatic interpretation.

Cao (2004) concluded that relative variability of F0 register reflected universal aspects of pragmatic intonation, where a raise in F0 register indicated prosodic prominence. However, what this conclusion failed to consider is that different Mandarin tones are prone to inherent, discrepant F0 values based on natural phonetic tendencies of F0 contour shape. For instance, the third tone tends to be expressed at lower F0 due to its dipping contour shape. If F0 register were calculated for a kontrast of this tone, its inherently low minF0 may cause an average F0 to appear lower than a non-focal syllable with an inherently higher F0 value like the fourth tone, regardless of the third tone’s kontrast possessing higher maxF0. Whilst this may not significantly influence global intonation trends, it may be troublesome for identifying exact kontrast locations within prosodic groupings that compose an utterance.

Mandarin focus may be more accurately realized through variation of F0 range expansion, or raising/lowering of minF0/maxF0 depending on a syllable's lexical tone. Chen

and Gussenhoven (2008) found that when discourse context was altered to prompt Mandarin speakers to respond with varied degrees of focus, focus-marked nouns possessed a larger F0 range than non-focal items of the same tonal contour, despite lexical contours of Mandarin. Although different tones inherently possess varied F0 ranges, the largest F0 range within a prosodic grouping appears to indicate prosodic prominence (Jin, 1996). That is, kontrast is marked by maximum F0 range relative to other lexical items. Using F0 range as a cue for prominence therefore seems more reliable than F0 register, as F0 range avoids the problematic effect of inherent F0 variation on F0 register calculation. Comparing the role of F0 range expansion to F0 accent by elucidating behavioural similarities, Flemming (2008) states that Post-focal F0 range compression creates the same sort of downward F0 trend movement as post-focal deaccenting, and F0 deaccenting may even be more accurately characterized as F0 compression. This cross-variable trend emphasizes the prominence of kontrast in both languages. Thus, based on the similar behaviours of F0 accenting and F0 range expansion, prosodic focus-marking in Mandarin and English appear fundamentally comparable. Schack's (2000) cross-linguistic comparison of intonation patterns further asserts the need for ToBi to add F0 range as evidence of Mandarin focus-marking. Thus, F0 range expression, with an expanded pitch range on focus-marked items followed by a post-focal compression of pitch, appears to function similarly as a signifier for focus across Mandarin and English speech.

In sum, this thesis concludes that F0 manipulation is used cross-linguistically in both Mandarin and English as a tool for focus-marking to direct pragmatic interpretation. English does so through F0 raising, and Mandarin through F0 range expansion. Given a congruous cross-linguistic global intonation trend, F0 raising in English and F0 range expansion in

Mandarin may be treated similarly as prosodic cues for pragmatic inference.[18] Thus, both these manipulations of F0 may be viewed as expressions of F0 range to signify focus-marking.

2.3. Effect of language knowledge on communication and discourse

This section investigates language knowledge's influence on discourse through L1 transfer and L2 proficiency.

2.3.1. Influence of L1 Mandarin transfer on EFL discourse

L1 transfer can affect L2 discourse in areas such as semantics, syntax, phonetics, prosody, and pragmatics[19] (Kasper & Rose, 1999:81). If L1 influence results in target-language-like behaviors in the L2, then positive transfer occurs and does not hinder L2 development; however, if L1 pragmatic influence causes non-native like behaviors in the L2, then negative transfer occurs and may subsequently hinder L2 communication (Wang, 2011). Whilst lexical tones exist in Mandarin but not English, they do not prevent the existence of intonation (Schack, 2000:25), as exhibited in the equity of F0 range to F0 accent. Thus, prosodic transfer across these languages should not present as a confounding factor. Nonetheless, because speakers learn to associate pragmatic intent with particular prosodic cues (i.e. pragmatic-prosodic mapping, as previously discussed) based on cultural input, L1

[18] This framework for prosodic variation will be addressed further in the methodology (section 4).
[19] Given the context of this thesis, the focus is L1 Mandarin transfer onto EFL discourse, and only areas of transfer relevant to pragmatic function of [I + verb] will be explored. Therefore, being only tangentially relevant, cross-linguistic lexico-semantic, phonetic, and syntactic similarities will not be discussed.

transfer may be possible regarding pragmatic knowledge in L2 form-function mapping. A speaker's use and acquisition of L2 pragmatic knowledge is referred to as interlanguage pragmatics (e.g. Cai &Wang, 2013; Kasper & Rose, 1999).

In particular, Kasper and Rose's (1999) meta-analysis of interlanguage pragmatic studies found that interlanguage pragmatics has persistently focused on pragmatic transfer from the L1 to L2; methodology from these studies ranged from judgment tasks to rating scales to both verbal and written discourse completion tasks (DCTs). Of particular relevance, Kerekes' (1992) study on perceived assertiveness used rating scales of audio-recorded native speaker dialogues to determine that EFL speakers' perceptions of qualifier speech acts such as "I think" became less native-like with decreasing L2 proficiency for males, whilst low proficiency female EFL speakers performed similarly to native and high proficiency speakers. L2 proficiency's effect on pragmatic knowledge will be further discussed in the next section.

Regarding how pragmatic transfer may affect conversational discourse, a speaker in one environment may associate a global F0 rise with an interrogative statement, while in another environment, interrogative statements may be signaled by downward intonation (e.g. American vs. British English). Given the sociocognitive nature of discourse, cross-linguistic variations in pragmatic-prosodic mappings may lead individuals to interpret the same prosodic contour in dissimilar ways. However, given focus-marking's universality and the overall similarities of prosodic cues (see section 3.3), this should not present as a large factor in Mandarin speaker's EFL discourse.

If anything, cultural variation may affect how pragmatic cues are perceived as discourse style between Mandarin/English speakers. Lu (2002) stresses how cross-cultural

miscommunications may arise since speakers follow their own cultural beliefs and norms in discourse, including culturally variable interpretations of universal cues. However, how may the amount of Time immersed in a new sociolinguistic environment affect pragmatic knowledge? Having conducted a discourse completion task (DCT) of request and apology situations, Blum-Kulka and Olshtain (1986) found that the longer the time that Non-native speakers of Hebrew resided in a Hebrew sociolinguistic community, the more native-like their performance regarding perception of pragmatic tools for directness and positive politeness, with L2 learners acquiring a native-like performance after five to seven yeas of stay in the L2 environment. These results are further supported by Blum-Kulka and Olshtain's (1984) study, which found that participants residing longer in an L2 environment were more likely to accept target speech-act behavior than those who had resided there for a shorter length of time, with native-like performance after eight years in the L2 environment. Together, these findings thereby suggest that Time immersed in a new L2 sociolinguistic environment may affect a speaker's L2 pragmatic knowledge, with increased time in an L2 environment corresponding with increased native-like norms of pragmatic language use.

Alternatively, Kasper and Rose (1999) surmise from their meta-analysis that L2 learners may possess a sort of "universal pragmatic knowledge," as long as their L2 proficiency reaches a certain threshold for linguistic competence (p. 87). They note that the studies they analyzed mostly used intermediate to advanced EFL learners, or at least participants whose L2 proficiency was advanced enough to complete the elicitation task presented to them. Perhaps if EFL learners were not all of a proficiency threshold that allowed them full access to the range of pragmatic speech acts, then there would be particular differences in the patterns of speech acts utilized by L2 learners. Whether or not this difference correlates to L1

pragmatic transfer is another interesting point of investigation. Accordingly, cross-linguistic pragmatic knowledge is investigated in Phase 1, and any variation found between pragmatic-prosodic mappings, as well as the potentially confounding variable of Time immersed in a L2 sociolinguistic environment, are taken into account for Phase 2.

The next section explores how L2 proficiency of EFL learners may influence their use of L2 pragmatic language.

2.3.2. Influence of L2 proficiency on EFL discourse

Pragmatic tool use of foreign language learners may differ from native speakers' in two manners. Firstly, in comparison with pragmatic marker use of native speakers, low L2 proficiency of a non-native speaker may correlate with underuse of pragmatic markers that fulfill discourse-organizing functions (Aijmer, 1997; Huang, 2011). These include modality (see section 3.2) and discourse markers fulfilling an interpersonal function (i.e. negative face, hedging). Indeed, this notion coincides with Trosberg's (1987) role-play study on apology strategies of Native English/Danish vs. Danish EFL learners: Results showed that the use of modality markers increased with proficiency for the EFL participants so that higher proficiency speakers more closely mirrored performance of native speakers, a point that formed a clear trend of frequency of use given that Native Danish speakers' usage was over double that of Native English speakers'. Thus, these results indicate a developmental pattern of L2 pragmatic acquisition in relation to increased L2 proficiency, thereby suggesting less likelihood of Ll pragmatic transfer.

Nonetheless, Kasper and Rose (1999) suggest that based on numerous studies regarding pragmatic use of apology strategies of various EFL learners such as native Cantonese (Rose, 1998), Japanese (Hill, 1997), and Danish speakers (Trosborg, 1987), EFL speakers appear to possess the same range of speech act realization strategies as native speakers, regardless of L2 proficiency level. Note that this notion does not necessarily contradict the above claims regarding dissimilarities in frequency of speech act use between native and non-native speakers. Indeed, Clark (1979) stated that while non-native speakers, regardless of L2 proficiency, did have access to the full range of apology strategies, their use and implementation of these strategies differed from native speakers depending on social or discourse context. Blum-Kulka and Olshtain's (1986) previously mentioned study furthers this notion, suggesting that at lower proficiency levels, L2 learners tend to avoid verbosity due to a limited knowledge of the L2 language. Therefore, mere access to range does not necessarily imply that L2 learners will perform similarly to native speakers regarding pragmatic knowledge and use of pragmatic speech acts, with low L2 proficiency potentially leading to less verbosity and consequently less use of pragmatic speech acts.

Secondly, low L2 proficiency may correlate with overuse of discourse markers as gap fillers, or what Baumgarten and House (2010) call "the verbal routine form" of a discourse marker. These markers present as speaker hesitation, de-emphasizing the assertiveness of the speaker's utterance (Khojastehrad, 2012:183). Verbal routine forms do not carry any pragmatic meaning and thus do not alter conversational implicature (Stenström, 1994); rather, they indicate "insufficient knowledge" of what to say next (Baumgarten & House, 2010:1197). These discourse markers are used exclusively for expressing difficulties with online-planning (cognitively determining a course of linguistic action), and are used textually

to provide processing time to think of the next statement (Wu, Wang, & Cai, 2010). Liu (2013) found that Mandarin EFL speakers overused the discourse marker "I think" in comparison to native speakers, but admitted that much of the overuse can be attributed to EFL speakers' use of "I think" to "collect thoughts, thus using it as a pause filler" instead of as a pragmatic tool (p. 160). Whilst Liu (2013) admitted that up to 28% of this overuse may be attributed to online-planning difficulties, the study incorrectly classified all instances of "I think," including main clause use and epistemic modality use, as discourse markers. Such a miscategorization of [I + verb] forms may have led to misrepresented results of frequency. This discrepancy with mislabeling of pragmatic forms of [I + verb] will be further discussed in section 3.

It is important to note the functional distinction between interpersonal (pragmatic) forms vs. verbal routine (gap filler) forms. While textual verbal routine discourse markers fulfill a different cognitive function than interactional/interpersonal discourse markers (Stenström, 1994:59), these forms are prosodically indistinguishable[20] (Wu, Wang, & Cai, 2010; Liu, 2013), with both forms receiving no focus-marking. Because native speakers use this prosodic variation to signify interpersonal discourse markers of negative face and hedging, they may misinterpret EFL speaker's use of verbal routine forms as interpersonal forms. This misinterpretation may lead native speakers to view EFL speaker's difficulties in L2 fluency as signs of weakness or lack of confidence, thereby leading to a potential misconstrued view of EFL speakers' communication styles. Such misinterpretation is especially significant given the importance of directness and confidence in communication in American culture (see section 2.1.3.1.) and may produce unwarranted negative judgments. The next section

[20] The prosodic nature of these forms will be discussed in section 3.

introduces a particular pragmatically ambiguous construction through which these areas of pragmatic inference and speaker perception will be explored.

3. CONCEPTUAL FRAMEWORK AND RATIONALE:

PRAGMATIC FUNCTION OF [I + VERB] BELIEF CONSTRUCTIONS

This thesis investigates the above topics through a specific syntactically ambiguous construct: The [I + verb] belief construction (see Dehe & Wichmann, 2010 for overview). While sociolinguistic environment's influence on pragmatic-prosodic mappings may be manifested through a host of linguistic forms, this particular [I + verb] structure was chosen based on three factors: Firstly, because multiple lexical item composition allows for variation in kontrast[21] placement and thus different prosodic contours. Secondly, its high usage frequency has allowed it to gain multiple cognitive forms whilst maintaining the same surface structure; this pragmatic ambiguity results in varied pragmatic function based on prosodic form, thereby allowing investigation of different pragmatic-prosodic mappings. Thirdly, this construct occurs frequently and in similar contexts (presupposing an opinion) in both Mandarin and English, thus allowing for cross-linguistic comparison.

Variation in kontrast placement may lead to three separate prosodic forms of this construct, each of which fulfills a different pragmatic function. Before describing these forms, it is important to discuss the nature of [I + verb]. In both English and Mandarin, indicative statements include a presupposed notion of speaker belief. Thus, actively presupposing an indicative statement of opinion with the matrix clause, "[I + verb]", such as "I think," "I feel," or "I believe" ('wo juede,' "wo ganjue," "wo xiangxin") appears semantically unnecessary, especially since the profile determiner of the statement is the indicative opinion clause (Langacker, 1987). That is, instead of saying "I think cats are nice", one could simply state, "Cats are nice," to convey the same meaning, as it is clearly

[21] Recall from the literature review that "kontrast" refers to "contextually contrastive", so that kontrast-marking is akin to focus-marking to signify prominence (Wichmann et al, 2009:10).

the speaker's own opinion. Because belief is implicit in the opinion clause itself, an [I + verb] construct may thereby appear to be semantically redundant, for then what new information could it possibly contribute? Nevertheless, despite their seemingly superfluous nature, [I + verb] may actually work to function pragmatically as a cognitive tool for facilitating discourse.

[I + verb] constructs in Mandarin are formed similarly to those in English. Because both are Subject-Verb-Object (SVO) languages, this construct is formed by combining the first person singular pronoun ("I" in English; "wo" in Mandarin) with a verb of cognitive belief. Just as the English verb, "think," is the most flexible and common mental belief verb used in these constructions, its Mandarin counterpart "juede," is the most common, flexible verb in Mandarin [I + verb] (Huang, 2011). While "think" may be the most flexibly used, other cognitive verbs, such as "guess," "believe," and "feel," and their roughly Mandarin counterparts "juede" ('think'), "xiang" ('guess'), "xiangxin" ('believe') and "renwei" or "ganjue" ('feel') are also susceptible to polysemous interpretations (Fus, 2008).

Based on these cross-linguistic similarities (further discussed in literature review), the syntactic forms of [I + verb] appear comparable between Mandarin and English. Subsequently, in the following sections, English [I + verb] constructs will be used in examples, which should correspond to their Mandarin counterparts[22].

3.1. Syntactic variation within [I + verb] belief constructions

[22] This assumption will be tested in Phase 1 of this thesis.

Syntactically, a matrix clause acts as the main clause of the utterance, with the embedded (complementary) clause designated as a subordinate clause (Dehe & Wichmann, 2010). Recall the previous example, "I think (that) cats are nice." This utterance would normally be analyzed so that the matrix clause, "I think," functions as the main clause, whilst the embedded clause, "cats are nice," acts as the subordinate clause. However, taken from a "cognitive-functional perspective," certain instances of indicative opinion statements being preceded by a matrix clause of belief may be alternatively classified (Kaltenbock, 2009). That is, if function overrides form, then this utterance may contextually be analyzed so that the embedded clause, "Cats are nice," functions as the main clause, rendering the matrix clause of belief to be a subordinate clause.

This derives from Langacker's (1991) theory that only the main clause may act as the profile determiner of the utterance; that is, only the main clause may carry semantic prominence. In this example, the semantic information contained in the embedded clause, "Cats are nice," may be deemed as more prominent to the utterance than the semantic information expressed by the "I presupposition" of the [I + verb] construct, the reasoning being that belief in an indicative opinion statement such as "cats are nice" is tacitly assumed in expression of the opinion itself. However, if the point of the utterance was for the speaker to express that this statement (i.e. that "cats are nice") is *his* opinion and not another speaker's, then in this context [I + verb] carries more semantic weight and retains the function of the matrix clause. Furthermore, what of "I think" as a hedger or filler? Carrying no semantic weight, this form may be used to save face or minimize confrontation in discourse.

These various analyses of the same utterance demonstrate the functional flexibility of [I + verb]. The next section will categorize these variations into syntactic forms.

3.1.1. Syntactic categorization of [I + verb] forms

[I + verb] may be classified into three pragmatic forms: main clause (MC), epistemic marker (EM), and discourse marker (DM) (Quirk et al., 1985:1112). Dehe and Wichmann (2010) relate these three cognitive categories to Diessel and Tomasello's (2001) three uses of mental verbs: MCs embody an assertive function, EMs a performative function, and DMs a formulaic function.

Accordingly, these classifications of MC, EM, and DM fulfill three pragmatic functions regarding speaker relation to his/her utterance. Assertiveness of language use may be linked to what Nikula (1996) dubs the modality of involvement vs. politeness, or directness vs. indirectness. Thus, more assertive functions of [I + verb] should indicate more confident, direct attitudes in communication. The parameter for each [I + verb] form is outlined in Table 3.1. MCs appear the most assertive, committing the speaker directly to his or her statement (Dehe & Wichmann, 2010:27); EMs appear the second most assertive, performing like an adverbial modifier (akin to “perhaps” or “maybe”) to lessen speaker commitment to the statement (Kärkkäinen, 2003:17); DMs, having been completely grammaticalized, appear the least assertive, functioning either as a hedger of passive politeness or as a formulaic filler for online-planning of speech (Fus, 2008:1702; Baumgarten & House, 2010; Aijmer, 1997).

Table 3.1: Variables determining [I + verb] form

Form	Pragmatic function	Semantic integrity vs. grammaticalization	Can occur with complementizer?	Syntactic position	Multiple occurrences per utterance?
Main clause (MC)	Expresses utterance's main proposition while focusing on speaker's (vs. listener's) mental state (Dehe & Wichmann, 2010:27).	Retains original semantic information	Yes (takes argument)	Sentence-initial	Not possible (acts adverbially on whole utterance) (Fus, 2008:1702).
Epistemic marker (EM)	Expresses degrees of speaker commitment or certitude toward truth of utterance (Fus, 2008:1700; Palmer, 1986:51).	Conventionalized meaning akin to adverbial modifiers (Kärkkäinen, 2003:17).	Yes (due to MC's frequent occurence with complementizer)		
Discourse Marker (DM)	Interpersonal hedger to save face	Completely grammaticalized (Fus, 2008).	No	Any	Possible (has lost all original meaning)
Gap Filler[23]	None: indicates difficulties in online planning and language processing (Baumgarten & House, 2010) akin to "um" or "well" (Aijmer, 1997).				

[23] This non-pragmatic DM form (see section 3.1.3) is included for comparison.

Note that whilst Dehe & Wichmann's (2010) research on sentence-initial [I + verb] distinguishes between MC vs. CC (comment clause; kontrast placed on verb much like EM in this study) forms of [I + verb], it seems unclear of what comprises a CC, with their work suggesting that CC may encompass both EMs functioning to expressing speaker attitude as well as DMs used for interpersonal/interactional purposes (i.e. encompassing two different [I + verb] pragmatic functions), but perhaps not including DMs used non-pragmatically as gap fillers (p. 37). However, Dehe (personal correspondence, 2016) notes that the term CC did not have a clear definition even at the time of publication. Furthermore, given that interpersonal/interactional DM use vs. gap filler DM use are virtually indistinguishable in terms of prosody and syntactic position (see section 3.1.3), adopting this term of CC would result in difficulties deciding whether a particular DM instance of [I + verb] belonged to this category. Thus, in order to refine a more defined prosodic-pragmatic framework, the umbrella term CC was foregone in lieu of the separate, prosodically-distinguishable pragmatic classifications of EM and DM. This distinctive cue of prosody to signify the pragmatic function of [I + verb] forms is discussed in the following section.

3.1.2. Alternative semantics revisited: Prosodic cues for [I + verb] forms

This section explores how different [I + verb] forms are inferred through prosodic cues. Because pragmatic tools create context-dependent implicatures, they are employed based on one-time situational use, making context a relevant cue for form (Fus, 2008; Sperber & Wilson, 1997). Regarding linguistic cues, Dehe and Wichmann's (2010) ToBi analysis of "I think" and "I believe" in British English found that prosody was the most reliable cue for

determining pragmatic function. Table 3.2 revisits alternative semantic forms for prosodic variations in focus-marking of the same statement (see section 2.2.).

Table 3.2: Kontrast placement to alternative semantics implicature for [I + verb] variations

Kontrast location	Example (italics = focus-marked)	Alternative form put forth by [I + verb] belief construction
Pronoun	*I* think cats are nice.	*X* think(s) cats are nice.
Verb	I *think* cats are nice.	I *y* cats are nice.
None	I think cats are nice.	None

Regarding linking prosodic form to pragmatic function, DM appears to correspond to the non-kontrast condition; this mapping is based on alternative semantic forms produced, as well as on DM research (refer to section 2.3.2). More focus indicates more semantic prominence; because DMs by definition retain no semantic information, they receive no focus-marking.

With DM forms corresponding to the no kontrast prosodic variation, rhis leaves two [I + verb] forms to be mapped to two prosodic variations. Given the alternative forms in Table 3.2, Table 3.3 uses focus semantic value and contrastive information theory to link these [I + verb] forms to pragmatic function, and thereby to syntactic form.

Table 3.3: Pragmatic function of [I + verb] prosodic variation via focus semantic value

Kontrast location	[[focus semantic value]]F = {set of semantic objects}	Pragmatic function	[I + verb] Form
Pronoun	[[I]]F = {me vs. anyone else}	Kontrast on speaker = increased commitment to statement	MC
Verb	[[think]]F = {think vs. know}	Kontrast on uncertainty verb[24] = decreased commitment to statement	EM
None	N/A	No kontrast = interpersonal marker of reluctancy to present one's statement	DM

Thus, alternative semantics correlates with Dehe and Wichmann's (2010) prosodic determination of how kontrast placement signifies [I + verb] form. MC corresponds to pronoun focus-marking, EM to verbal focus marking, and DM naturally to no focus-marking.

3.1.3. The DM conundrum: Contextual clues for [I + verb] forms

While MC and EM may correspond to distinct prosodic variations, DM (interactional, interpersonal) and gap filler forms both share the same intonation pattern (no focus). Thus, teasing apart these forms in discourse appears complex. Differentiation may best be determined through qualitative, contextual clues, as there is no clear-cut rule for

[24] Note that in the case of "know", being the one certainty verb of belief, kontrast on this verb indicates an increased commitment to statement as per alternative semantics (see section 2.2): Opposite to the alternative semantic possibilities set forth by uncertainty verbs such as "believe" (Cappelli, 2007:155), focus-marking the strong verdical verb of "know" indicates, "I know factively, as opposed to merely believing." Thus, it may be concluded that the EM form of "know" may be treated as a certainty marker, which indicates an explicit move by the speaker to point out belief and commitment toward statement and therefore exhibits high speaker confidence (Cappelli, 2007:156). Hence, EM form of "know" appears to be akin to its MC form in terms of high speaker commitment toward statement and shall be treated as such.

quantitatively distinguishing them. Others have also found this challenging, with many (e.g. Huang, 2011; Liu, 2013; Fus, 2008) lumping all DM forms (and sometimes even including EM and MC forms) as an overarching "DM" label.

While this issue may appear problematic, it closely mirrors people's real-time pragmatic process: a listener can only rely on contextual clues to determine interactional/interpersonal DM from gap filler use, and communication is just as dependent on speaker intent as listener interpretation. However, fillers may still be considered one DM form, as they still function pragmatically (see section 2.3.2). Both gap filler and interactional/interpersonal DMs result in decreased speaker confidence through varied hedging. Thus, because the end result is the same in terms of speaker stance, these forms may be treated similarly.

3.2. Developed framework for determining pragmatic function of [I + verb]

The following framework for classifying [I + verb] instances in discourse has subsequently been developed (see Table 3.4a/3.4b). Both prosodic and syntactic cues are utilized to determine pragmatic function of [I + verb]. However, given the use of prosody as a universal cue for prominence, combined with the potential for syntactic movement during spontaneous speech, prosodic cues take precedence in the event of conflicting prosodic-syntactic cues.

Table 3.4a: Prosodic cues for determining pragmatic function of [I + verb]

	Main clause	Epistemic marker	Discourse marker
Stress on PRN	X		
Stress on V		X	
No stress			X

Table 3.4b: Syntactic cues for determining pragmatic function of [I + verb]

	Main clause	Epistemic marker	Discourse marker
potential "that"	X	X	
p medial/final			X
2+ times use per utteance (Fus, 2008:1702)			X

From this framework, a measure of communication style was developed and labelled as confidence (Table 3.5), where confidence decreases from left-to-right. This continuum equates to speaker confidence since [I + verb] forms function pragmatically to convey speaker stance.

Table 3.5: Continuum of speaker stance and confidence for [I + verb]

Deliberate	Uncertain	Reluctant
MC	EM[25]	DM

[25] Excluding the certainty verb of "know," for whom the EM form, being deliberate, is treated like its MC form (refer to Table 3.3). Given this similarity in deliberateness and resulting high level of speaker

In summary, [I + verb] seems an opportune construct for observing deviations in pragmatic inference and cross-cultural speaker perception. As no studies as of yet have quantitatively investigated syntactic-prosodic mappings of [I + verb] between American English and Mandarin Chinese, this thesis aims to fill this research gap by comparing and contrasting pragmatic interpretation and then use of [I + verb] in expressing an opinion in a cross-cultural and cross-linguistic analysis. English speakers may use more assertive [I + verb] forms (MC), or simply not precede opinions with [I + verb], such directness resulting in a more assertive utterance. Conversely, Mandarin speakers may employ more tentative, cautious forms (EM, DM). Additionally, EFL learners may misinterpret the intent of native speakers and vice-versa based on linguistic behaviour and L1 pragmatic knowledge, as cultural influences from an L1 environment may shape pragmatic knowledge in the L2 English so that it develops differently from a native speaker's. These pragmatic misinterpretations may be further augmented by limitations of L2 proficiency, leading to misjudgment of speaker intent. These notions will be further discussed in the following methodology chapter.

---

confidence between EM and MC forms of “know,” kontrast on the verb for “know”(EM) is thus classified an MC from hereon out.

4. METHODOLOGY

This chapter begins with an outline of the research design, followed by presentation of research questions and hypotheses. Next, Phase 1 participants, materials, procedure, analysis, and pilot data are described, subsequently followed by Phase 2 participants, material, procedure, analysis, and pilot data.

4.1. Research design

This study investigates how sociolinguistic background affects pragmatic knowledge and discourse style, and how variations in pragmatic interpretations between speakers of different sociolinguistic backgrounds affect perceived speaker confidence. These factors are investigated through a particular pragmatically ambiguous construct: [I + verb] preceding an opinion.

Recall from section 1 that because opinions inherently express speaker belief, the addition of [I + verb] belief constructs is redundant. Pragmatically, however, they may perform either boosting (MC) or hedging (EM, DM) functions. While frequency of [I + verb] has previously been analyzed for high frequency verbs like "think" and "believe," these studies have primarily focused on their function as discourse markers[26]. Additionally, viewpoints differ in what constitutes DM vs. non-DM (EM, MC) use, resulting in inconsistent analyses of [I + verb]'s performative functions (see section 3.1.3). Furthermore, studies have not always addressed the variable of L2 proficiency on usage frequency, resulting in misinterpreted findings.

[26] See Huang, 2011 and Liu, 2013 for review.

Consequently, given these incongruous classifications and methodological shortcomings, this thesis proposes a Proof of Concept phase to identify the cognitive function of each [I + verb] form through participants' interpretations of [I + verb] prosodic variation. After validating the conceptual framework, Phase 2 examines potential differences in [I + verb] usage between English and Mandarin native speakers, as well as Mandarin EFL speakers in English, and whether language knowledge and sociolinguistic background correlate with these pragmatic variations. Additionally, two sets of opinion prompts are analyzed (Self-opinion vs. Opinion of others) to examine any potential differences in expressing opinions regarding the self vs. social judgment, given that these two cognitive processes have been the primary focus of most cross-cultural research on self-construal to date (Cross, Hardin, & Gercek-Swing, 2011).

Variables that may affect frequency and form of [I + verb] usage include sociolinguistic background, L1 pragmatic transfer, and L2 proficiency. Considering these variables may help determine whether a speaker's intent of [I + verb] pragmatic usage differs from its perceived output. Low-proficiency speakers of English may overuse [I + verb] due to lack of oral fluency; this overuse of self-referral may cause one to be misperceived as self-centered, or overuse of gap filler use may cause one to seem more hesitant or reluctant than intended. Conversely, one might predict that US sociolinguistic background may motivate English speakers to use the assertive [I + verb] function, while CHI sociolinguistic background would motivate speakers to use the tentative [I + verb] function more often. However, these notions of directness vs. indirectness in expressing an opinion may be context-dependent, and may potentially differ between expressing opinions regarding the self vs. social judgment.

When bilingual speakers leave their native L1 environment for an L2 cultural environment, differences in the new sociolinguistic environment may influence how speakers communicate in L2 discourse by affecting pragmatic knowledge. This effect of sociolinguistic experience on cognition may augment with increased time exposed to the new L2 environment. In order to better understand the communicative difficulties that inevitably occur when individuals immigrate, this thesis contributes a novel empirical Proof of Concept experiment examining universality of pragmatic-prosodic mappings via the effect of focus-marking in [I + verb] on perceived speaker confidence, and a novel mixed-methods experiment investigating how frequency and form of pragmatic [I + verb] usage may vary cross-culturally and cross-linguistically when expressing Self-opinion and Opinion of others. Through investigating speaker intent vs. listener interpretation, a better understanding of how L1 sociolinguistic background may pragmatically affect how L2 communication can be obtained.

4.1.1. Research questions

The following research questions (RQ), as previously stated in section 1.1, were investigated:

Phase 1:

1) Are specific prosodic variations of [I + verb] belief constructions linked to different pragmatic interpretations in discourse?

1a) For Native vs. Mandarin speakers of English, do these pragmatic-prosodic mappings vary as a function of sociolinguistic background?

2) Does sociolinguistic background influence pragmatic attitudes toward speaker confidence?

Phase 2:

3) Does sociolinguistic background influence frequency of functional [I + verb] usage when expressing Self-opinion vs. Opinion of others?

3a) For the "Opinion of others" prompt set, does sociolinguistic background influence whether participants express a negative appraisal when asked for their impression of the character in the story?

4) Does frequency of functional [I + verb] usage vary as a function of Time immersed in a new sociolinguistic environment?

5) For Mandarin EFL speakers, does L1 pragmatic knowledge transfer influence the functional use of [I + verb] belief constructions in expressing an opinion in English?[27]

6) For Mandarin EFL speakers, does frequency of functional [I + verb] usage in expressing an opinion in English vary as a function of L2 proficiency?[28]

Table 4.1 presents an overview of tasks and which RQs they address.

[27] Note: RQ 5 depends on the results from Phase 1. If pragmatic-prosodic mappings are found to vary cross-linguistically between Mandarin and English, then L1 negative pragmatic transfer on SLA may act as a confounding variable, and RQ 5 will thereby be investigated in Phase 2. However, if mapping is found to be cross-linguistically similar, then L1 transfer is positive and thereby rendered irrelevant as a confounding variable, given that both languages draw upon similar pragmatic-prosodic mappings.

[28] RQ 6 does not directly address functional use but correlation of frequency of use for different prosodic variations (see section 3.1.3 for reference).

Table 4.1: Task descriptions, purpose, and relevant RQs addressed

Phase	Task	Description	Purpose	RQ addressed
1	Language questionnaire	Survey regarding linguistic/cultural background	Screen participants for group inclusion	[screen for participants]
	Perception task: quantitative questions	Main trials of confidence rating	Validate conceptual framework; determine relevancy of RQ 5	1, 1a, 5
	Perception task: qualitative questions	Open-ended definition of "speaker confidence"	Investigate sociolinguistic background's effect on speaker confidence	2
2	Lex30	Productive vocabulary test of L2 proficiency	Screen participants for group inclusion; quantitatively evaluates proficiency of L2 language use in discourse	6
	Language questionnaire	Survey regarding linguistic/cultural background	Screen participants for group inclusion; gather information on Time immersed in new sociolinguistic environment for EFL US group	[Screen for participants] 4
	Production task	Open-ended prompts regarding opinions	Quantitative analysis: Determine frequency of usage of different [I + verb] forms	3, 4, 5, 6
			Qualitative analysis: Investigate sociolinguistic background's effect on expression of appraisal	3a

4.1.2. Hypotheses

Based on the conceptual framework and past literature, the following hypotheses were formed regarding each RQ (Table 4.2). Predictions are made where stated.

Table 4.2: Hypotheses for each RQ

RQ	Hypothesis, null hypothesis
1)	H1: Prosodic variation and pragmatic attitude of [I + verb] statements are correlated in terms of confidence rating (CR). Specifically, it is predicted that CR will decrease from DIR to PRN to V to no kontrast conditions for prosodic variation, and from Deliberate to Uncertain to Reluctant conditions for pragmatic attitude. H0: CR will not decrease from DIR to PRN to V to no kontrast conditions for prosodic variation; and from Deliberate, Uncertain to Reluctant conditions for pragmatic attitude.
1a)	H1: Sociolinguistic background affects trend of CR of [I + verb] statements. H0: There is no effect of sociolinguistic background on trend of CR of [I + verb] statements.[29]
2)	H1: Sociolinguistic background affects pragmatic attitudes toward speaker confidence. Specifically, it is predicted that US individuals will mirror the IndSC mindset salient within an individualist society, whilst Chinese

[29] Based on past research on universality of focus-marking (see section 2.2), this thesis predicts that results will support H0 instead of H1 for 2).

	individuals will mirror the RelSC mindset salient within a relational-collectivist society. H0: There is no effect of sociolinguistic background on pragmatic attitudes toward speaker confidence.
3)	H1: Sociolinguistic background affects the frequency of functional [1 + verb] usage. Specifically, it is predicted that higher levels of US sociolinguistic background will lead to using more assertive and less reluctant forms of [I + verb] when expressing Self-opinion, whilst higher levels of CHI sociolinguistic background will lead to using more assertive and less reluctant forms of [I + verb] when expressing Opinion of others. H0: There is no effect of sociolinguistic background on frequency of functional [I + verb] usage.
3a)	H1: Sociolinguistic background affects the expression of negative appraisals in the Opinion of others prompt set. Specifically, it is predicted that frequency of negative appraisals increase as the level of CHI sociolinguistic background increases and the level of US sociolinguistic background decreases. H0: There is no effect of sociolinguistic background on the expression of negative appraisals in the Opinion of others prompt set.
4)	H1: Frequency of functional [I + verb] usage varies as a function of Time immersed in a new sociolinguistic environment. Specifically, it is predicted

	that the longer a speaker is immersed in a new sociolinguistic environment, the more a speaker's performance regarding frequency of functional [I + verb] usage will mirror a native speaker's frequencies. H0: A speaker's performance regarding frequency of functional [I + verb] usage does not increase in similarity to a native speaker's performance the longer one is immersed in a new sociolinguistic environment.
5)	H1: For Mandarin EFL speakers, L1 pragmatic transfer affects the functional use of [I + verb] belief constructions in expressing an opinion in English. H0: For Mandarin EFL speakers, L1 pragmatic transfer does not affect the functional use of [I + verb] belief constructions in expressing an opinion in English.
6)	H1: For Mandarin EFL speakers, usage frequency of different prosodic [I + verb] belief constructions in expressing an opinion in English varies as a function of L2 proficiency. Specifically, it is predicted that DM use of [I + verb] constructions will increase as L2 proficiency decreases. H0: For Mandarin EFL speakers, DM use of [I + verb] constructions does not increase as L2 proficiency decreases.

4.2. PHASE 1: Proof of Concept

An interpretive, mixed-methods task was designed to determine Proof of Concept for pragmatic interpretation of prosodic variations of [I + verb]. In the following sections, information regarding participants, materials, procedure, analysis, and pilot data are presented.

4.2.1. Participants

Participants were fluent native speakers who have resided most of their lives (<2 years in any other country[30]) in the US (Group 1) or China (Group 2) to account for variables of L2 language knowledge and additional sociolinguistic background[31]. All participants were aged 18+. Participants were recruited via snowball sampling with personal contacts in both countries, as well as through recruitment at universities via email adverts. Due to the online nature of this Proof of Concept task, group 1 participants were recruited from various locations within the US, whilst group 2 participants were recruited from various locations across China.

Because pure random sample selection is not usually possible, a large enough sample size was gathered to allow variation of gender, age, and geographical location within each target country (Paltridge & Phakiti, 2015:85, 105). A power effect size calculation indicated a minimum sample size of N = 52 (two groups of 26) necessary to achieve the

[30] As there does not appear to be any established minimum amount of time immersed in a new environment before significant cultural influence occurs (given cross-cultural and between-person variability of personal characteristics like age and degree of interaction with native speakers), this thesis has established an minimum amount of time immersed as 2 years based on Gladwell's (2008) 10,000 hour rule for learning a new (pragmatic) skill. Because this minimum applies equally to all participant groups, it should not present as a confounding factor.

[31] See Table 4.9 for details on participant characteristics for group inclusion.

.80 observed power threshold for reliability (Cohen, 1988). 60 participants completed the task in total. However, one participant within group 2 was excluded upon further investigation (see section 5.1). Thus, remaining sample size was n = 30 in Group 1 and n = 29 in Group 2, thereby satisfying the sample size requirement.

Participants were aged 18 to 71, of which 34 were males and 26 were females. Mean age was 35.27 (std. error = 2.54) for US group 1 and 44.21 (std. error = 2.67) for China group 2. Levene's test showed equality of variances (p = .86), but since an independent samples t-test showed a significant between-group age difference of 8.94 years (F(1,57) = .03, std. error = 3.680, p = .02), age was run as a covariate in analyses. Given that there were the exact same number of males and females in each group, a one-way ANOVA showed no significant effect of group on Gender (p = 1.00), thereby suggesting that gender should not act as a confounding variable. Additionally, given that gender is not relevant in answering any of the research questions, it was not further included in analyses.

4.2.2. Materials

Language questionnaire: An online-language questionnaire screened for participant-group inclusion based on language and sociolinguistic background (Appendix 3). This questionnaire has been adapted from Zhao's (2013) study on bilingual French-English vs. monolingual English children. This questionnaire appears as the first page before the perception task survey. Participants may only proceed to the English or Mandarin

perception task if their answers indicate that they are a monolingual speaker of English or Mandarin, respectively, resulting in all data being produced by relevant participants.

Perception task: An online survey was created via Qualtrics. This survey was validated through pilot testing (see section 4.2.5) (Paltridge & Phakiti, 2015:35). This task resembles a combination between Meyer and Mleinek's (2006) study, which used an electronic survey to present audio recordings in a perception categorization task involving prosodic interpretation of questions/statements in Russian, and Kurumada, Brown and Tanenhaus' (2012) electronic Human Intelligence Task involving likelihood ratings of semantic interpretations of prosodically varied "It looks like X" statements, which similarly examined prosodic variation's effect on semantic interpretation.

Based on research on syntactic form and function of [I + verb] (see Dehe & Wichmann, 2010), six possible [I + verb] prosodic-syntactic variations were determined[32] (Table 4.3).

[32] Note that because gap filler form does not function prosodically, it is not included in Phase 1's study of pragmatic-prosodic mappings.

Table 4.3: Prosodic-syntactic [I + verb] variations and labels

Prosodic-syntactic form (based on kontrast placement)	Label for study's use	Associated pragmatic form	Example of construct (***kontrasts in bold-italics***)
Direct statement (no [I + verb])	DIR	Direct	"It's a #."
Kontrast on pronoun	PRN	Main clause	"***I*** think it's a #."
Kontrast on verb	V	Epistemic marker	"I ***think*** it's a #."
No kontrast ([I + verb] sentence-initial)	FRONT	Discourse marker	"I think it's a #."
No kontrast ([I + verb] sentence-medial)	MID		"It's, I think, a #."
No kontrast ([I + verb] sentence-final)	END		"It's a # I think."

Different prosodic-syntactic variations of target nouns in [I + verb] utterances were audio-recorded from a native speaker (the researcher) with a "typical" accent in each language. To confirm a standard speaker accent, at the end of the pilot was an open-ended question in both the Chinese/English survey versions asking whether the speaker in the audio recordings sounded like a typical native speaker. Taking into account the numerous Mandarin dialects existing in China, responses from the pilot phase confirm overall standardness of accent (see section 4.2.5). Notably, using the same speaker for both surveys beneficially removes speaker variation as a confounding variable.

"Think" ('juede' in Mandarin), the most commonly used belief verb, was used across trials. All target nouns were imagable, common words and were chosen for their neutrality and contextual plausibility. All target nouns belonged to the OED's frequency

band 5 or 6, which should ensure comprehension from such monolingual participants. These were audio recorded with one of the six prosodic variations and presented for the listener to click and hear. Each recording was followed by a question asking the participant to choose whether the speaker in the recording sounds, "Deliberate," "Uncertain," or "Reluctant", with a short definition following each term. This question is based on the structure of Meyer and Mleinik's (2006) categorical decision task. This was then followed by a second question asking the participant to rate the confidence of the speaker in the recording on a Likert scale from 1-7. This question is derived from Kurumada, Brown and Tanenhaus' (2012) confidence rating task in Experiment 2[33]. Originally, it was not specified as to which end of the scale indicated a higher vs. lower confidence rating. However, one participant during pilot testing remarked that not every individual may come to the assumption that 7 indicates highest confidence (although he noted that he did correctly make this assumption). Thus, for the main phase a note has been inserted in question one stating that 1 = lowest and 7 = highest confidence rating. Each recording presented contrastive information through prosodic variation, creating different conversational implicatures through focus-marking. This focus-marking was used to alter the linguistic salience of possible interpretations (Dehe & Wharton 2010:1559), so that different prosodic variations should indicate particular syntactic forms and functions. In turn, this should have directed participants toward one choice. All focus-marking in audio-recordings was checked and confirmed through prosodic analysis on Praat, to ensure that all recordings consistently followed the specific pitch-contours of each [I + verb] variation.

[33] Note: Kurumada, Brown and Tanenhaus (2012) asked participants to rate self-confidence in regard to their answer choice. Whilst the goal of their question differs, the process of confidence rating should hold sufficient for this study.

Originally, a second prosodic variation of the same target noun was presented after these questions, along with a third question asking participants to rate the second recording as more, less, or the same in confidence as the first recording. The foundation for this question was based on Kurumada, Brown and Tanenhaus' (2012) 3-Alternative forced decision task in Experiment 1, but measured confidence instead of likelihood. However, pilot testing showed a large attrition rate of 67%, with many participants stopping their survey before the halfway mark; presumably, a survey of 48 recordings was too long for many voluntary participants to finish. Additionally, when undergoing analysis of pilot data, data from this third question was not necessary to achieve valid results to answer any of the RQs, with all RQs addressed using data from the first two questions. Given this redundancy, the second recording and this corresponding question were removed from the main phase with the benefit of shortening survey length by 50% to avoid such large participant dropout.

Questions and format, as well as a screenshot from an example prompt, may be found in Appendix 4. This task included 24 target nouns with each of the six prosodic variations appearing four times. Originally the pilot tested 24 target nouns, from which the 18 target nouns resulting in the least response deviation were to be included in the main phase. However, because all 24 target nouns tested similarly in terms of standard deviation (SD), there was no reason to choose six for exclusion. The condition distribution for prosodic variation is provided in Appendix 5.

4.2.3. Procedure

Participants were recruited via verbal/email enquiry through academic contacts at American and Chinese universities, and further on through snowball sampling. Universities were selected for A) oversampling of the particular participant specifications (see section 4.2.1) and B) pragmatism (where academic contacts may be reached). Schools fitting these criteria included Washington University and University of Michigan in America, and universities in international cities like Beijing, Hangzhou and Nanjing in China. Because this was an online survey, this task may be completed from anywhere. Qualtrics IP address tracking of participant submissions ensured that all participants completing the task resided in the appropriate country.

Before data collection, participants received written explanation of the study (Appendix 1) and signed an electronic consent form (Appendix 2). The project title on the information sheet and consent forms is abbreviated to avoid participants determining the study's precise topic and providing answers based on perceptions of that topic.

Next, participants filled out an online language questionnaire (Appendix 3). If responses indicated appropriate linguistic and cultural background for group inclusion (see section 4.2.1), participants were immediately directed to the online perception task in their native English/Mandarin.

Participants were provided with written instructions (Appendix 4) and then presented with 24 consecutive trials administered on separate pages[34] (Appendix 5). Trials were

[34] While Meyer and Mleinek (2006) presented all their audio recordings and choices simultaneously on the same page, in this research each audio recording and answer set will be provided on separate consecutive pages so as to minimize distractions and influence from previous answer choices being visible. Additionally, this allows unique randomization of order of questions for each participant, thus minimizing the potential confounding variable of question position.

followed by an open-ended question asking participants to type in their definition of speaker confidence.

4.2.4. Analysis

Participant ID codes were assigned to each participant to ensure anonymity (Paltridge & Phakiti, 2015:40).

Quantitative analysis for the perception task consisted of three mixed-design repeated measures ANCOVAs, as well as a Spearman rank-order correlation test coinciding with the third ANCOVA. For any data with potentially non-parametric data distribution, a Kruskal-Wallis test was also run alongside the ANCOVA and presented in an Appendix. Age was analyzed as a covariate due to a significant mean difference in age between groups (see section 4.2.1). All confidence rating (CR) values were compared within-group and between-group, as presented in Table 4.4.

For Analysis 1, a mean CR value was assigned to each prosodic variation based on participants' CRs for the audio recordings of each prosodic variation. These mean CRs produced a CR trend between levels, which was then compared to the predicted trend in the conceptual framework (see section 3.2) as well as between-groups.

For Analysis 2, a similar mean CR value was assigned to each pragmatic attitude based on participants' CRs for the audio recordings rated with that specific pragmatic attitude; from this, a CR trend of perceived pragmatic attitude in relation to speaker confidence was observable. This resulting trend was then compared to the predicted trend in the conceptual framework as well as between-groups.

Analysis 3 mirrors the between- and within-group factors of Analysis 1, differing only in its dependent variable: Instead of using participants' CR responses given for each prosodic variation, it used mean CR values for each pragmatic attitude (Deliberate, Reluctant, or Uncertain), assigning a CR to each prosodic variation based on participants' pragmatic attitude choice for that prosodic recording. For instance, a recording of a prosodic variation was rated for both direct CR of the recording (Analysis 1), as well as for perceived pragmatic attitude. Based on which perceived pragmatic attitude the participant chose for that prompt, a CR value was assigned from the mean CR values of that perceived pragmatic attitude (Analysis 2). Emulating Analysis 1, resulting trends in CR were compared within-group and between-group. A one-tailed Spearman rank-order correlation test was also conducted to further establish the link between prosodic variation and pragmatic attitude. In summary, Analysis 3 served to substantiate a correlation between CR and pragmatic-prosodic variation trend, thereby strengthening the pragmatic-prosodic mapping for [I + verb] constructs.

Table 4.4: Phase 1 analyses and variables

Analysis	Dependent variable	Within-group factor	Within-group levels	Between-group factor	Between-group levels
ANCOVA 1	Confidence rating	Prosodic variation	6 (Direct, Pronoun, Verb, no kontrast (sentence-initial/mid/final)	Cultural Background	2 (US, CHI)
ANCOVA 2	Confidence rating	Perceived pragmatic attitude	3 (Deliberate, Reluctant, Uncertain)		
ANCOVA 3	Calculated confidence rating (based on ANOVA 2's CR values)	Same as ANOVA 1	Same as ANOVA 1		
Spearman rank-order Correlation test	Calculated confidence rating (based on ANOVA 2's CR values)	Same as ANOVA 1	Same as ANOVA 1		

4.2.4.1. Examination of timing

For each participant, timing of answer submissions for each prompt was measured on the Qualtrics survey page. Prior to running analyses, timing of response submissions was examined to ensure that participants performed the task appropriately. If responses were found to be submitted in a time frame shorter than the time required to listen to the audio

clip, then these participants were excluded from data analysis. The justification for this participant elimination (Paltridge & Phakiti, 2015:15) lies in the notion that the participant cannot complete the task properly if a response is submitted before the prompt has been fully played, thereby rendering their performance unreliable as the prompt has not been considered when choosing a response.

4.2.4.2. Coding scheme for open-ended speaker confidence definitions

The coding scheme for this open-ended question was constructed based on the pilot phase results. Five categories were determined. These categories encompassed every participant's response. These five nominal categories are presented in Table 4.5.

Table 4.5: Coding scheme for open-ended speaker confidence definitions

Category	Inclusion criterion
Directness/tone of language	Refers to the language or speech of the statement itself, especially assertiveness or brevity of statement.
Speaker belief toward statement	How much the *speaker* believes his or her statement to be true, regardless of actual correctness or truth or statement.
Listener belief toward statement	How much the *listener* believes the statement to be true, regardless of actual correctness or truth or statement.
Appeal/reputation of speaker	Refers to the reputation or credibility of the speaker's character, vs. the statement itself.
Comfortableness of speaker to share statement	Refers to the speaker's willingness or desire to share his statement.

Due to the open-ended nature of the question, certain participant's responses encompassed 2 categories; this overlap was due to divided multiple part responses

wherein a single participant's response presented two separate ideas (e.g. "Self belief in what they are saying, direct wording without ambiguity"), or compound responses wherein one idea would be embedded in another (e.g. "The words and intonation of the speech convey that the speaker believes what he is speaking is the truth"). No response fell under three or more categories. Thus, this non-mutual exclusivity may be due more to the nature of open-ended responses and unequal observations (Miles & Huberman, 2013:279) rather than anything detrimental to the cluster of semantic categories. Furthermore, categories were exhaustive, and no responses were left unassigned to a category.

While Cohen's kappa may not be directly calculated with non-mutually exclusivity, given that no responses belonged to more than two groups, modified mutually exclusive categories were created from the intersect between each two categories (see Table 4.6). Cohen's kappa was then successfully calculated from these categories and interrater agreement obtained.

Table 4.6: Modified mutually exclusive categories for speaker confidence definition

Categorical coding	Inclusion
1	Directness/tone of language
2	Speaker belief toward statement
3	Listener belief toward statement
4	Appeal/reputation of speaker
5	Comfortableness of speaker to share statement
6	{1, 2}
7	{1, 3}
8	{1, 4}
9	{1, 5}
10	{2, 3}
11	{2, 4}
12	{2, 5}
13	{3, 4}
14	{3, 5}
15	{4, 5}

Mandarin responses were translated into English by two native Mandarin speakers who compared and checked one another's translations. Finalized translations were then reviewed by the researcher for plausibility and accuracy; no discrepancies were found. Interrating was performed by a native English speaker on all responses. As per Paltridge and Phakiti's (2015:41) methods, results for both simple percentage agreement and Cohen's Kappa were calculated. Interrater agreement calculated using simple percentage agreement (# differences / # total units) was 94%, and Cohen's kappa calculated with modified mutually exclusive categories was 90% (std. error = .04, $p < .001$).

Intrarater agreement calculated using simple percentage agreement (# differences / # total units) was 97%, and Cohen's kappa calculated with modified mutually exclusive categories was 96% (std. error = .03, $p < .001$).

Due to the possibility of unequal observations of response data, results were analyzed via non-parametric Kruskal-Wallis testing, with analysis details presented in Table 4.7.

Table 4.7: Analysis for open-ended definition of speaker confidence

Analysis	Dependent variable	Within-group factor	Within-group levels	Between-Group factor	Between-group levels
Kruskal-Wallis test	% of total responses	Semantic category	5 (Categories 1-5 in Table 4.5)	Sociolinguistic Background	2 (US, CHI)

4.2.5. Pilot data

The pilot differs from the main phase in sample size (N = 38 for pilot) and inclusion of enquiry regarding speaker accent of recordings (see section 4.2.2). Otherwise, materials and procedure remain similar. Pilot results are provided in Appendix 6.

To summarize, pilot data substantiated all hypotheses overall. Congruent CR trends in prosodic variation and perceived pragmatic function support a cross-linguistic pragmatic-prosodic link, thereby suggesting the main phase may bear similarly fruitful results. Given cross-linguistic similarity in pragmatic-prosodic mappings, L1 pragmatic transfer should not act as a confounding variable (RQ 5). Even so, between-group variations were

visible through "speaker confidence" definition, and variations effectively matched predicted independence-interdependence differences between US vs. CHI cultural environments; this suggests sociolinguistic background's effect on socio-cognitive perception, despite universality of pragmatic-prosodic mappings.

Originally 18 out of 24 target nouns were to be chosen for the main phase; however, because all 24 target nouns possessed similar SDs ($.95 < SD < 1.43$), there seems no reason to not include all 24. While slightly lengthier, result validity benefits from inclusion.

Qualitative analysis of pilot data resulted in the creation of a semantic coding scheme for the open-ended speaker definition question, wherein five exhaustive categories (see section 4.2.4.2) encompassed all responses.

Participant feedback successfully clarified specified high-low levels of the CR scale for the main phase, and instructions will accordingly label “1” as lowest CR and “7” as highest CR. Additionally, overall “standardness” of accent in recordings was confirmed.

4.3. PHASE 2

A mixed-method, open-ended production task examined the influence of sociolinguistic background on communication style in the genre of expression of opinion. Additionally, given that there is no research on how EFL proficiency may influence frequency of MC/EM [I + verb] forms before an opinion, L2 proficiency was also evaluated as a variable of communication style. Information regarding participants, materials, analyses, and pilot data are presented below.

4.3.1. Participants

Originally, sample size for phase 2 was to be 120 individuals (4 groups of 30) based on a power effect size calculation suggesting 27 participants per group. However, given pragmatic constraints of time (both for data collection and transcription analysis) and participant availability, the final sample size was 100 individuals (4 groups of 25). Group inclusion of participants was dependent on sociolinguistic background and language knowledge, as presented in Table 4.8.

Table 4.8: Group labels and inclusion based on sociolinguistic background

Group	Description
Native US	Monolingual native English speakers in America.
EFL US	EFL Chinese speakers in USA. They will have immigrated to USA and have lived in the USA for 2+ years.
EFL CHI	EFL Chinese speakers in China. These participants will have lived most of their lives in China but have learned (are learning) English as a foreign language.
Native CHI	Monolingual native Chinese speakers in China[35].

Monolingual status was determined by an online language-screening questionnaire (Appendix 3) and was defined as not being conversationally proficient in any other language. [1] Since it may prove difficult to find pure monolinguals given the prevalence of

English as a global language, this group included very low-proficiency EFL speakers, defined as individuals whose English ability was not proficient enough to allow simple conversation. Table 4.9 examines variables determining group inclusion. The choice to examine EFL US participants separately from Native US and EFL CHI groups is based on the notion that these participants would have exposure to both CHI and US cultural influences, a distinction which Heine and Hamamura (2007) justify in their meta-analysis of East Asian vs. Western cross-cultural research. The Native CHI group serves to act as a monolingual control group for EFL CHI participants, given that they both share the same Chinese cultural influence, without having resided in and been exposed to US cultural influence. The two EFL groups differ in the potential amount of US sociolinguistic influence present: the EFL US participants, currently residing in the new L2 environment, are assumed to possess more US sociolinguistic influence than the EFL CHI group, the latter of which should only be exposed to US sociolinguistic influence through learning the L2 in their L1 home environment of China.

Table 4.9: Sociolinguistic background of groups, broken down into linguistic and cultural background

	Native English speaker	EFL Chinese speaker	Native Chinese Speaker
Individualist (IDV) cultural environment = salience of Independent Self-construal (IndSC)	Native US group		
Relational-collectivist (REL-COL) cultural environment = salience of Relational Self-construal (RelSC)		EFL CHI group	Native CHI group (acts as control for EFL groups)
REL-COL + IDV = RelSC + IndSC		EFL US group	

Whilst the final sample size for data analysis was 100 participants, there were also 3 participants who originally undertook the task but were eventually excluded. This included one participant from the Native US group, who upon completion and debriefing of the task revealed to the researcher that he had actively decided to stop using the exact [I + verb] constructions from his speech in an effort to sound more "direct" and "clear"; accordingly, his performance on the production task was deemed as unnaturally altered (by himself) and his data was excluded. Another participant from the EFL CHI group and one participant from the Native CHI group failed to complete the production task due to what they cited as difficulties thinking of enough things to say to fill up the whole 15-second response timeframe; these participants' incomplete data sets were thereby

excluded as well. Given that participant attrition only amounted to 2.91% of the total sample size, and that attrition occurred almost equally across all groups as opposed to being biased within one group, there should be no confounding effect of selective attrition on overall data collected.

Participants were aged 18 to 71, of which 53 were males and 47 were females. Mean age per group is provided in Table 4.10. A one-way ANOVA showed a statistically significant main effect of group for age of participants ($F(1,3) = 7.37$, $p < .001$), with post-hoc paired t-tests revealing significant mean differences in age between groups 2 and 3 ($p < .001$) and 2 and 4 ($p = .03$). Thus, Age was run as a covariate of analysis to avoid a potentially confounding variable.

Table 4.10: Means and SDs for age of participants per group

Group	Mean age	SD
1	36.04	14.09
2	43.80	11.29
3	27.85	9.88
4	34.00	12.84
Total (average)	35.42	13.25

A one-way ANOVA showed no significant effect of group for Gender ($p = .71$), thereby suggesting that gender should not act as a confounding variable in analyses. While some may suggest that gender may play a role in self vs. relational construal and

how this may relate to in-group vs. out-group interactions in Western (US) culture (see Cross & Madson, 1997), the authors also stress that this gender divide may not necessarily assume a similar construct in other cultures, with Cross, Hardin and Gercek-Swing (2011) asserting that gender stereotypes should thereby be independent of theories of self-construal (p. 148). Moreover, given that gender is not relevant in answering any of this thesis' research questions and no significant effect of gender between groups was found, it was not further included in analyses.

4.3.2. Materials

Materials have been updated in relation to pilot study findings (see section 4.3.5). Changes are noted where made.

Lex30: This online vocabulary production test uses simple word association to rate a speaker's English proficiency, and works similarly for both native and non-native English speakers (Fitzpatrick & Meara, 2004). Participants are presented with a list of high-frequency stimulus words and then instructed to produce four typed responses that they feel are related to the word. The high-frequency nature of the stimuli allows the test to be accessible by speakers from a wide range of proficiency (Meara & Fitzpatrick, 2000). Given that for Phase 2 of this study proficiency levels are not known during recruitment of EFL speakers, this wide range is advantageous so that all participants are able to take it. Moreover, because there is no set list from which to choose vocabulary, this test is a free production task (Meara & Fitzpatrick, 2000). Scores are given for the responses

produced based on lemmatization and frequency analysis, with high-frequency function words and the 1,000 highest frequency content words receiving no points and responses outside these categories comprising the Lex30 score (Walters, 2012:175). Reliability testing of the Lex30 suggested validity in its scoring, emphasizing that it worked better in distinguishing low-level proficiency speakers from proficient speakers (Fitzpatrick & Meara, 2004). Criteron-validity for the Lex30 (Paltridge & Phakiti, 2015:14) is also strengthened given a significant correlation between the Lex30 test and a receptive yes/no vocabulary test as well as the EVST test, as shown through Fitzpatrick and Meara's (2004) findings.

Originally, the Oxford Online Proficiency Test (OOPT) was chosen to test for proficiency for this study, given its computer-adaptability, online accessibility, and examination of a speaker's ability to read "between the lines", or implied knowledge (OOPT, 2013). However, in pilot testing, participants voiced concern that the OOPT, at 50 minutes, was too long, especially given their voluntary status in the study; some refused to take it altogether, resulting in the pilot not involving any analysis of EFL proficiency (Sse section 4.3.5). Given this issue of length, the Lex30 replaced the OOPT in the main phase.

The benefits of the Lex30's accessibility are three-fold: Firstly, its short length renders it a much more user-friendly test; whilst similar vocabulary production tests like the LFP and P_Lex require an hour's time (Walters, 2012), completing the Lex30 only takes approximately 15 minutes. Secondly, since the Lex30 is an online test, it may be taken anywhere at anytime. Thirdly, the site tabulates marks instantly. This is especially useful given the pragmatic constraints of time, budget, and travel. Recall that the Lex30

is completed by the participant in the presence of the researcher so as to ensure proper completion; with the Lex30, overall fieldwork length is shortened with travel costs thereby minimized, and results are produced instantly before proceeding to the oral production task. This instant result may be beneficial if a participant has an abnormally low score, which may indicate an inability to complete the production task.

Furthermore, the Lex30 was chosen for its productive vocabulary nature, which coincides with the Phase 2 oral production task. Whilst the Lex30 measures receptive knowledge, Meara and Fitzpatrick (2000:20) argue for the implicit assumption of a correlative relationship between receptive and productive vocabulary knowledge. Thus, an individual's productive vocabulary score on the Lex30 may be implicitly correlated with active productive vocabulary proficiency in the phase 2 oral production task.

Language questionnaire: Questionnaires for the monolingual groups (Native US and Native CHI) are the same as the corresponding monolingual groups in Phase 1. For the EFL groups (EFL US and EFL CHI), questions have been added regarding age, gender, age first exposed to English, language of school and teacher instruction, language spoken at home, and preferred language in which to socialize. This information was collected so as to control for random factors that may otherwise threaten internal validity (Paltridge & Phakiti, 2015:15). For EFL individuals who have immigrated to America (EFL US group), an additional question is included regarding age immigrated to the U.S., so as to calculate the amount of Time immersed in the new cultural environment and identify whether a length of time spent in an L2 environment influences pragmatic [I + verb] usage (RQ 4). All questionnaires may be found in Appendix 3.

Production task: An oral discourse completion task (DCT) was created in order to illicit the [I + verb] construct under investigation. To complement the perception task of Phase 1, this task was a production task. Participants are asked to listen to pre-recorded audio prompts and then respond verbally with no rejoinder.

An oral DCT was chosen over a written one given that according to Yuan's (2001) findings, oral DCTs generate a "significantly larger number of natural speech features than the written DCT does," thus more closely mimicking real-time conversational discourse (Yuan, 2001:271). Additionally, the highly-controlled nature of the audio-recorded DCT prompts renders it easily replicable, therefore strengthening its validity (Paltridge & Phakiti, 2015:15) as well as allowing for comparison of speech acts cross-linguistically (Bardovi-Harlig & Hartford, 2005:11), thereby making it an ideal candidate for gathering data on this study's particular speech act of [I + verb] constructs in both English and Mandarin. Furthermore, this study's oral DCT was validated through pilot testing to ensure successful elicitation of all variations of [I + verb] constructs (see section 4.3.5) (Paltridge & Phakiti, 2015:35).

This production task contains one pre-recorded set of direct opinion questions (N = 9) to elicit responses of self preference, and one pre-recorded set of questions regarding the participant's opinion of another individual (N =9; 2 sets of 3 questions per scenario) to elicit responses of social judgment. In all three question sets of the Opinion of others prompt set, a story is provided in which a fictional character has made a mistake, although the nature of why the mistake was made was purposefully kept vague to allow participants to interpret it as they saw fit. Given that the protagonist in the story is

presented as an abstract, fictional character, they are not presumed to be someone with whom the participant has any direct relationship. Thus, the "other" who the participant is to appraise may be assumed as a stranger and therefore an out-group member, designating high social distance between the character and the participant (Brown & Levinson, 1987).

For the pilot phase, N = 12 for both sets to test which prompts best elicited the desired constructs (see section 4.3.5), with the 9 most successful from each piloted set being chosen. This task has both an English and Mandarin version. Prompts were designed to illicit pragmatic [I + verb] usage by prompting expressions of opinion and thus providing a suitable context for [I + verb] constructs. Prompts specifically do not include any of the verbs that may be included in an [I + verb] construct (e.g. think, feel, believe, guess) so as to avoid any confounding effect of priming. While prompts all express the same sentiment of requesting the participant to offer an opinion, they are worded in various ways to avoid being too repetitive. The finalized list of prompts may be found in Appendix 10.

4.3.3. Procedure

Recruitment for Phase 2 was identical to that of Phase 1. Participants were provided with a written explanation of the study (Appendix 7) and asked to sign a written consent form (Appendix 2).

All participants were administered online language questionnaires (Appendix 3). EFL US and EFL CHI groups were administered the Lex30 in the presence of the researcher.

They were asked to finish the Lex30 as quickly as possible and given the timeframe of 15 minutes as a target; however, as per the procedure for the Lex30 there was no enforced time limit for completion. Every participant completed the task within 30-35 minutes.

All participants were then administered the production task (Appendix 10) in a quiet, undisturbed location. These locations included empty classrooms in university buildings, quiet café settings, and home sitting rooms. Native US, EFL US and EFL CHI- groups were administered the English version; Native CHI group received the Mandarin version. 18 prompts (2 sets of 9) regarding opinion expression were audio-played by the researcher for the participant to hear. Participants were instructed to continue talking until verbally stopped by the researcher, and informed that any inquiry may only be answered with "However you wish to interpret it." Participants were told to inform the researcher if they did not recognize words in the prompts; these instances were noted for analysis. Prompts were then audio-played and paused between each presentation to allow the participant to respond. Prompts were played once so as to ensure spontaneous, natural discourse. However, if the prompt was misheard, participants could request to rehear it once. Upon completion of the task, participants were debriefed on the nature of the research regarding cross-cultural [I + verb] pragmatic variation.

Order of prompts within each set was randomized for each participant. Additionally, order of sets was counterbalanced. Prompts were recorded on a Macbook Pro via Praat at 96000 hz to ensure high quality audio for pitch analysis on Praat.

4.3.4. Analysis

Participant ID codes were assigned to each participant to ensure anonymity (Paltridge & Phakiti, 2015:40). Audio-recorded data from the production task was first transcribed. Then, data for each participant was quantified via frequency counts of each [I + verb] construct variation. [I + verb] frequency counts were tabulated using AntConc 2 software for English transcriptions and AntPConc 2 software for Mandarin transcriptions (Anthony, 2011). Prosodic and syntactic analyses were then performed to determine specific [I + verb] variation (see section 4.3.4.2).

The production task was quantitatively analyzed for frequency of pragmatic forms of [I + verb] usage, as validated in Phase 1 and presented in the conceptual framework (see Chapter 3). To reiterate, based on the pragmatic-prosodic mapping for [I + verb] substantiated in the Proof of Concept phase, Direct (DIR) and Main clause (MC) were considered as 'Deliberate', Epistemic marker (EM) as 'Uncertain,'[36] and no kontrast Discourse marker (DM) at sentence-front (DMF), mid (DMM) and end (DME) positions as 'Reluctant'.

Table 4.11 presents analyses and their variables. Age was run as a covariate (see section 4.3.1). The following analyses were run twice separately: once for part 1 and once for part 2 of the production task.

[36] Except for EM form for the certainty verb of "know," which through alternative semantics (see section 2.2) indicates semantic deliberateness (Cappelli, 2007:156), and is therefore treated as its deliberate MC form in analyses (refer to section 3.1.2. in conceptual framework).

Table 4.11: Phase 2 analyses and variables

Analysis	Dependent variable	Within-group factor	Within-group levels	Between-group factor	Between-group levels
ANCOVA 1	% of usage to total utterances	Pragmatic function	3 (Deliberate, Reluctant, Uncertain)	Sociocultural background (written as Language-Residence for each group).	4 (US-US, EFL US, EFL CHI, CHI-CHI)
ANCOVA 2	% of usage to total constructs				
ANCOVA 3	Calculated CR (ANOVA 2's % * CR value of prosodic variation from phase 1)	Pragmatic-syntactic variation	5 (DIR, MC, EM, DMF/DME, DMM)		

For both sets of the production task, Two ANCOVAs were run with the same within-group factor of pragmatic function and compared between-groups, differing only in the dependent variable (DV) being measured. For each prompt set (Self-opinion, Opinion of others), the number of times a participant used each prosodic-syntactic variation of the [I + verb] construct was counted. From this frequency count, each number was then divided by the total number of utterances produced by the participant and multiplied by 100, resulting in the percentage of each variation's [I + verb] usage to total utterances; this number was also divided by the total number of constructs produced by the participant and multiplied by 100, resulting in the percentage of each variation's [I + verb] usage to total constructs.

These percentages were then grouped according to pragmatic function based on results of Phase 1's pragmatic-prosodic mapping of pragmatic attitudes to variations in focus-marking within the [I + verb] construct: Deliberate (functioning to strengthen speaker commitment to statement), Uncertain (functioning as a face-saving adverbial to lessen speaker commitment), and Reluctant (functioning passively as a hedger). The total percentage of DIR and MC = % Deliberate pragmatic condition; EM percentage = % Uncertain pragmatic condition; total percentage of DMF, DMM, and DME = % Reluctant pragmatic condition. Thus, the DV for ANCOVA 1 was Percentage of pragmatic function of [I + verb] to total utterances, and the DV for ANCOVA 2 was Percentage of pragmatic function of [I + verb] to total [I + verb] constructs used.

These percentages were used to examine the distribution of pragmatic [I + verb] constructs between individuals of different sociolinguistic backgrounds, with the first DV being calculated the same way as the analysis in Lin, Ren and Woodfield's (2012) study regarding frequency percentages of various pragmatic speech acts in a written DCT; the second DV was developed by the researcher as another method for analyzing distribution of pragmatic constructs in regard to total constructs produced. In order to strengthen validity of findings, results from this second DV were compared to the first DV to see whether the two DVs produced similar results regarding distribution frequencies of [I + verb]. Results were then analyzed regarding whether frequency of functional [I + verb] usage corresponded to predicted sociolinguistic behaviours regarding directness and speaker attitude (RQ 3)[37] (see sections 2.1.3.1 and 2.1.3.2).

[37] Refer back to Table 4.1 for detailed task-to-RQ mappings.

Next, mean CR values of each prosodic-syntactic variation[38], as calculated in Phase 1, were used to calculate each participant's overall level of speaker confidence (see Table 4.12). Because Phase 1 results showed no significant between- or within-subject mean difference of CR for No kontrast: sentence-mid (DMM) and No kontrast: sentence-end (DME) conditions, the CR values of these conditions were averaged and these two conditions merged under one category for Phase 2 (DMM/DME). Each prosodic-syntactic variation was multiplied by the percentage of total constructs comprised of that particular prosodic variation for each participant, resulting in an assigned average CR for speaker confidence for each participant. These CRs were then compared within- and between- groups via a third ANCOVA. Results were analyzed to check whether mean CRs of speaker confidence corresponded to predicted individualistic vs. collectivistic differences in mindset based on sociolinguistic background.

Table 4.12: Mean CR values of prosodic-syntactic variations for [I + verb] statements

Prosodic variation: focus marking	Corresponding syntactic function	CR
Direct statement	Direct statement (DIR)	6.17
Pronoun	Main clause (MC)	5.59
Verb	Epistemic marker (EM)	4.67
No kontrast: sentence-front	Discourse marker: sentence-front/ sentence-end (DMF/DME)	3.93
No kontrast: sentence-mid/end	Discourse marker: sentence-mid/sentence-end (DMM/DME)	3.37

[38] Note that mean CRs of prosodic variation vs. perceived pragmatic function are being used due to Phase 1 results indicating significant differences in mean CR between DIR vs. MC, and DMF vs. DMM/DME conditions. While this does not affect the overall CR trend from high-to-low for Deliberate to Uncertain to Reluctant, CR will be calculated via CR of each prosodic variation to account for any potentially confounding variability from these between-level differences.

Next, a dependent samples t-test was conducted for each group at Deliberate, Uncertain, and Reluctant conditions between frequency results from prompt set 1 vs. prompt set 2, in order to identify whether frequency of functional [I + verb] usage varies across groups regarding Self-opinion vs. Opinion of others (RQ 3). These comparisons were conduced for each DV in the analyses above (Percentage of pragmatic function of [I +verb] to total utterances, Percentage of pragmatic function of [I +verb] to total constructs, Calculated confidence rating).

Following this, two in-depth sub-analyses were conducted to investigate whether frequency of functional [I + verb] usage aried as a function of L2 proficiency or Time immersed in a new sociolinguistic environment, respectively. Firstly, Lex30 score for the two EFL groups was analyzed to examine whether L2 proficiency had an effect on frequency of usage for DM (Reluctant condition) [I + verb] forms (RQ 6). Because RQ 6 only concerns itself with DM variation (Reluctant condition) of [I + verb] (see section 3.1.3 for overview of relevant literature), this in-depth sub-analysis involves the DV of frequency at Reluctant condition only. However, in the interest of research, an analysis conducted on all pragmatic [I + verb] conditions is included in Appendix 16 for reference. For Analysis 1 above, the same analysis was run with EFL US and EFL CHI participants only (N = 50) and with Reluctant condition of each prompt set as DVs, subsequently followed by a similar analysis with the addition of the between-subject factor of Lex30 score, as well as the potentially confounding between-subject factors of Age of L2 acquisition and Native language of L2 instructor(s)[39]. The first N = 50 sub-

[39] Originally, Language spoken at school was also considered as a between-subject factor. However, lack of variance (N = 47 Mandarin, N = 1 English, N = 2 Both Mandarin/English) meant a comparison could not be successfully conducted. Thus, this variable was dropped from analysis. Similarly, Preferred language for socialization was also measured, but once again lack of variance within each group (N = 2 English, N = 23

analysis serves to act as a comparison model for the corresponding N = 50 sub-analysis with additional covariates, which acts to answer RQ 6 well as to control for potential influence from such variables.

A second in-depth sub-analysis was conducted on US EFL group only. Quantified values of Time immersed in the new L2 sociolinguistic environment for EFL US group were analyzed by years to see if this variable significantly affected frequency of functional [I + verb] usage (RQ 4). For the second sub-analysis run on N = 50 in Appendix 16 (repeated measures ANCOVA with additional covariates of Lex30 score, Age of L2 acquisition, and Native language of L2 instructor), the same analysis was run again but with EFL US participants only (N = 25, meaning that the between-subject factor of sociolinguistic background was removed due to there only being one group involved); this was subsequently followed by a similar analysis with the addition of the between-subject factor of Time immersed in new sociolinguistic environment. The first N = 25 analysis serves to act as a comparison model for the second analysis, which acts to answer RQ 4 as well as to control for potential influence from this variable.

Table 4.13 presents analyses and their variables. Age was run as a covariate (see section 4.3.1). The following analyses were run twice separately: once for part 1 and once for part 2 of the production task.

Mandarin, N = 0 Both Mandarin/English for EFL US group; N = 0 English, N = 23 Mandarin, N = 2 Both Mandarin/English for EFL CHI group) meant a comparison could not be successfully conducted and resulted in this variable being dropped from analysis as well. Finally, Preferred language spoken at home was also dropped from analysis due to lack of variance (N = 1 English, N = 22 Mandarin, N = 2 Both Mandarin/English for EFL US group; N = 0 English, N = 24 Mandarin, N = 1 Both Mandarin/English for EFL CHI group).

Table 4.13: Phase 2 in-depth sub-analyses and variables

Analysis	Dependent variable	Within-group factor(s)	Within-group levels	Between-group factor(s)	Between-group levels
Sub-analysis 1a	% of usage to total utterances	Pragmatic function	2 (Reluctant for Self-opinion, Reluctant for Opinion of others)	Sociolinguistic background	2 (EFL US, EFL CHI)
Sub-analysis 1a2		Pragmatic Function, Native language of L2 instructor	2 (Reluctant for Self-opinion, Reluctant for Opinion of others); 2 (US/non-Chinese, Chinese only)	Sociolinguistic background, Lex30 score (continuous), Age of L2 acquisition (continuous)	
Sub-analysis 1b			3 (Deliberate, Uncertain, Reluctant); 2 (US/non-Chinese, Chinese only).	Lex30 score (continuous), Age of L2 acquisition (continuous)	None
Sub-analysis 1b2				Lex30 score (continuous), Age of L2 acquisition (continuous), Time immersed in new sociolinguistic environment (continuous)	

4.3.4.1. Length of audio recording

Because this was an open-ended task, it was difficult to set a standard length for each prompted response. Thus, data was compared based on a mean number of utterances per prompt[40].

[40] An utterance is defined as a statement with a subject and verb (whether explicitly spoken or spoken with ellipsis).

Inevitably, real-time measurements of number of utterances spoken cannot be as accurate as measurements of utterance frequency via transcription. However, based on the pilot study (see section 4.3.5), it appeared that most responses contained approximately four to five utterances, with no statistically significant difference in mean utterances per prompt between groups. In order to guarantee enough data was collected from each participant, a minimum length of 15 seconds was accordingly established for each prompt, which was a length that pilot results showed the majority of responses met or exceeded.

If the participant for whatever reason stopped short of the 15 second response timeframe (e.g. they could not think of anything else to say, leaving only silence instead), then these responses were noted as incomplete and excluded from data analysis, so as not to bias the mean number of utterances and thereby amount of data for each full response given.

In sum, for each prompt set of the production task, a one-way ANOVA was conducted to investigate any significant differences in mean number of utterances between groups.

4.3.4.2. Orthogonal and prosodic transcription

Data was transcribed based on Tones and Break Indices (ToBi) guidelines. For pragmatic purposes, ToBi tiers irrelevant to this study's analysis were subtracted, with only relevant transcription indexes (orthogonal, tones) retained.

For orthogonal transcription, relevant information included mispronunciations, punctuation, length marking, filler words, stuttering, pitch rise, unintelligible speech, non-verbal sounds, and silence. Originally, timing down to .1 second was transcribed for each recording with anything less written as (.); however, the laboriousness of this process, coupled with the notion that timing was not necessary for analysis, led to a revision after the pilot phase so that only silences of 1.0 second or greater were noted. A random selection of 10% of English and Mandarin transcriptions may be found in Appendix 7.

Transcription accuracy was calculated via both simple percentage agreement (# edits / # total units) and minimum string difference calculated from Levenshtein distance between two strings (Levenshtein, 1996). Edits encompassed addition, subtraction, and substitution (Kruskal, 1983). Both interrater and intrarater reliability were conducted in order to confirm reliability of results (Paltridge & Phakiti, 2015:15).

For English transcriptions, interrating was performed by a native English-speaking Linguistics student at a US university. Interrater agreement calculated with simple percentage agreement was 92%, and calculated with Levenshtein distance was 93%. Intrarater agreement calculated with simple percentage agreement was 96%, and calculated with Levenshtein distance was 97%.

For Mandarin transcriptions, interrating was performed by a native Mandarin-speaking computer programmer from China. Interrater agreement calculated with simple percentage agreement was 87%, and calculated with Levenshtein distance was 91%. Intrarater agreement calculated with simple percentage agreement was 92%, and calculated with Levenshtein distance was 94%.

A unique computer code for calculating reliability was created for each language. Coding accounted for all similarity and differences necessary for this study's analysis: Punctuation is included in calculations, as utterance parsing may be significant for mean utterance count and semantic interpretation. Transcription features irrelevant to this study, including capitalization, spelling of filler words (um, em, uh, mm, eh), length markers (:), unintelligible speech within double parentheses, and nonlinguistic sounds in brackets ([laughs], [coughs], [sighs]) are not included. These codes may be found in Appendix 8.

For all utterances containing an [I +verb] construct, further prosodic analysis was conducted via Praat (Boersma & Weenik, 2016). Both the global intonation of the utterance and the phrasal intonation of [I + verb] constructions were analyzed, thereby allowing for observation of both pitch (F0) range and F0 accent. Recalling that both Mandarin and English speakers appear to manipulate focus-marking by using F0 as a primary source for conveying prominence[41] (Chen & Gussenhoven, 2008; Zou, Bao, & Luo, 2010), such an analysis should work for prosodic analysis of all groups. As with the orthogonal transcription, only ToBi features relevant to this thesis' analysis were retained. By utilizing basic tone elements of high (H) and low (L), pitch accent and focus-marking may be determined (Niemann et al., 1998). Thus, since levels of stress (i.e. variation in *types* of pitch accents) were not directly relevant for determining [I + verb] variations, inclusion of ToBi's complex pitch accents (L*+H, L+H*, H+!H*) did

[41] Recall that while English uses F0 raising (i.e. F0 accent) and Mandarin uses F0 expansion (i.e. F0 range) to convey focus-marking and prominence, literature suggested a congruous cross-linguistic global intonation trend: Post-focal F0 range compression and post-focal F0 deaccenting both create the same sort of downward F0 trend movement, and according to Flemming (2008), F0 deaccenting may even be more accurately characterized as F0 compression. Subsequently, this similarity indicates that there should be no cross-linguistic limitation to investigating F0 manipulation as a tool for focus-marking in both Mandarin and English.

not appear necessary; furthermore, whilst these complex pitch accents may apply to English, such distinctive variations may not directly transfer to a tonal language such as Mandarin, in which F0 movement to signify focus appears to be a binary distinction between focal/non-focal words with no significant set difference between levels of emphasis. Given these factors, a functional approach was taken in which the two tones of H and L were adopted as tonal targets for analysis, in which H represents a pitch rise and therefore focus-marking (examples in Table 4.14 below). Use of H and L tones should ensure cross-linguistic compatibility whilst still providing the data necessary for determining specific [I + verb] construct variation. That is, treating H/L as tonal targets provides sufficient information for kontrast placement and pitch movement necessary to address the RQs, whilst at the same time this functional approach works for both English and Mandarin, thereby allowing for a cross-cultural system of comparison between these typologically different languages.

4.3.4.3. Coding scheme for [I + verb] constructs

[I + verb] constructs were coded under 5 categories (MC, EM, DMF, DMM, DME) as per the conceptual framework, plus 5 categories for direct opinion utterances (DIR (N is), DIR (partial), DIR (command), DIR (fact), Not DIR), with the last two categories being added for interrating in case the utterance was deemed not to be a DIR opinion statement (For an overview of construct pitch contours, refer to the conceptual framework in Chapter 3).

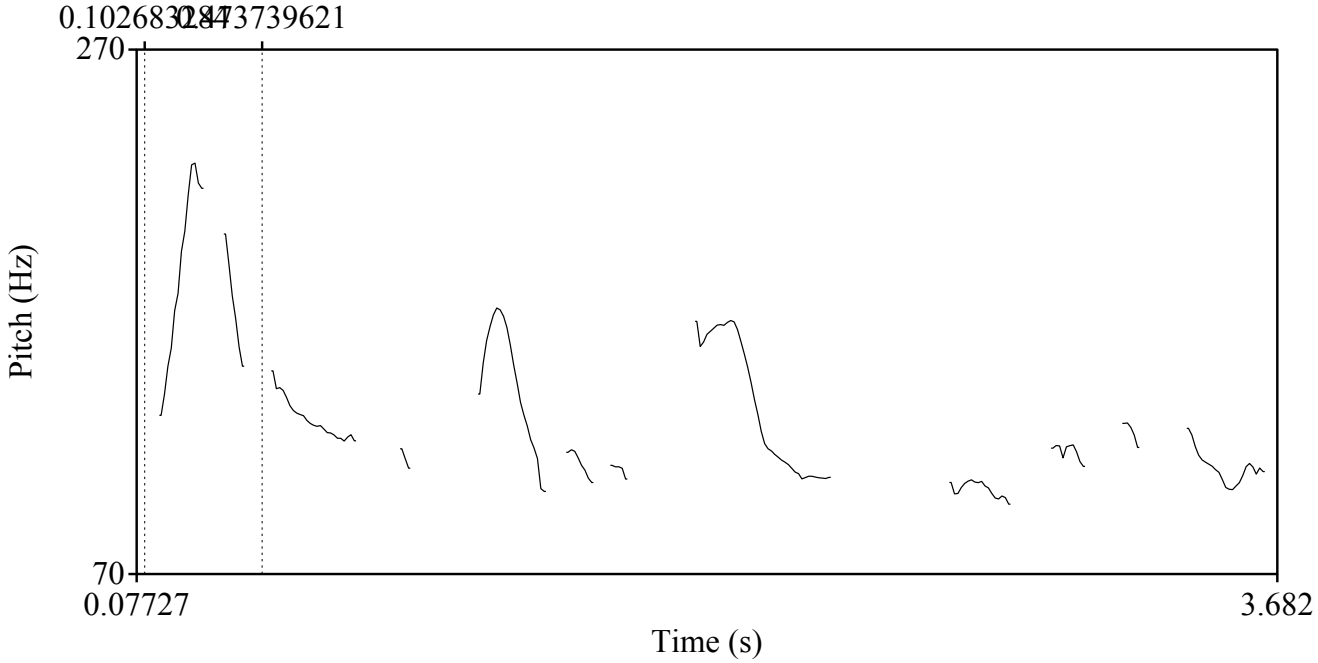
Table 4.14 presents each construct variation's intonation marking. Because determining DMF vs. DMM vs. DME classification of a construct is a syntactic matter, this was done via global pitch analysis of the utterance.

Table 4.14: [I + verb] construct intonation markings

Construct	Pitch accent	Intonation marking
Main clause (MC)	I	HL
Epistemic marker (EM	[verb]	LH
Discourse marker (DM)	None	[None]

Figure 4.1 provides an example of a prosodic analysis for an [I + verb] construct from English transcription; Figure 4.2 presents a Mandarin transcription (both in Mandarin characters and in pinyin spelling) with English translation. Word placement along the bottom correlates to the pitch contour directly above. Constructs are underlined. The last line corresponds to tone markings for the pitch contours of [I + verb] constructs. The pitch contour construct is sectioned off by the dotted lines where the phrasal intonation can be viewed, while the full utterance's pitch contour is presented so that the global intonation may be viewed.

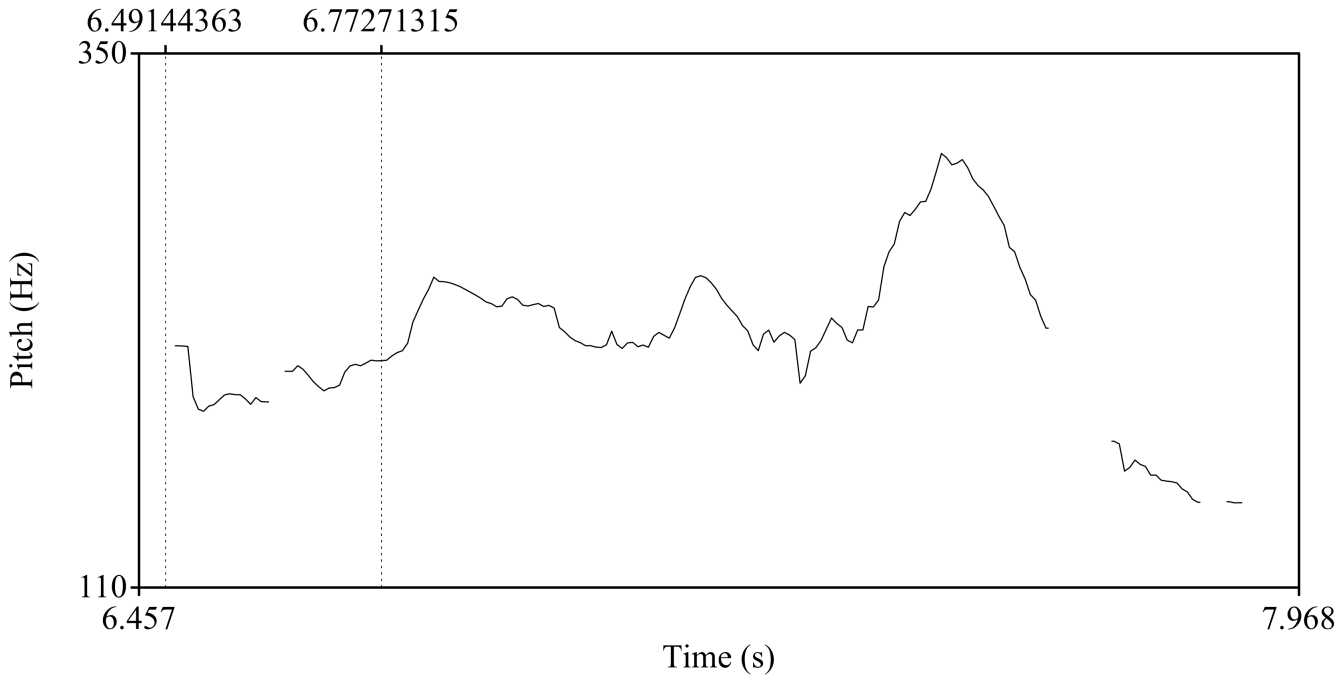
Figure 4.1: Prosodic analysis for English transcription containing MC [I + verb] variation



In Figure 4.1, the speaker began his utterance with an [I + verb] construct ("I think", sectioned off by the dotted lines). Based on prosodic analysis of the pitch contour, it was determined that this was an MC construct. This was evident due to the pitch accent being placed on the "I" (H), with a lowering in pitch afterward for "think" (L). Furthermore, when viewing the global pitch contour, the construct here had the highest F0 of the utterance, making the "I" of the construct the nuclear peak accent and thereby focal point

of the utterance. This peak accent placement was in line with the parameters of an MC construct in the conceptual framework.

Figure 4.2: Prosodic analysis for Mandarin transcription containing DMF [I + verb] variation



我	觉得	外	面	太	热	了
wo	juede	wai	mian	tai	re:	le
I	feel	outside		too	hot:	(modal particle)

[no intonation marking]

In Figure 4.2, the speaker also began her utterance with an [I + verb] construct ("wo juede", or 'I feel'). For this case, however, prosodic analysis showed that there was no pitch accent within the pitch contour of the [I + verb] construct, with both "wo" and "juede" receiving no intonation marking and therefore no focus-marking. Thus, this stimulus was classified as a DM construct; upon syntactic analysis, this [I + verb] instance was further confirmed as a DMF construct due to its sentence-initial position. Global pitch analysis suggested that the peak accent of the utterance occurred on "re" ('hot'), signifying this as the focus of the utterance. This peak accent placement coincided with the conceptual framework's expectation regarding a DMF construct.

Cognitive belief verbs used by participants included "think," "feel," "believe," "imagine," "suppose," "know," "assume," "guess," and "reckon" in English, and "juede" ('think'), "xiang" ('guess'), "ganjue" (feel), and "renwei" (feel) in Mandarin. Both interrater and intrarater reliability were conducted in order to confirm reliability of results (Paltridge & Phakiti, 2015:15). Interrating for English coding (Native US, EFL US, EFL CHI groups) was conducted by the same researcher that did interrating for English transcriptions, and interrating for Mandarin coding (Native CHI group) was done by the same researcher that did interrating for Mandarin transcriptions.

Interraters were given a spreadsheet with four columns: Name of file, timing (at which stimulus occurs), transcription of stimulus, and a blank column to write down their chosen annotation. For English and Mandarin examples of coding set up, see Appendix 11. They were provided with two files: one containing the audio clips and one containing screenshots of the construct's pitch contour (vs. full utterance) stimuli. Based on the timing given in column 2, interraters were instructed to play and view the pitch contours

of the audio clip at that time on PRAAT, using the transcription in column 3 as a guide. They were then told to pick which construct or statement type they thought the stimulus to be. Choices for utterances also included "Direct (fact)" and "Non-direct" in case they felt that the utterance did not fall under any direct opinion category. The instruction sheet given to interraters is provided in Appendix 11. Training was conducted with interraters via Skype on 10 practice stimuli, with one stimulus corresponding to each possible category.

Interrating and intrarating were done based on a random selection of 10% of total constructs used (230 out of 2301 constructs used for English coding; 65 out of 649 constructs used for Mandarin coding). This count thereby surpassed the minimum threshold needed to calculate Cohen's kappa with the expectation of 80% agreement and 20% relative error (Gwet, 2010). As per Paltridge and Phakiti's (2015:41) methods, results for both simple percentage agreement (i.e. before applying Cohen's correction) and Cohen's Kappa were calculated. For English coding, interrater reliability calculated using Cohen's kappa was 89% (std. error = .02, $p < .001$); interrater agreement before applying Cohen's correction for chance was 91%. Intrarater reliability calculated using Cohen's kappa was 96% (std. error = .01, $p < .001$); intrarater agreement before applying Cohen's correction for chance was 97%.

For Chinese coding, interrater reliability calculated using Cohen's Kappa was 90% (std. error = .04, $p < .001$); interater agreement before applying Cohen's correction for chance was 92%. Intrarater reliability calculated using Cohen's kappa was 94% (std. error = .03, $p < .001$); intrarater agreement before applying Cohen's correction for chance was 95%.

4.3.4.4. Coding scheme for Expression of negative appraisal

For each of the three stories in the Opinion of others prompt set, participant responses were analyzed for expression of negative appraisal. Each story set of three prompts contained the same third prompt, which asked the participant, "What is your impression of *X*?" wherein *X* is the abstract fictional character in the story. Participant responses were then semantically coded into two categories: Negative appraisal, and non-negative appraisal. This dichotomous coding was mutually exclusive and exhaustive, with all responses falling under a category and belonging to one category only. Given the open-ended nature of the prompts, it was possible for participants to offer mixed sentiments toward one character (e.g. a negative statement followed by a positive statement). Responses were nonetheless coded based on overall attitude of the response as a whole, such as the response is either more negative than non-negative (e.g. a negative statement with a neutral statement), or more non-negative than negative (e.g. a very positive statement with a hedged negative statement).

Mandarin responses were translated into English by two native Mandarin speakers who compared and checked one another's translations. Finalized translations were then reviewed by the researcher for plausibility and accuracy; no discrepancies were found.

Interrating was performed by a native English speaker on 20% of the responses, resulting in interrating of 60 responses and thereby surpassing the minimum threshold needed to calculate Cohen's kappa with the expectation of 80% agreement and 20% relative error (Gwet, 2010). As per Paltridge & Phakiti's (2015:41) methods, results for both simple percentage agreement (i.e. before applying Cohen's correction) and Cohen's Kappa were calculated. Interrater reliability calculated using Cohen's Kappa was 86%

(std. error = .07, $p < .001$); intratater agreement before applying Cohen's correction for chance was 93%. Intrarater reliability calculated using Cohen's kappa was 93% (std. error =.05, $p < .001$); intrarater agreement before applying Cohen's correction for chance was 97%.

Results were analyzed using a Chi-square test of independence. The association between sociolinguistic background (measured at 4 levels: Native US, EFL US, EF CHI, Native CHI) and expression of negative appraisal (measured at two levels: Negative, Non-negative) in each of the stories was investigated for significance, with post-hoc paired chi-squared tests being performed between each group on significant results.

4.3.5. Pilot data

Differences between the pilot and main phase, as well as a summary of pilot results, are presented below. Full pilot data can be found in Appendix 12.

The pilot survey differed from the main phase in the following ways:

1) Sample size and mode of conduct: Whilst the main phase survey was conducted in-person with 25 participants per group (N = 100), the pilot survey was conducted via Skype with 4 participants per group (N = 16). Skype was chosen instead of in-person testing due to pragmatic travel constraints. While Skype transmission may slightly affect quality of recorded audio analysis, this variable presents across all participants and thus should not present as a group confound.

2) <u>Covariate of age:</u> Whilst for the main phase mean age was found to be significantly different between groups, there was no statistically significant between-group difference in mean age for the pilot sample. Thus, pilot data was analyzed without the addition of age as a covariate, and analyses were run as ANOVA vs. ANCOVA.

3) <u>Length of production task:</u> The pilot survey contains N = 12 for both sets (4 sets of 3 questions for the Opinion of others prompt set), with the 9 most successful prompts being chosen for the main phase. The three prompts omitted from each set with reason for removal are presented below in Table 4.15.

Table 4.15: Prompts omitted from Phase 2 main phase based on pilot results

Prompt	Reason omitted
Set 1: What is your impression of celebrity culture?	A portion of the EFL speakers (38%) did not know the meaning of "celebrity culture".
Set 1: What makes someone a great writer and why?	This prompt was found to be too closely related to another prompt ("What makes a story interesting and why?") but was not as successful in frequency as the other prompt at eliciting [I + verb] constructs.
Set 1: What is your favorite movie and why?	The word "favorite" did not prompt constructs, as the speaker's "favorite" is a subjective but definite fact (vs. an arguable belief).
Set 2: Ella is at the airport to check in for her flight. When the flight attendant asks her for identification, Ella reaches into her pocket but doesn't find her passport. -What has happened? -What should Ella do? -What is your impression of Ella?	Many participants did not respond with an opinion regarding the subject and situation at hand, and instead just reworded the question as a statement to question 1 (e.g. "What has happened is that she cannot find her passport.") Additionally, a few (19%) of speakers said that they did not have an opinion what the subject should do given that they did not know what has happened in the scenario, rendering this scenario less clear than the others.

4) Use of OOPT vs. Lex30: As indicated above, originally the OOPT (Oxford Online Placement Test) was chosen to test EFL proficiency of EFL US and EFL CHI groups. However, 40% of the participants commented that the test, at 50 minutes, was too long, and did not complete the OOPT online before the scheduled Skype interview for the production task; furthermore, given their voluntary participation in this study, these participants never ended up completing the OOPT after the production task, either. Given this large attrition, no EFL proficiency measure was included in the analysis of the pilot data. Consequently, the OOPT was replaced with the Lex30 for the main phase, which at approximately 15 minutes, all participants of EFL US and EFL CHI groups successfully completed in the presence of the researcher prior to administering the production task.

Note that due to the small sample size, the following pilot results must be interpreted with extreme caution. Additionally, given N = 4 for each group, in-depth sub-analyses could not be conducted with validity and were thus not analyzed for the pilot. Nevertheless, the purpose of the Phase 2 pilot was not to garner parametrically sound results but more to verify and hone the instrumentation of the production task and procedural pragmatism, which it successfully tested.

In summary, results from ANOVA testing showed no mean difference for utterances per prompt for both Self-opinion and opinion of other prompt sets, thereby accounting for an otherwise potentially confounding variable. Table 4.16 summarizes results from the three ANOVAs run on both sets.

Table 4.16: ANOVA results from Phase 2 Pilot testing

Analysis	Prompt set	Trend	Significant effects
ANOVA 1: % Pragmatic function of [I + verb] to total utterances produced	Part 1: Self-opinion	Perceived confidence increases for D and U and decreases for R as US linguistic/cultural influence increases.	Significant between-subject effect of sociolinguistic background; interactional effect between pragmatic function and sociolinguistic background.
	Part 2: opinion of others	Perceived confidence increases for D and U as US linguistic/cultural influence increases; no particular trend for R.	Significant within-subject effect of pragmatic function; interactional effect between pragmatic function and sociolinguistic background.
ANOVA 2: % Pragmatic function of [I + verb] construct to total constructs used	Part 1: Self-opinion	Perceived confidence decreases for D and U and increases for R as US linguistic/cultural influence increases.	Significant interactional effect between pragmatic function and sociolinguistic background.
	Part 2: opinion of others	Perceived confidence increases for D and U as US linguistic/cultural influence increases; no particular trend for R.	Significant within-subject effect of pragmatic function.
ANOVA 3: Calculated CR of speaker confidence	Part 1: Self-opinion	CR trends from high-to-low from group 1 to 3 to 2 to 4.	Significant between-subject effect of sociolinguistic background.
	Part 2: opinion of others	CR trends from high-to-low from group 1 to 2 to 3 to 4; increase in CR corresponds to increase in US sociolinguistic background.	No between-subject effect of sociolinguistic background.

Whilst overall frequency trends did surface as predicted, many of these trends did not show significant mean differences between groups. This may be due to such a small sample size of n = 4 per group, wherein any outliers may have skewed significance of the ANOVA results. However, prompts overall did successfully elicit various pragmatic usages of [I + verb], with all prosodic-syntactic variations of [I + verb] produced within each group. Thus, pilot testing succeeded in substantiating the effectiveness of the production task, instilling hope that main phase testing will produce significant results.

4.4. Limitations of studies

Potential limitations encountered regarding this research, and how they are addressed in this study, are listed below.

Teasing apart language and culture: Given that language, thought, and culture are intrinsically intertwined, it becomes increasingly difficult to tease apart influences of culture vs. language. Recall how the definition for cultural environment includes how language is used for communication--in fact, many definitions of "culture" include "language" underneath its umbrella. However, given that a language carries many socio-contextual nuances and connotations, culture appears inherent to language and vice versa. Thus culture might more determine how language acquisition occurs, in turn shaping socio-cognitive interpretation and pragmatically influencing cognition.

In order to accommodate for the fluidity between these concepts, the term "sociolinguistic background" was adopted (see Definition of terms), thereby referring to language and culture as an intertwined, ever-evolving experience. With this perspective, the effect of both language and culture on communication may be investigated without the confound of differentiating between the (potentially overlapping) effects of each.

Homogeneity of culture: Whilst neuro-cultural and anthropological research has sought to pinpoint and define distinct cultural regions, there remains the fact that not every person within a socio-cultural community possesses the same neuro-cultural thought patterns and behaviours. In order to account for individual cognitive variation within a sociolinguistic environment, this study incorporates a sample size near the calculated power estimate that aims to draw from diverse parts of the cultural whole. Inclusion of participants from a range of age, geographical, and gender backgrounds should help to garner a more accurate depiction of the sociolinguistic environment overall and avoid misrepresentation of any niche cognitive attitudes as general thought, thereby controlling for these factors to maintain internal validity (Paltridge & Phakiti, 2015:14).

Effect of age: Age may play a role in L2 speaker performance in multiple facets, such as age of L2 acquisition and current age of participant (Paltridge & Phakiti, 391:2015). Whilst most research agrees that these variables of age are influential in L2 acquisition, there is no consensus yet as of how the exact nature of L2 acquisition varies as a function of these variables (Dornyei, 2009). Nevertheless, to mitigate against any influence from

such variables, both age and age of L2 acquisition were measured and included as covariates of analysis where appropriate.

Variation of individual speaker style: Inevitably speaker variation will be present between all individuals, even between speakers of the same language. Moreover, speaker style may differ within the same speaker at different points in time, and such differences may include inconsistencies with prosodic marking (Niemann et al., 1998). Nonetheless, results regarding a universal pragmatic-prosodic mapping from Phase 1 may lend support to a consistent focus-marking of [I + verb] pragmatic-prosodic implicatures in Phase 2; furthermore, given the high frequency of [I +verb] usage in expressing an opinion, it is anticipated that participants are familiar in their use of these constructs, and thereby variability should be limited.

Variation in environmental interaction: Assuming that overall cultural environment can be largely predicted by country, how people choose to interact in their external environment may still vary. For instance, a Chinese immigrant may prefer the company of other Mandarin speakers and may thus speak Mandarin with peers, despite group identification with an English-speaking environment. Furthermore, cultural and linguistic environment *within* the household may be difficult to discern, particularly with mixed-sociolinguistic background participants (EFL US group). Thus, these variables of Preferred language of socialization and Language spoken in the home were measured through the language questionnaire of Phase 2 and included when relevant for analysis (see section 4.3.3).

Nonetheless, whilst qualitative surveys do help shed light on L1 vs. L2 usage and influence in various environments, these qualitative self-descriptions may differ not only from reality, but the same descriptor (e.g. "predominantly Chinese household") may be used by two individuals to describe different environments. To combat such confounding variation, the language questionnaires developed for this study originally contained carefully tailored *quantitative* questions with an optional column for *qualitative* explanation. However, due to sample sizes of N = 25 per group, a mixed-design such as this would not have yielded sufficient observed power for inferential analyses. Thus, a nominal measurement was used to ensure higher validity at the expense of a less accurate quantification of meaning.

Teaching style within formal education environments: Teaching style encompasses both variables of Language of L2 instruction and Native language of L2 instructor. For example, lexical items may be defined through direct L1 translation, L2 paraphrasing, or through placing the term in a real-world context in either the L1 or L2. Whether teaching style variation presents as confounding may be difficult to determine, but specified language questionnaires (Appendix 3) aim to account for these factors by quantifying whether participants' mode of English learning instruction was from only native English speakers, only Mandarin speakers, or both; the decision to include this variable in analysis is based on Medgyes' (1992/1994) work regarding behavioral differences in Native vs. Non-native foreign language teachers.

While Non-native instructors may not necessarily be less efficient at English language teaching (indeed, Medgyes (1992:346) states that non-native competency may

yield alternative benefits in teaching strategy), there may still be differences in teaching practice due to divergent language backgrounds and potential differences in teacher cognition between native and non-native instructors (Aslan, 2015; Medgyes, 1992). Additionally, results were analyzed based on whether participant's L2 instruction was done in an English or Chinese school. These variables, along with other external variables, were measured and included in analyses to further strengthen internal validity.

Participant dropout rate: Because no incentive was offered and voluntary participation may be terminated at any time, participants may be inclined to stop undertaking the tasks before completion. Phase 1 is shorter than Phase 2 in length, but because it is an online task done outside the presence of the researcher, participants may be more likely to quit the task than if it was an in-person test. This was indeed found for the pilot phase, wherein many participants did not even finish half the survey before stopping. This attrition may influence the results due to exclusion of unfinished data; however, assuming this exclusion is done completely at random and not systematically, there should be no bias due to such attrition.

Because Phase 2 is an in-person task, participants may feel more motivated to finish the task when in the researcher's presence. Indeed, only three people failed to complete the task in the main data collection phase. Additionally, the Lex30 was administered prior to undertaking the oral production task in order to make the participant feel more inclined to partake in the production task, given that he or she had already invested time in the vocabulary test prior to the proficiency test.

Naturalness of interaction: Because participants had not previously made acquaintance with the researcher, they may have felt apprehensive or self-conscious in conversing during the oral survey. Nonetheless, the production task aimed to recreate natural spontaneous discourse by choosing in-person conduct and limiting prompt repetition. Furthermore, participants may adjust how they speak based on speaker perception of the researcher (e.g. gender, age, ethnicity). However, because the same researcher was involved in every interaction and prompts of the researcher's voice were recorded/played, this perception should remain constant across all participants; hence, since results are compared between-groups, this variable should not affect findings.

4.5. Ethical considerations

As this is not an intervention study, minimal ethical risk is presented. A CUREC 1/A ethics and signed consent informed individuals of potential risks. To avoid misuse of participant information, data was kept confidential in accordance with the Data Protection Act of 1998; additionally, the conducting researcher has undergone ethics training[42]. In an effort to ensure that participants are protected from any personal or social ramifications concerning their responses, data was treated confidentially and anonymized via participant ID codes assigned prior to analysis and write-up (Paltridge & Phakiti, 2015:40). To mitigate against involuntary participation, individuals received information stating that participation is voluntary, and may end whenever desired. Thus, participants should be exposed to minimal ethical risk.

[42] Investigating Research (Ethics course), UCL; 05/2013; Oxford Research Skills and Toolkit Online; 04/2014.

5. PHASE 1 RESULTS

Results for Phase 1 are presented regarding examination of response timing, perception task analysis, and coding of speaker confidence definition. Significance was set at the .05 value (Berry, 1997:260).

5.1. Examination of timing

Timing of response submission was analyzed to check for improper completion of the perception task (i.e. submitting a response without listening to the full recording). For participant 32, answers were submitted before the minimum time required to listen to the audio clip. Audio clips ranged from two-to-three seconds, but 16/24 of participant 32's answers were submitted within three seconds, 7 of which were less than two seconds, thereby indicating improper performance of the task. Thus, final sample size for analysis was N = 30 for US speakers, N = 29 for Mandarin speakers.

5.2. Perception task

Analysis included three mixed-design repeated measures ANCOVAs (see section 4.2.4 for details). For each analysis, descriptive statistics are first presented, followed by inferential statistics.

Confidence Rating of [I + verb] prosodic variations:

The first ANCOVA served to investigate RQ 1 and 1a (refer to section 4.1.1). This analysis explored whether prosodic variation and pragmatic attitude of [I + verb] statements are correlated in terms of confidence rating (CR), and whether sociolinguistic background influences the CR of [I + verb] statements. This analysis was conducted with one within-subject factor (prosodic variation) tested at six levels (Direct, Pronoun, Verb, Front, Mid, End) and one between-subject factor (sociolinguistic background) tested at two levels (US, CHI), with age as covariate and confidence rating (CR) as dependent variable. Table 5.1 presents the mean CRs of prosodic variations by group.

Table 5.1: Means, SDs, and confidence intervals of prosodic variations

Kontrast Placement	Socio-linguistic background	Mean	SD	95% Confidence internal	
				Lower bound	Upper bound
Direct (DIR)	US	6.16	.73	5.86	6.46
	CHI	6.19	.87	5.88	6.50
	Total	6.17	.80	5.96	6.39
Pronoun (PRN)	US	5.94	.87	5.61	6.27
	CHI	5.22	.89	4.90	5.56
	Total	5.59	.96	5.54	5.82
Verb (V)	US	4.47	.82	4.16	4.78
	CHI	4.89	.83	4.57	5.20
	Total	4.67	.86	4.46	4.89
No Kontrast: Sentence initial (FRONT)	US	3.61	1.12	3.14	4.08
	CHI	4.25	1.38	3.77	4.72
	Total	3.93	1.30	3.60	4.26

No Kontrast: Within-sentence (MID)	US	3.20	1.07	2.80	3.62
	CHI	3.37	1.10	2.96	3.78
	Total	3.29	1.09	3.00	3.57
No Kontrast: Sentence final (END)	US	3.41	.82	3.08	3.75
	CHI	3.50	.96	3.15	3.84
	Total	3.45	.89	3.22	3.69

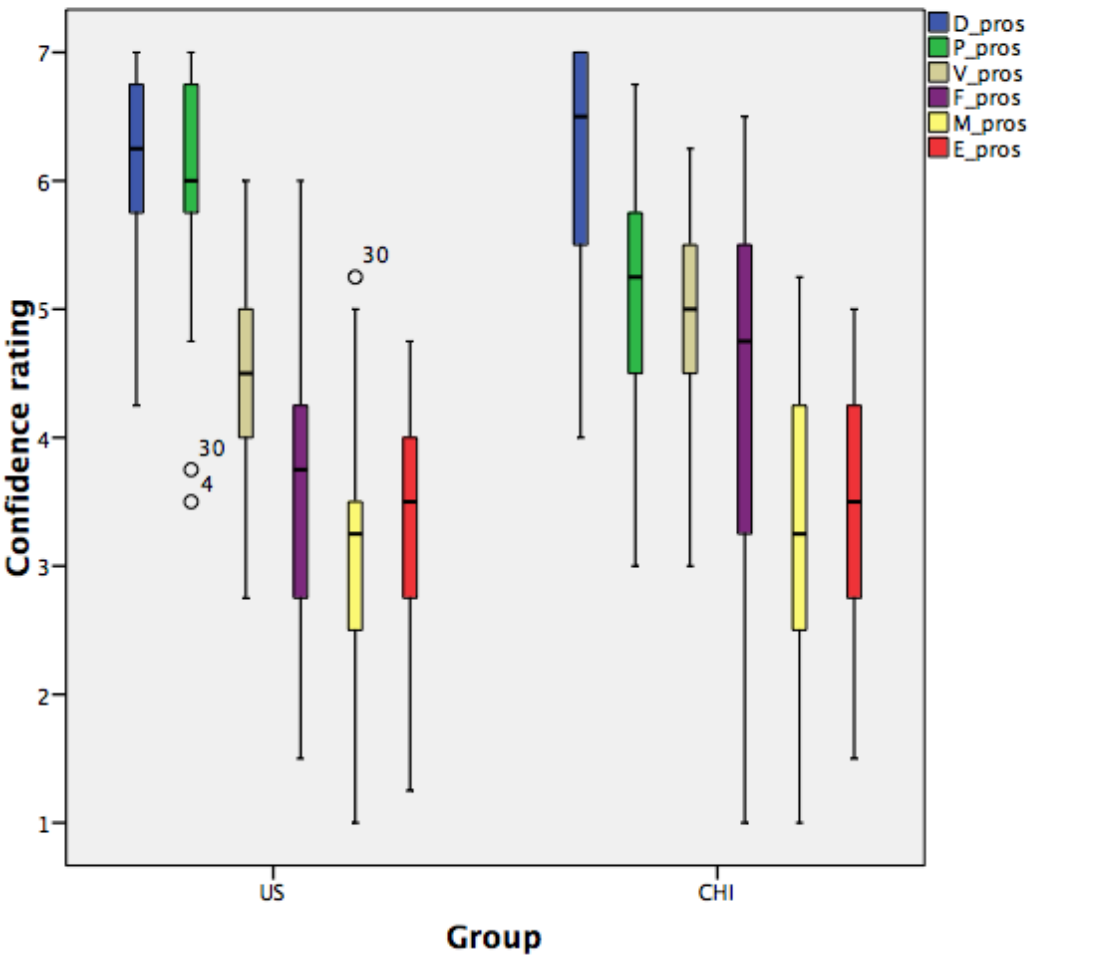
Shapiro-Wilk results indicated normality for all conditions between groups except DIR level for US ($p = <.01$) and CHI ($p = .001$) and PRN level for US ($p < .01$). Explorative analysis showed normal distribution overall, with some negative skew in the DIR level for US (skewness = -1.00, std. error = .43) and CHI (skewness = -1.10, std. error = .43) and PRN level only for US (skewness = -1.39, std. error = .43). However, levels for 95% confidence intervals (Table 5.1) show that there is no ceiling effect given a continuous scale up to 7. Thus, because the majority of data is distributed normally and a small portion merely has negative skew, normality is not violated and parametric tests will be run.

Whilst within-level and between-group boxplots showed no outliers in CHI and two cases with potential outliers in US (4, 30) (Figure 5.1), Z scores (Participant 30:Mid level SD = 1.81, Participant 4:Pronoun level SD = -2.17, Participant 30:Pronoun level SD = -1.91) showed that no data points lied more than the 3 standard deviations from the mean (based on Howell's (1998) practiced notion that 99.87% of a normal distribution falls within this range), thereby observing normality of distribution[43]. Moreover, ANCOVA

[43] Field (2003:217,698) defines potential outliers as being between 1.96 – 2.58 SDs away from the mean, with probable outliers being between 2.58 – 3.29 SDs away, and deeming anything 3.29 SDs or above

results carried out both with and without these data points do not differ in terms of statistical significance. In addition, due to small N of sample size removing these two cases only created more potential outliers. Thus, statistics below are run with all cases included, with a version of ANCOVA run with potential outliers removed in Appendix 13 for reference.

Figure 5.1: Boxplots for Prosodic variation between groups



away from the mean as an extreme outlier. This appears to correspond with Howell's (1998) 3 SDs notion, and thereby strengthening the decision to use 3 SDs as a determiner for a data point being treated as an outlier for removal.

Whilst sphericity cannot necessarily be assumed (X^2 (1,14) = 25.43, p = .03), results calculated with Mauchly's sphericity vs. Greenhouse-Geiser (.85) or Huynh-Feldt measurements (.96) did not differ regarding significance. Since the Greenhouse-Geiser measurement did not fall below .75, the Hunyh-Feldt estimate was chosen as a correction for sphericity given that Greenhouse-Geiser measurements may be too conservative (Hyunh & Feldt, 1976, Field, 2013:1778). Positive results for near-sphericity are furthermore signified by the close proximity of the Hunyh-Feldt measurement to 1 vs. the lower-bound estimate (1/([# conditions] – 1) = .2) (Field, 2013:1777).
Levene's test (see Table 5.2) indicated homogeneity of variances between all levels of prosodic variation.

Table 5.2: Levene's test of Equality of Error Variances

Prosodic variation	F value	Degrees of freedom	Significance
Direct	1.13	1, 57	.29
Pronoun	.03		.86
Verb	.01		.94
Front (no kontrast)	2.14		.15
Mid (no kontrast)	1.25		.27
End (no kontrast)	.77		.38

Results showed a significant within-subject effect of prosodic variation (F(1,4.81) = 24.74, η_p^2 = .31, p < .001). Conversely, no between-subject effect of sociolinguistic

background ($F = .32$, $\eta_p^2 = .01$, $p = .58$) or covariate of age ($F = .71$, $\eta_p^2 = .01$, $p = .40$) appeared. Additionally, age did not interact significantly with prosodic variation ($p = .107$). There was, however, a moderate interaction between sociolinguistic background and prosodic variation ($F(1,4.81) = 4.76$, $\eta_p^2 = .08$, $p < .001$).

The large effect of prosodic variation may be especially telling given a small sample size[44]. Referring back to Table 5.1, mean CRs are shown to significantly decrease from DIR to PRN to V to FRONT to no kontrast MID/END, thereby supporting predictions. To examine this effect, post-hoc paired t-tests were run with a Bonferroni adjustment for multiple comparisons[45] (see Table 5.3); whilst this "simple but effective correction" implies a more conservative type 1 error rate and thereby a higher probability for rejecting an effect that exists, it also strengthens the validity of the significant findings below (Field, 2013:73, 2723). Significant mean differences were marked with an asterisk (key at bottom of table) and emboldened for reference.

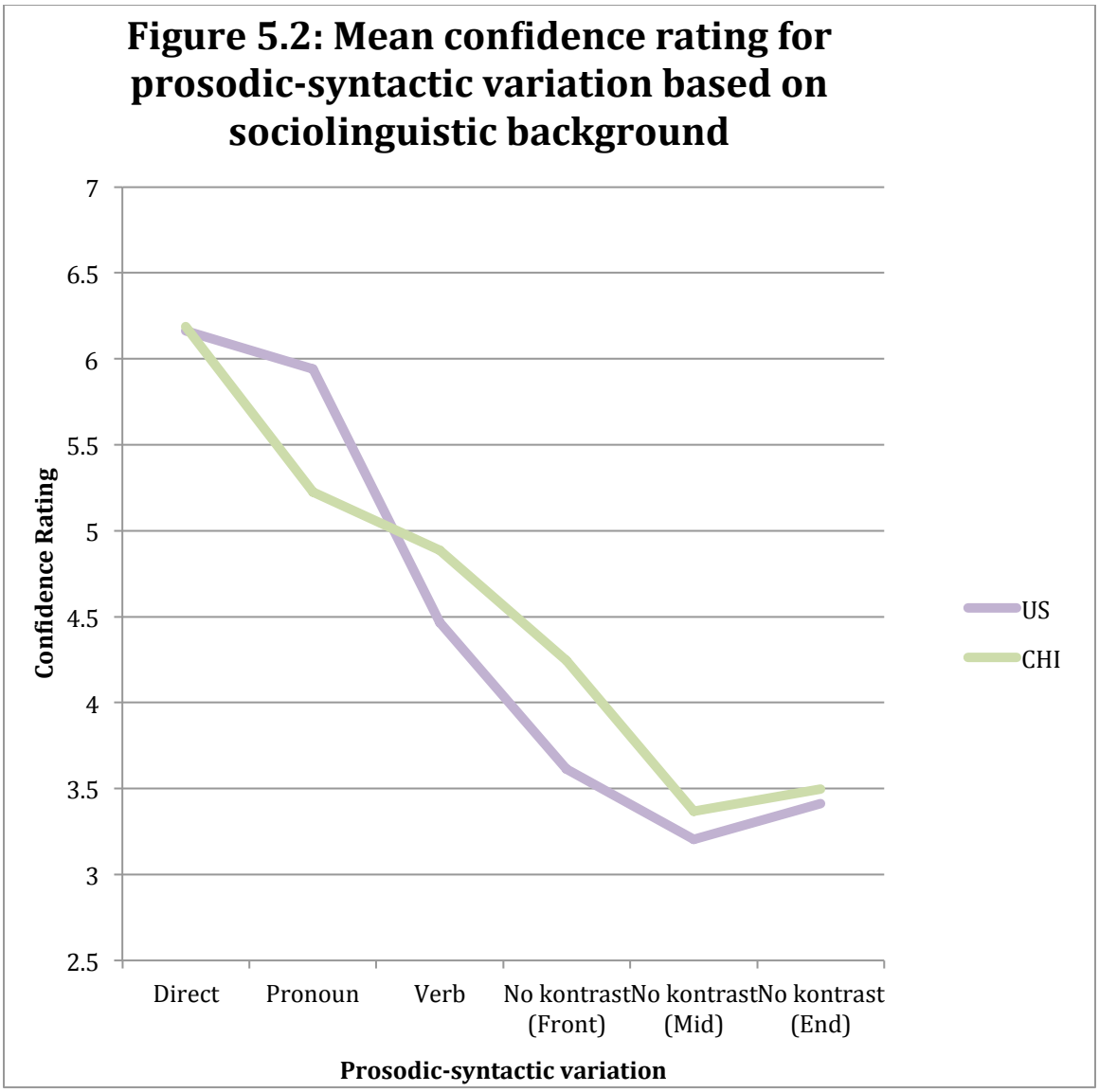
[44] "According to Cohen and Miles & Shevlin (2001) measurement guidelines for effect sizes are as follows: Small = .01, Medium = .06, Large = .14; these guidelines will thereby be adopted throughout this study.
[45] I.e. for post-hoc testing, SPSS calculations multiplied each of the significance levels from the unadjusted t-test by the number of tests performed. Note that this does not refer to Bonferroni correction in the sense of dividing alpha by the number of groups; rather, with this calculated post-hoc adjustment to results the alpha level is still .05. All further post-hoc paired t-tests are Bonferroni-adjusted unless otherwise stated.

Table 5.3: Post-hoc paired t-tests with Bonferroni adjustment

Prosodic variation	Compared with	Mean Difference	95% Confidence Intervals	
			Lower Bound	Upper Bound
DIR	PRN	**.59****	.15	1.03
	V	**1.50***	1.12	1.88
	F	**2.25***	1.74	2.75
	M	**2.89***	2.43	3.35
	E	**2.72***	2.35	3.09
PRN	DIR	**-.59****	-1.03	-.15
	V	**.91***	.48	1.34
	F	**1.66***	1.12	2.20
	M	**2.30***	1.81	2.80
	E	**2.13***	1.74	2.53
V	DIR	**-1.50***	-1.88	-1.12
	PRN	**-.91***	-1.34	-.48
	F	**.75***	.34	1.15
	M	**1.39***	1.00	1.78
	E	**1.22***	.88	1.57
F	DIR	**-2.25***	-2.75	-1.74
	PRN	**-1.66***	-2.20	-1.12
	V	**-.75***	-1.15	-.34
	M	**.64***	.14	1.15
	E	.48***	.07	.88
M	DIR	**-2.89***	-3.35	-2.43
	PRN	**-2.30***	-2.80	-1.81
	V	**-1.39***	-1.78	-1.00
	F	**-.64***	-1.15	-.14
	E	-.17	-.55	.22
E	DIR	**-2.72***	-3.09	-2.35
	PRN	**-2.13***	-2.53	-1.74
	V	**-1.22***	-1.57	-.88
	F	**-.48****	-.88	-.07
	M	.17	-.22	.55

* = p < .05; ** = p < .01' *** = p <.001.

Table 5.3 illustrates significant differences between all levels except MID/END ($p = 1.00$), which differ by only .17 out of the 1-7 rating scale. Recall that the predicted hypothesis for RQ 1 predicts that CR differs only for kontrast variation; thus, given that MID and END are both no kontrast conditions which differ only in syntactic placement, this result does not refute the hypothesis. However, FRONT, which also is a no kontrast condition, was significantly higher in confidence rating than MID and END, although only by .64 for MID and .48 for END. This difference was not significant in the pilot phase, and in relation to the 1-7 rating scale, this appears as a relatively small mean difference. Given the presence of aural variation when making each recording for the prosodic variation prompts, human error and vocal variability may account for this slight difference. Nonetheless, the overall trend of CR (high-to-low) appears similar between US and CHI groups (Figure 5.2), with CR trending from high-to-low as levels significantly shift from Direct to Pronoun, to verb, to no kontrast conditions.



Because there was no main effect of sociolinguistic background, results support both the hypothesis for RQ 2 and the predicted cross-linguistic congruence of pragmatic-prosodic mappings[46]. This renders influence from L1 pragmatic knowledge transfer (RQ

[46] Note that in Hypothesis 2), the effect of sociolinguistic background only relates to Proof of Concept (i.e. universality of pragmatic-prosodic mapping), and does not relate to trends regarding frequency of usage that may be more linguistically-specific to tonal vs. accent language characteristics.

5) irrelevant, for both CHI and US speakers should possess congruent pragmatic knowledge regarding prosodic focus-marking's effects on perceived speaker confidence.

To investigate the interaction between prosodic variation and sociolinguistic background, post-hoc paired t-tests were conducted on means adjusted for covariate of age on PRN, FRONT, and MID conditions between-groups where any gap in confidence rating may have occurred. PRN level results show equal variances within each group ($F(1,57) = .70$, $p = .53$), but between US and CHI a significant mean difference ($p < .01$) of .71 appeared on a scale of 7. An independent sample t-test further suggests a significant between-group difference at PRN condition ($t(1,57) = 3.54$, $p = .001$). FRONT condition results showed no significant mean difference between groups ($p = .07$). Additionally, no significant difference between-groups occurred at MID level ($p = .58$).

These findings suggest that results do not contradict hypotheses of a universal pragmatic-prosodic mapping between Mandarin and American English: it appears that kontrast placement, and the alternative semantic connotations plus modality it invokes, may universally determine perceived confidence of [I + verb] statements. Accordingly, the mean CRs of these prosodic variations (refer to Table 5.1) will be used in Phase 2 in which frequency of different [I + verb] prosodic variations are analyzed.

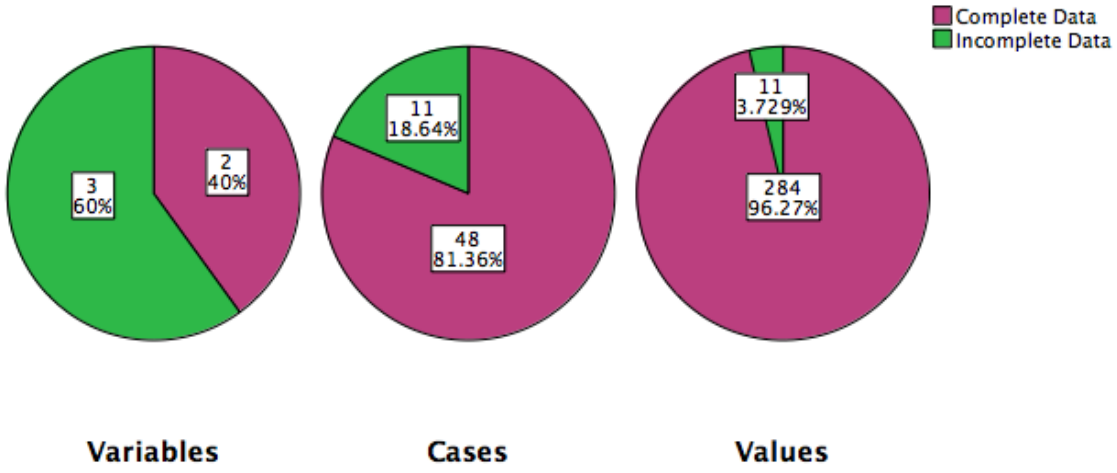
Confidence rating of [I + verb] perceived pragmatic function:

The second ANCOVA investigated RQ 1 and 1a (refer to section 4.1.1), to explore whether pragmatic attitude and perceived pragmatic function of [I + verb] statements are correlated in terms of confidence rating (CR), and whether sociolinguistic background

affects trend of CR of [I + verb] statements. This analysis was conducted with one within-subject factor (perceived pragmatic function) tested at three levels (Deliberate, Reluctant, Uncertain) and one between-subject factor (Sociolinguistic background) tested at two levels (US, CHI).

In the pilot phase, only 2/38 cases included missing values, and thus listwise deletion was used as a solution. For missing values in this data set (see Figure 5.3), although only 6% of data was incomplete, 11/59 cases (19%) included missing values. Analysis of missing data revealed that 9/11 missing values were in the R level, with 1/11 in each of the D and U levels. Thus, given the predictor variable of R for missing values, data appears to be missing at random (MAR) vs. missing completely at random (MCAR), meaning that listwise deletion may lead to invalid inferences (Schafer, 1999:7). In addition, listwise deletion would render too much of the original data obsolete. Therefore, it is necessary to use another method to replace missing values. Since simply replacing missing data with the mean would suppress the true SD and standard error (especially given a small sample size), multiple imputation (MI) was implemented for the repeated measures analysis to produce estimates "far better than the mean" (Field, 2013:1293; Rubin, 1987; Schafer, 1999). In addition to avoiding data loss, MI also makes less stringent assumptions about the nature of missing data and thereby is used well with MAR data (van Ginkel & Kroonenberg, 2014:79).

Figure 5.3: Overall summary of missing values



In terms of the number of MI to run, some theorists say Rubin's (1987) original estimate of only 3-5 suffices (with less than 5% loss of power between M = 5 and M = ∞) (Schafer & Olsen, 1998), some say there is negligible benefit for using more than 5-10 (Schafer, 1999), and others who say up to 100 is the safest option (Graham, Olchowski, & Gilreath, 2007). A more modern approach appears to tailor the number of imputations to the amount of missing data: White, Royston and Wood (2011) state that the number of imputations should be equal to at least the percentage of missing data cases. Thus, given 19% missing cases, this analysis used M = 20 with a random number generator set at 2,000,000 for the Mersenne Twister. For theorists who believe that M = 100 is safest to avoid loss of power, these results are also reported in Appendix 14.

Given that SPSS is unable to pool MI results for exploratory analysis or normality testing, van Ginkel (personal correspondence, 2016) suggests that the only method for detecting variation from normality is via observing the histograms. All 60 imputed histograms (20 per level) appeared to be relatively bell-shaped with no ceiling or floor

effects, albeit in some cases slight negative skew was shown in the US Uncertain level. In the interest of analysis of normality prior to MI, both listwise and pairwise deletion resulted in the same significance for Shapiro-Wilk results. Since excluding pairwise may cause bias in estimates and possibly alter effect sizes (Field, 2013:2532), exploratory results are reported using listwise deletion. However, it must be noted that due to numerous cases being excluded (18.64%) and such a small remaining sample size (N = 48), estimates may be skewed.

Shapiro-wilk results show normal data at all levels between groups except for US Deliberate level ($p = .013$) and Uncertain level ($p = .001$). Interestingly, pairwise deletion results as well as M = 1 to 20 results for US Deliberate level were all nearly normally distributed ($p < .05$). Imputed histograms for both these conditions appear with some negative skew (US Deliberate level skewness = -1.143, US Uncertain level skewness = -2.039), but 95% confidence intervals for both US Deliberate level (5.83, 6.34) and US Uncertain level (3.96, 4.50) indicate no ceiling effect present on a scale up to 7. It must be noted that skewness was significantly reduced using pairwise deletion (US Deliberate level skewness = -.911, US Uncertain level skewness = -1.475), indicating that without MI pooled results the exact amount of negative skew cannot accurately be estimated and may potentially be less. Nonetheless, neither analysis shows any ceiling effect. Thus, because the majority of data is distributed normally with small portion merely showing negative skew with no ceiling effect, normality is not violated and parametric tests will be run.

Next, van Ginkel's (2014) macro was run to implement MI for repeated measures in SPSS. Mauchly's test indicated that sphericity may be assumed ($\chi^2(2) = 1.163$, $p = .56$).

Being the first SPSS macro of its kind, it is not yet able to correct SPSS' inability to produce pooled results for Levene's test. Levene's test for original data using listwise deletion showed homogeneity of variance in both Deliberate ($F(1,46) = 2.270$, $p = .139$) and Reluctant ($F(1,46) = 2.608$, $p = .113$) levels but not Uncertain ($F(1,46) = 8.192$, $p = .005$) levels. However, given that minute differences in group variances can produce significant results in Levene's test, and that the test is not ideal for small sample sizes with uneven groups (Field, 2013:795), this test may prove too conservative; thus, in this case with heavy listwise deletion, a significant result shall be read cautiously and does not necessarily negate equality of variance.

Means of CR for each perceived pragmatic function are presented in Table 5.4[47]. Results reveal the same trend of CR between all levels for US and CHI groups. CR means are highest for Deliberate, then Uncertain, and lowest for Reluctant.

Table 5.4: Pooled mean confidence ratings and standard deviations for perceived pragmatic function based on sociolinguistic background

Perceived pragmatic function	Sociolinguistic background	Mean	Std. Error
Deliberate (D)	US	6.08	.11
	CHI	6.09	.12
	Total	6.08	.08
Uncertain (U)	US	4.25	.13
	CHI	4.27	.18
	Total	4.27	.11
Reluctant (R)	US	2.64	.15
	CHI	3.17	.19
	Total	2.90	.13

[47] SD and confidence intervals cannot be calculated due to pooled results. Standard error is given instead.

Pooled F-test results revealed a significant large effect of perceived pragmatic function ($F(2,106.46) = 307.58$, $\eta_p^2 = .85$, $p < .001$). For between-subject effects, sociolinguistic background was not found to be significant ($F(1,50.62) = .83$, $p = .38$). Covariate of age was also found to be non-significant[48]. No interactional effect was found between pragmatic function and sociolinguistic background ($F(2,105.09) = 2.76$, $p = .07$). To investigate the main effect of perceived pragmatic function, post-hoc paired t-tests were conducted (Table 5.5). As predicted, results revealed significant differences between all levels of perceived pragmatic function ($p < .001$), thereby rejecting the null hypothesis and supporting Hypothesis 1 for RQ 1.

Table 5.5: Post-hoc paired T-tests for pooled MI results

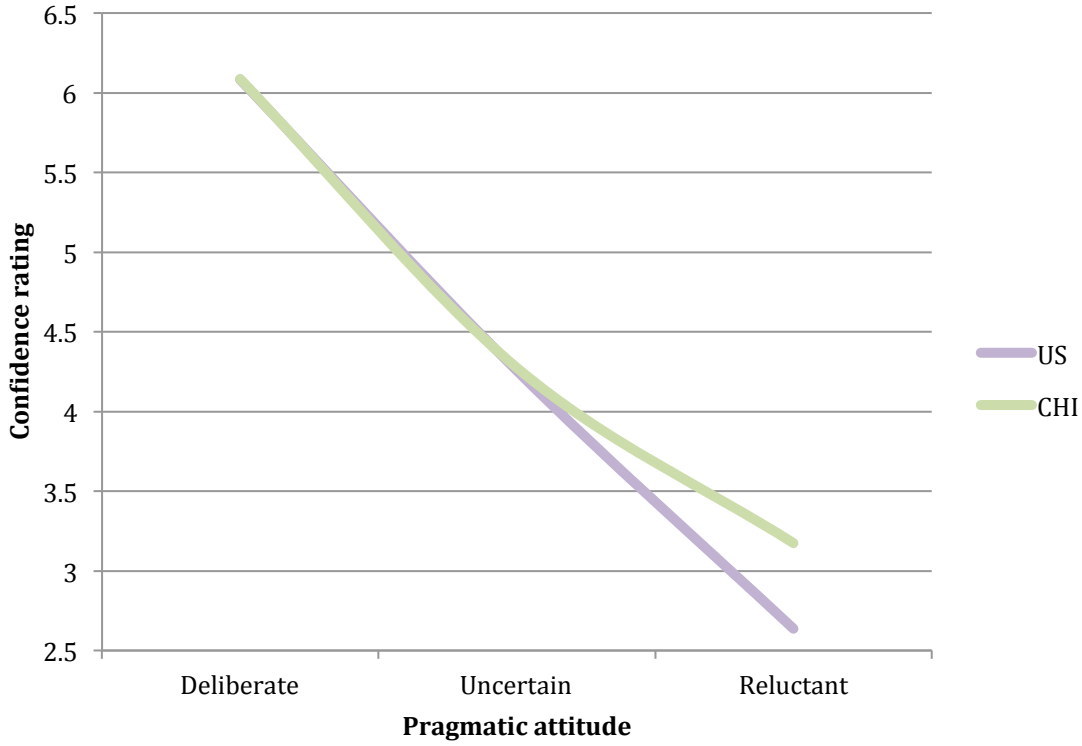
Pair	Mean difference	Significance	95% confidence intervals	
			Lower bound	Upper bound
D – U	1.82	< .001	1.46	2.16
U – R	1.36		.76	1.44
R – D	-3.18		-3.32	-2.50

The universally parallel movement of CR between-groups is visible in Figure 5.4. As predicted in Hypothesis 1, a visible difference of means appeared between each level trending from Deliberate, down to Uncertain, down to Reluctant. For Deliberate and Uncertain levels, US and CHI mean CRs are nearly identical. While US participants rated

[48] Due to the nature of the MI macro, covariates cannot be inserted; age was therefore run as a between-subject variable and no interactional effect was investigated with age.

Reluctant level as less confident than did CHI participants, this difference was not found to be significant. Both groups thereby exhibited the same trend of confidence rating between levels, thereby supporting this thesis' hypothesis for RQ 1 and indicating a universal perception of high confidence for deliberate statements, medium confidence for uncertain statements, and low confidence for reluctant statements.

Figure 5.4: Mean confidence rating for perceived pragmatic function based on sociolinguistic background



The lack of a significant effect of sociolinguistic background mirrored results of ANCOVA 1, thus also supporting the predicted hypothesis for RQ 1a. Similarly, the lack

of a significant interactional effect between sociolinguistic background and perceived pragmatic function indicated no significant mean difference in CR at any levels between US and CHI groups, which furthermore supports the predicted hypothesis for RQ 1a.

Relationship of CR between prosodic variation and pragmatic attitude:

A third analysis was conducted to establish a CR correlation between prosodic variation and pragmatic attitude, thereby strengthening the existence of a pragmatic-prosodic link and supporting hypotheses for RQ 1 and 1a (refer to section 4.1.1). It is predicted that when calculated via CR of perceived pragmatic attitude, CR of [I + verb] statements will follow the same prosodic variation trend as the first ANCOVA, and this trend will not be affected by sociolinguistic background.

Each pragmatic type was assigned a value based on mean CR calculated in ANOVA 2 (see Table 4.4 in section 4.2.4). Next, each participant's average CR per prosodic variation was tallied via average CR of pragmatic attitude chosen for that contour. This analysis was run with the same within-subject and between-subject factors as the first ANOVA, only with newly calculated CR values for prosodic variation based on pragmatic attitude CRs. Additionally, a Spearman's rank-order correlation test was conducted to further establish correlation.

Levene's test indicated homogeneity of variance at all conditions except V and was nearly present for P (see Table 5.6).

Table 5.6: Levene's test of Equality of Error Variances

Prosodic variation	F value	Degrees of freedom	Significance
Direct	.18	1, 57	.68
Pronoun	4.23		.04
Verb	12.82		<.01
Front (no kontrast)	1.02		.32
Mid (no kontrast)	1.87		.18
End (no kontrast)	.60		.44

Whilst Mauchly's test indicated that sphericity may not be assumed (X^2 (1,14) = 35.64, $p < .001$), results calculated with Mauchly's sphericity vs. Greenhouse-Geiser (.789) do not differ regarding significance. Positive results for near-sphericity are furthermore signified by the close proximity of the Greenhouse-Geiser measurement to 1 vs. the lower-bound estimate (1/([# conditions] – 1) = .2) (Field, 2013:1778).

Shapiro-Wilk testing (see Table 5.7) suggested potential non-parametric distribution within 8 out of 12 levels between both groups. Whilst histogram examination revealed mere skew in some conditions (DIR, PRN for US group; MID for CHI group), the other five potential conditions did not appear to have normally distributed data.

Table 5.7: Shapiro-Wilk normality test for calculated CR

Prosodic variation	Sociolinguistic background	Statistic	Significance
Direct (DIR)	US	.40	<.001
	CHI	.36	<.001
Pronoun (PRN)	US	.61	<.001
	CHI	.93	.07
Verb (V)	US	.82	<.001
	CHI	.91	.02
No Kontrast: Sentence initial (FRONT)	US	.93	.05
	CHI	.96	.37
No Kontrast: Within-sentence (MID)	US	.88	<.01
	CHI	.87	<.01
No Kontrast: Sentence final (END)	US	.90	<.01
	CHI	.94	.09

This sort of distribution may be expected given that the CR for prosodic variation is calculated from the CR of ANOVA 2, which was rated according to pragmatic attitude instead of prosodic variation. Nonetheless, this analysis serves to seek whether a correlation between the two variables exists.

Given these factors, the Kruskal-Wallis test is included in Appendix 15. However, given the presence of overall homogeneity of variance, the robustness of the ANOVA test should be able to account for the above variables. Thus, ANOVA testing was still carried out. It can be promisingly noted that results did not differ in significance between parametric and non-parametric testing. Means of calculated CR of prosodic variations by

group are presented below in Table 5.8.

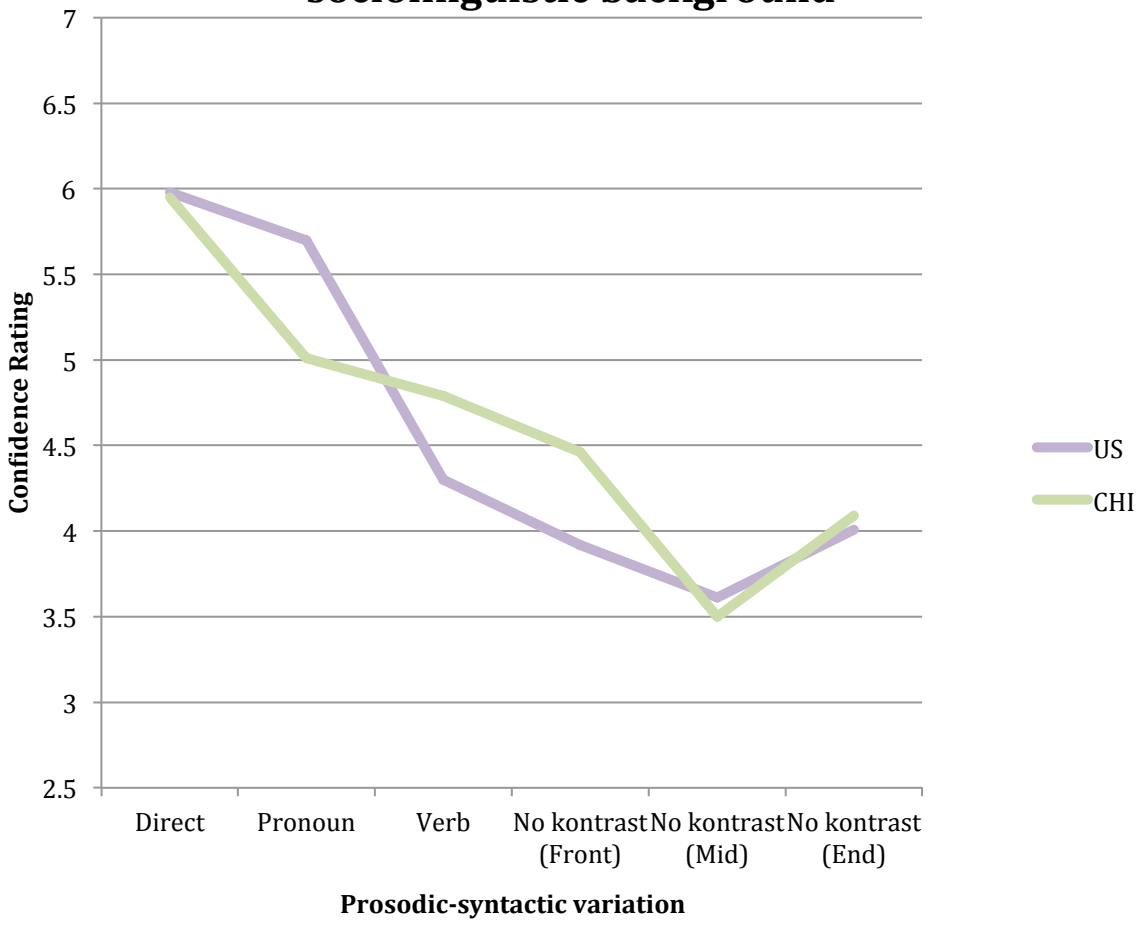
Table 5.8: SD, Std. Error and Means for calculated CRs of prosodic variations

Kontrast placement	Sociolinguistic background	Mean	SD	Std. Error
Direct (DIR)	US	5.98	.31	.07
	CHI	5.95	.39	.07
	Total	5.97	.35	.05
Pronoun (PRN)	US	5.70	.57	.13
	CHI	5.01	.81	.13
	Total	5.35	.81	.09
Verb (V)	US	4.30	.42	.12
	CHI	4.80	.78	.12
	Total	4.55	.67	.08
No Kontrast: Sentence initial (FRONT)	US	3.92	.72	.15
	CHI	4.46	.85	.15
	Total	4.19	.83	.10
No Kontrast: Within-sentence (MID)	US	3.61	.64	.11
	CHI	3.50	.50	.11
	Total	3.56	.57	.08
No Kontrast: Sentence final (END)	US	4.01	.73	.14
	CHI	4.09	.72	.14
	Total	4.05	.72	.94

Results appeared to exhibit the same trend of CR between prosodic variation as in the first analysis: CR shifts from high-to-low from DIR to PRN to V, with the no kontrast positions at the bottom. This trend can be seen in Figure 5.5. Note the similarity in shape

and trend here as in Figure 5.4 for ANOVA 1's CR of prosodic variation. This trend will be discussed more thoroughly after presenting ANOVA results below.

Figure 5.5: Mean calculated confidence rating for prosodic-syntactic variation based on sociolinguistic background



ANOVA results revealed a large within-subject effect of prosodic variation ($F(1,5) = 21.93$, $p < .001$, $\eta_p^2 = .28$). There was no statistically significant between-subject effect of sociolinguistic background ($F(1,1) = .21$, $p = .65$, $\eta_p^2 = < .01$) or covariate of age ($F(1,1) = .24$, $p = .62$, $\eta_p^2 = < .01$). Additionally, while age did not interact significantly with

prosodic variation ($F(1,5) = 1.01$, $p = .11$, $\eta_p^2 = < .03$), there was a significant interaction between group and prosodic variation($F(1,5) = 9.01$, $p < .001$, $\eta_p^2 = < .14$).

Post-hoc paired t-tests run with covariate of age-adjusted means suggested between-group differences in PRN, V, and FRONT conditions. Independent sample t-tests were conducted on PRN, V, and FRONT conditions to investigate this interaction. Results showed significant differences at PRN ($t(1,57) = 4.52$, $p < .001$), V ($t(1,57) = -3.31$, $p < .01$), and FRONT ($t(1,57) = -2.61$, $p = .01$) conditions. Results appear similar to the investigation of the interaction in Analysis 1, where PRN condition also showed a between-group difference, but differ in a significant between-group difference in FRONT and V conditions. Nonetheless, much like in Analysis 1, CR trends stay congruent from DIR down to MC, to E, then to the DMs for both groups, further providing evidence for this kind of universal pragmatic-prosodic mapping.

Post-hoc paired t-test results with calculated CR mirror that of ANOVA 1's prosodic variation CR results in terms of significance: All conditions showed significant differences ($p = .01$ between V and FRONT; $p < .001$ for all the rest) except between FRONT and END conditions ($p = 1.00$). It may be noted that Analysis 1 and 3 in the pilot showed the same post-hoc results in terms of significance. Analysis 1 for the main phase showed similar results, with the only difference being that no significant difference was found between the MID and END no kontrast conditions vs. the FRONT and END no kontrast conditions; however, given that the hypothesis only predicts a CR trend from DIR down to PRN to V to no kontrast conditions, what appears meaningful is that all three no kontrast conditions are still significantly lower in CR than the rest of the conditions. These results thereby further strengthen the CR relationship between

pragmatic attitude and prosodic variation by confirming significant differences between DIR, PRN, V, and no kontrast conditions as stated in the conceptual framework.

A Spearman's rank correlation test was conducted to further determine the correlation between prosodic variation and pragmatic attitude. Results showed significant one-tailed correlations at all conditions between pragmatic variation CR (Analysis 1) and pragmatic variation with pragmatic attitude-calculated CR (see Table 5.9).

Table 5.9: Spearman's correlation coefficient between calculated CR and CR of prosodic variation

Condition	r_s	Significance
Direct	.38	<.01
Pronoun	.69	<.001
Verb	.48	
No Kontrast: Front	.76	
No Kontrast: Mid	.52	
No Kontrast: End	.66	

In summary, when calculated via CR of perceived pragmatic attitude, CR of [I + verb] statements follows the same prosodic variation trend as Analysis 1. Thus, these results altogether indicate a pragmatic-prosodic link in the use of [I + verb] constructs in expressing an opinion. The particular pragmatic-prosodic mapping designed in the conceptual framework (see Section 3.2) is supported by evidence from this data, with CR trending from high-to-low as it moves from Deliberate to Uncertain to Reluctant on a pragmatic level, and from DIR to PRN to V to no kontrast forms on a prosodic level. Prosodic variation, even with pragmatic attitude-calculated CR, was once again found in Analysis 2 to have a significant main effect, just as the effect of sociolinguistic

background was again found to be insignificant. While some between-group differences seem possible within certain prosodic conditions, these differences do not influence or contradict the trend of the pragmatic-prosodic mapping overall for either group.

Thus, Analysis 3 serves to both help establish a pragmatic-prosodic link as well as to strengthen the existence of a universal trend of CR in syntactic-prosodic variations of [I + verb] constructs. Therefore, Phase 1 perception task results appear to support the predicted hypotheses for RQs 1 and 1a. The frequency of functional [I + verb] construct usage and how this may differ between speakers of US vs. CHI sociolinguistic backgrounds will be further investigated in Phase 2. The next section analyzes the results from the open-ended question regarding speaker confidence definition appearing at the end of Phase 1's perception task, which serves to investigate RQ 2.

5.3. Coding of speaker confidence definitions

Sociolinguistic background's affect on perception (RQ 2) was investigated through participants' definitions of "speaker confidence". Remarks were semantically annotated into five categories (Table 5.10). Descriptive statistics are first presented below, followed by inferential statistics.

Table 5.10: Frequency of speaker confidence definitions per semantic category

	Direct/firm tone of language	Belief of speaker toward statement	Belief of listener toward statement	Appeal/ Reputation of speaker	Comfortableness of speaker to share statement
US	16	19	1	1	6
CHI	4	6	11	16	0

Due to the open-ended nature of responses and the potential for multiple utterances within a single response (see section 4.2.4.2.), each participant's response may fall under more than one category. Twenty responses fell into two categories; none fell into three or more. Each group's mean percentage of response per category is presented in Table 5.11.

Table 5.11: Mean percentage of response inclusion based on semantic annotation

Semantic category	Sociolinguistic background	Mean percentage of response inclusion	SD
Direct/firm tone of language	US	.53	.51
	CHI	.14	.35
	Total	.34	.48
Speaker belief toward statement	US	.63	.49
	CHI	.21	.41
	Total	.42	.50
Listener belief toward statement	US	.03	.18
	CHI	.38	.49
	Total	.20	.41
Appeal/ Reputation of speaker	US	.03	.18
	CHI	.55	.51
	Total	.20	.41
Comfortableness of speaker to share statement	US	.20	.41
	CHI	.00	0
	Total	.10	.31

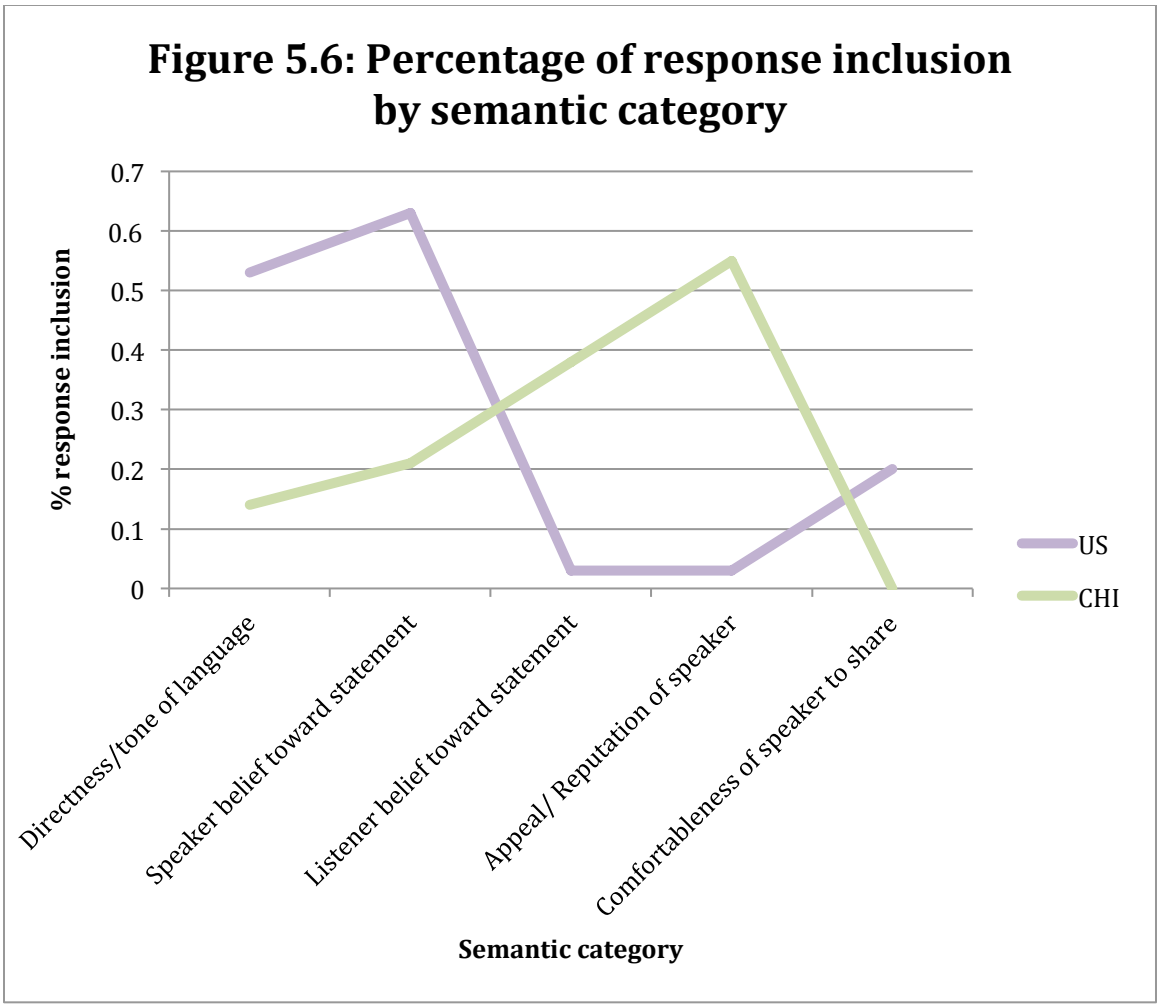
Due to an unequal observation of data from non-mutually exclusive semantic categories (i.e. each open-ended response may belong to multiple categories), a non-parametric Kruskal-Wallis test was conducted. Results (see Table 5.12) revealed

statistically significant differences between US and CHI groups at all levels. These differences are presented in Figure 5.6, which illustrates an inverse correlation of frequency between levels. Peak frequencies for US annotations (Speaker belief toward statement, Directness/tone of language) appear in the same category where CHI annotations exhibit lower frequencies, while peak frequencies for CHI annotations (Appeal/Reputation of speaker, Listener belief of speaker) appear in the categories with the bottommost frequencies for US annotations.

Specific between-group differences as well as within-group analyses are conducted by semantic category below.

Table 5.12: Kruskal-Wallis results for semantic categories of speaker confidence definition.

Semantic category	Chi-squared	Significance value
Direct/firm tone of language	10.11	.001
Speaker belief toward statement	10.80	.001
Listener belief toward statement	10.71	.001
Appeal/ Reputation of speaker	18.99	.000
Comfortableness of speaker to share statement	6.35	.01



Direct/firm tone of language: 53% of US participants' responses indicated that the definition of speaker confidence relates to direct/firm tone of language, which comprised 16 out of 43 (37.21%) total US semantic annotations. In comparison, this was referred to by only 14% of CHI participants' responses, which comprised a mere 4 out of 37 (10.81%) total CHI semantic annotations. This data indicates that US individuals are three times more likely to judge speaker confidence based on a speaker's direct or firm tone of language.

Speaker belief toward statement: The majority of US participants' responses fell under this category at 63%, which at a count of 19 also comprises the majority of total US semantic annotations at 44.19%. Only 21% of CHI participants' responses belonged to this category, making up 16.22% of total CHI semantic annotations. Thus, US participants appear three times more likely to link speaker belief toward statement to speaker confidence.

Listener belief toward statement: Only one US response mentioned listener belief, comprising 3% of participants and 2.33% of total annotations. In contrast, 11 CHI responses fell under this category, comprising 38% of participants and 29.73% of total CJI annotations. Chinese speakers therefore appear more likely to correlate listener belief toward statement with speaker confidence whist US individuals do not.

Appeal/Reputation of speaker: Like listener belief toward statement, appeal/reputation of speaker does not appear to be a prominent determinant of speaker confidence for US speakers at only one response (3% of participants; 2.33% of total annotations), making it tied with listener belief for the least cited semantic category for US participants. Conversely, this semantic category encompasses the largest amount of CHI responses with 55% of participants and 43.24% of total CHI annotations. CHI responses therefore dominate this semantic category.

Comfortableness of speaker to share statement: 20% of US speakers cited

comfortableness of speaker to share statement as a determinant of speaker confidence. Whilst this may comprise only 12.77% of total US annotations, this percentage is significant compared to the 0 CHI responses. This was the third most cited category for US speakers, whilst being the least cited category for CHI speakers.

In summary, there seemed a converse relation between US and CHI participants for each semantic category of speaker confidence definition; furthermore, all differences between levels are proven to be statistically significant via Kruskal-Wallis testing. Whilst US participants appear to focus on the language of speech itself and on speaker mindset (both speaker belief and comfortableness of speaker to share), CHI participants may focus instead on the listener, both in terms of listener belief and how appealing the speaker appears to the listener. These results appear to mirror the mindsets of IDV vs. COL sociolinguistic environments, with IDV societies like the US focusing on the self and COL societies like China focusing on other-relation (further discussed in Chapter 6 below). In terms of speaker confidence definition, semantic relation across groups does seem to correlate with Hofstede's individualist and collectivist perspectives, thus supporting this thesis' hypothesis that sociolinguistic background may indeed affect pragmatic attitudes toward speaker confidence and expression of opinion.

6. PHASE 1 DISCUSSION

This chapter offers a detailed discussion of the main findings for Phase 1 (Proof of Concept) presented in Chapter 5, with reference to each of the relevant RQs. Firstly, results from the perception task are examined and interpreted, followed by discussion of speaker confidence definitions.

6.1. Perception task

The empirical Phase 1 perception task served to address RQs 1, 1a, and 4 by examining individuals' confidence ratings and perceived pragmatic function ratings of different [I+ verb] variations. Each of these RQs and their hypotheses is reiterated below, with the hypotheses supported by results labeled in bold and followed by discussion of interpretation of data.

RQ 1: Are specific prosodic variations of [I + verb] belief constructions linked to different pragmatic interpretations in discourse?

H1: Prosodic variation and pragmatic attitude of [I + verb] statements are correlated in terms of confidence rating (CR). Specifically, it is predicted that CR will decrease from DIR to PRN to V to no kontrast conditions for prosodic variation, and from Deliberate to Uncertain to Reluctant conditions for pragmatic attitude.

H0: CR will not decrease from DIR to PRN to V to no kontrast conditions for prosodic variation; and from Deliberate, Uncertain to Reluctant conditions for pragmatic attitude.

As predicted, results regarding CR and perceived pragmatic function appear to indicate a pragmatic-prosodic link within interpretation of CR of different [I + verb] variations. In Analysis 1's investigation of CR of [I + verb] pragmatic variation, pragmatic function was found to have a significant effect on CR. Additionally, Analysis 2's examination on CR of perceived pragmatic function found a significant effect of perceived pragmatic function on CR of [I + verb], with Deliberate statements being perceived as most confident, Reluctant statements as least confident, and Uncertain statements in the middle across both groups. Moreover, a third analysis establishing a confidence rating correlation between pragmatic attitude and prosodic variation via mean CR for each participant mirrored results of Analysis 1, confirming a significant within-subject effect of prosodic variation on CR.

Thus, all three analyses provided data suggesting a significant link between pragmatic attitude and prosodic variation, wherein every analysis found the same specific confidence rating trend decreasing from for DIR to PRN to V to no kontrast (DMF, DMM, DME) conditions. Furthermore, post-hoc paired t-testing from all three analyses showed significant mean differences between DIR to PRN to V to no kontrast conditions in all three analyses of both pilot and main phase testing, thereby confirming a CR trend which varies based on pragmatic-prosodic mapping of [I + verb] variations.

Whilst mean differences between no kontrast conditions did not always reach significance (no significant mean difference between DMM/DME for main phase results and between DMF/DME for pilot phase results), these results do not contradict the hypotheses: The hypothesis for RQ 1 predicts a trend of high-to-low CR from DIR to PRN to V to no kontrast conditions, on the basis of the same prosodic variation for all no kontrast conditions (with variation being purely syntactic). Thus, results from both pilot and main phase testing support the predicted hypothesis of RQ 1, establishing a proof of the concept regarding a specific pragmatic-prosodic link for perceived confidence of [I + verb] belief constructs.

These results therefore support the notion that focus-marking does create systematic and consistent conversational implicatures; such findings are in line with theories of alternative semantics as presented by Rooth (1985/1992/1996) (refer to section 3.12), in which all non-focal lexical items maintain their original semantic values, whereas focus-marked items take on a different semantic value than their original lexico-semantic identities (Zimmermann & Onea, 2011; Rooth, 1985). Thus, different [I + verb] variations possess semantic contrasts created by kontrast placement in [I+ verb], coinciding with Dehe and Wichmann's (2010) analysis of "I think" and "I believe": DIR statements, by rule of Occam's Razor, are most confident, with opinions being stated frankly as if fact. PRN kontrast placement in [I + verb] (equating to "Main clause" [I + verb] constructs) is next most confident, with focus-marking of the "I" creating a semantic contrast with any other subject and thereby establishing speaker confidence through a strong speaker commitment to the statement (Nikula, 1996). V kontrast placement (equating to "Epistemic marker" [I + verb] constucts), at second least

confident of the levels predicted, creates a semantic contrast between "think/feel/believe" etc. and "know" when stressing an uncertain belief verb, thereby leading to decreased confidence and decreased speaker commitment to one's statement (Kärkkäinen, 2003). Finally, no kontrast marking of [I + verb] (equating to "Discourse marker" [I + verb] constucts), indicating less semantic prominence than the focus-marked variations above (Dehe & Wichmann, 2010; Huang, 2011), renders these conditions least confident of all the prosodic variations.

In summary, results from all three analyses support the notion of a distinctive form-function mapping of [I + verb] variations based on pragmatic meanings created through kontrast placement, with the consistent CR trend in pragmatic-prosodic mapping of [I + verb] decreasing from DIR to PRN to V to no kontrast conditions, thereby supporting this thesis' hypothesis and successfully answering RQ 1. Exploration of RQ 1a will be subsequently discussed.

RQ 1a: For Native vs. Mandarin speakers of English, do these pragmatic-prosodic mappings vary as a function of sociolinguistic background?

H1: Sociolinguistic background affects trend of confidence rating (CR) of [I + verb] statements.

H0: There is no effect of sociolinguistic background on trend of CR of [I + verb] statements.

As predicted, results supported the null hypothesis that there was no effect of sociolinguistic background on trend of confidence rating of [I + verb] statements. In Analysis 1, whilst an interactional effect between pragmatic variation and sociolinguistic background showed significant mean differences between groups at PRN condition, the CR trend across pragmatic-prosodic mapping nevertheless remained congruent for both groups, decreasing from DIR to PRN to V to no kontrast conditions with significant mean differences between these levels. Similarly, whilst in Analysis 3 an interactional effect between pragmatic variation and sociolinguistic background did show significant mean differences at PRN condition, the CR trend nevertheless was congruent across both groups, similarly decreasing from DIR to PRN to V to no kontrast conditions with significant mean differences between these levels. Moreover, Analysis 2 showed no interactional effect between sociolinguistic background and perceived pragmatic function.

Thus, because there was no significant effect of sociolinguistic background, as well as no difference in trend of CRs between Native English and Native Mandarin speakers across all three analyses, results support the predicted cross-linguistic congruence of pragmatic-prosodic mappings. These findings thereby reinforce the notion of focus-marking as a universal cue for prominence as expressed by Zimmermann and Onea (2011) and Rooth (1985/1992) (see section 2.2).

However, it must be noted that a congruent trend in form-function interpretation of [I + verb] does not essentially necessitate universality of pragmatic-prosodic mapping. Further research may choose to focus on testing of other languages outside of Mandarin

and English in order to examine whether the same trend of high-to-low confidence rating from DIR to PRN to V to no kontrast conditions is found. With a greater array of languages examined, there can be more proof for (or potentially against) a theory of universal pragmatic-prosodic mapping of [I + verb] in expressing an opinion.

RQ 5: For Mandarin EFL speakers, does L1 pragmatic knowledge transfer influence the functional use of [I + verb] belief constructions in expressing an opinion in English?

H1: For Mandarin EFL speakers, L1 pragmatic transfer affects the functional use of [I + verb] belief constructions in expressing an opinion in English.

H0: For Mandarin EFL speakers, L1 pragmatic transfer does not affect the functional use of [I + verb] belief constructions in expressing an opinion in English.

Recall from the original presentation of RQs (see section 4.1.1.) that the significance of RQ 5 depends on the results of RQ 1 and RQ 1a: If pragmatic-prosodic mappings are found to vary cross-linguistically between Mandarin and English, then L1 pragmatic knowledge transfer on SLA may act as a confounding variable, and RQ 5 will thereby be investigated in Phase 2. However, if mapping is found to be cross-linguistically similar, then L1 transfer is rendered irrelevant as a confounding variable, given that both languages draw upon similar pragmatic-prosodic mappings.

Thus, since results and discussion for RQ 1 and RQ 1a above indicate a congruous pragmatic-prosodic mapping of [I + verb] between Mandarin and English speakers, then RQ 5 is thereby deemed irrelevant as a factor in Phase 2, That is, given that any effects of pragmatic L1 transfer would be positive transfer (Wang, 2011), L1 pragmatic knowledge would thereby not differ from the L2's pragmatic knowledge. Therefore as predicted, results subsequently support the null hypothesis, and L1 pragmatic transfer should not affect this thesis' findings.

6.2 Speaker confidence definitions

Phase 1 employed mixed-method coding to semantically categorize and tabulate participant's open-ended definitions of speaker confidence. This analysis of speaker confidence definitions served to answer RQ 2. This question and its corresponding hypotheses are reiterated below, with the hypotheses supported by results labeled in bold and followed by discussion of interpretation of data.

RQ 2: Does sociolinguistic background influence pragmatic attitudes toward speaker confidence?

H1: Sociolinguistic background affects pragmatic attitudes toward speaker confidence. Specifically, it is predicted that US individuals will mirror the IndSC

mindset salient within an individualist society, whilst Chinese individuals will mirror the RelSC mindset salient within a relational-collectivist society.

H0: There is no effect of sociolinguistic background on pragmatic attitudes toward speaker confidence.

Analysis of speaker confidence definitions supported the predicted hypothesis that US individuals' definition of speaker confidence more closely mirrored the mindset of an individualistic (IDV) society, whereas Chinese individuals' definitions more closely mirrored the mindset of a relational-collectivistic (Rel-COL) society. Out of the five semantic categories that definitions could fall under (Direct/firm tone of language, Speaker belief toward statement, Listener belief toward statement, Appeal/Reputation of speaker, Comfortableness of speaker to share statement), A Kruskal-Wallis test revealed significant mean differences between groups at all levels. Results for each level are discussed below.

Direct/firm tone of language: US responses in this category were three times as frequent as CHI responses, indicating that US individuals were three times more likely to associate speaker confidence with directness and tone of language than CHI individuals. US responses expressed notions of directness and clarity of speech, with frequently cited words including "clear" (participant no. 7, 11, 17), "assertive" (participant no. 17, 21), and "direct" (participant no. 3, 9, 14, 16), with mention of speaking "without any pauses or moments of hesitation" (participant no. 15, 24) and without "falter[ing] in

communicating an idea" (participant no. 3). Additionally, there was focus on minimizing ambiguity through "direct wording" and particular word choice (participant no. 10, 16, 23). There were also references to the tone, such as "strong and unwavering voice" (participant no. 9, 11, 23, 29), "the level of conviction heard in the speaker's voice" (participant no. 26), or simply "firm[ness]" of tone (participant no. 14, 15). These results coincide with Fox's (2001) findings that Americans may incorporate higher degrees of modality that lead one to appear more assertive or direct, and further correspond with Kim, Sharkey, and Singelis' (1994) and Markus and Kitayama's (1991) claim that an individual's independent self-construal in an individualistic society like the US would lead to more concern for clarity, assertiveness and directness in conversation. Findings further parallel results from Chai and Fontaine's (2007) study, which showed that US individuals preferred high-context communication involving direct verbal communication style, whereas Chinese individuals preferred low-context communication involving indirect verbal communication style.

The few CHI responses in this category referred more to the content of the speech, denoting "successful" or "effective speech" (participant no. 31, 43, 44), or simply stated that speaker confidence involves "tone of voice" (participant no. 34), but without specifying in which manner the tone should be produced or how the speaker's speech should appear in order for success. Such results correspond with Lu's (2002) findings that Chinese speakers may give the impression of being vague or indirect, in order to save face and "manoeuvre" the conversation in the future (p.215).

In summary, the high frequency of "direct" and "assertive" references for US individuals, as well as the low frequency and low specificity of tone-related responses for

Chinese individuals, coincides with the IDV vs. (Rel-)COL attitudes of each respective sociolinguistic environment, as well as with each culture's philosophy toward value on self-expression in US culture and saving-face in Chinese culture.

Speaker belief toward statement: Once again, US speakers appeared three times as likely as Chinese speakers to associate this individualist-focused category with speaker confidence, given that individualist societies emphasize the self and self-belief over belief of others, whilst collectivist individuals tend to define themselves relationally through other's perceptions (Sedikides, 2013, Markus & Kitayama, 1991).

US responses invoked notions of a speaker's "certainty" (participant no. 1, 7, 24), "assuredness" (participants no. 9, 20, 29), "confidence" (participant no. 4, 22, 23), "conviction" (participant no. 19), and "belief" (participant no. 8, 9, 10, 13, 15, 16, 27, 28) that their statement is correct or true[49]. These results correspond to the US culture's focus on developing a "positive self-image" found by Katriel and Philipsen (1981), as well as to Sedekide's (2013) finding of a higher self-esteem in US vs. Chinese individuals, which both indicate a focus on high self-esteem of the individual in US culture. Such high confidence and self-esteem accordingly correlate to higher speaker confidence and thereby more certainty in expressing one's belief.

While a portion of the small amount of CHI responses in this semantic category correspondingly refer to "confidence" (participants no. 46, 49) and "certainty" (participants no. 38, 45), a large portion link the notion of "judgment" (participants no 56, 57, 59) to correctness, a sentiment not expressed by US individuals: Specifically, CHI

[49] Note that a mention of "confidence" or "conviction" without relation to "truth" or "correctness" is not categorized as "speaker belief" but is instead categorized semantically as "speaker comfortableness to share a statement," since it is then a reference to attitude and not a reference to belief.

responses indicate the need for the speaker to have "grasp of their own judgment" and enough "self-confidence [to] impact for the judgment of things" to be correct. Note that while US responses related to self-belief, CHI responses appear to veer more toward the idea of confidence in the absolute correctness or one's judgment; such results correspond to Zou et al.'s (2010) notion that collectivist listeners are less tolerant of individual opinion and more interested in "epistemically sound judgments" which coincide with the social norm (p. 580). These findings furthermore coincide with Kim and Markus' (1999) study, which found that Asian advertisements condoned conformity, whereas American advertisements bolstered uniqueness as the "right way" to be, thereby emphasizing compliance with adopting a general opinion over one's personal opinion. In other words, research suggests that in Chinese culture the "correct" opinion one should put forth is the consensual opinion of the in-group, rather than the opinion of the self, which correlates with Gao's (1998) claim that communication in a collectivist society is not used to affirm self-identity but is instead used to preserve and strengthen communal relations. Indeed, it follows that in the next semantic category that CHI individuals may be more prone to linking speaker confidence to the belief of the *listener* vs. the speaker.

Listener belief toward statement: When associating speaker confidence with listener belief, only one US individual's response fell under this category, compared to 11 CHI responses (nearly 30% of total CHI responses), thus indicating that CHI individuals may be far more likely to associate speaker confidence with the belief of the listener instead of the speaker.

Interestingly, the one US response referencing listener belief (participant no. 27) also references speaker belief in the same statement, defining speaker confidence as "How much they believe what they are saying and how much they are interested in persuading the listener to believe it." Hence, it appears that the only US response in this semantic category focuses not specifically on listener belief like all the CHI responses do, but instead focuses on the *speaker*'s mindset and how the speaker may *persuade* the listener to believe him or her; this is evident by the words, "how much they [the speaker] are interested in persuading the listener" instead of focusing on how the listener may feel and in turn grants control to the speaker. Whilst conversational discourse is always a two-part negotiation of intent and inference, it appears that here this US response focuses more on speaker intent and ability of speaker to shape listener inference.

CHI responses invoked similar sentiments regarding judgment as they did in the previous semantic category, only this time the judgment falls solely with the listener: CHI individuals referred to notions such as "accuracy of judgment by listener"(participant no. 35, 47) and "level of listener interest/concentration" (participants no. 48, 58) in "accepting the speaker's judgment" (participant no. 52, 58) as correct, which may depend on the statement's "plausib[ility] to the audience" (participant no. 54). There is also a notion of listener trust toward the speaker, with responses relating "how effective [the speaker] made you accept his opinions" (participant no. 47) and with a large portion of responses citing how the speaker must "convince the listener" to accept what he said as true (participant no. 49, 50, 51, 60). Whilst these responses and the one US response above both reference the speaker, the difference lies in that these CHI responses concern themselves with the *listener's* interpretation of the speaker's statement and do not place

focus on the speaker's mindset. This finding directly mirrors Yum's (1988) claim that in contrast to US culture's self-centered speaker perspective of communication, Chinese culture emphasizes listening and the listener's interpretation above speaker intent. That is, it is the judgment of the listener that determines the appraisal of the speaker. This other-focused tendency may once more relate to the (Rel-)COL focus on face-saving as previously stated by Liu (2013), for even if a speaker may believe his or her statement to be correct, ultimately the appraisal of the speaker depends on whether that statement is accepted by the audience regardless of speaker belief; thus, the speaker must in turn and understand what the listener believes to be correct.

Overall CHI participants' definition of speaker confidence appears more other-dependent, aligning with a collectivist other-centric focus and salience of relational self-construal. A dearth of US emphasis on the listener coincides with an individualistic perspective of self-centric focus and salience of independent self-construal, as viewed by the numerous US definitions of belief of one's *own* statement; this strongly coincides with Choi, Nisbett and Norenzayan's (1999) as well as Choi and Markus' (1998) results on internal behavioural attribution of US individuals vs. external behavioural attribution of East Asian individuals (See section 2.1.3.1).

Appeal/Reputation of speaker: Mirroring results for the category above, only one US response fell under this category, in comparison to 16 CHI responses. This category also happened to be the semantic category most cited by CHI speakers.

The only US response in this semantic category also fell into the following category of Comfortableness to share; in relation to Appeal/reputation of speaker, this US response

states that even if one's statement is wrong, it "won't diminish your [the speaker's] standing with the person you're speaking to." Once more it appears that US individuals inevitably link their focus regarding speaker confidence back to the speaker rather than focusing on the listener.

As the prominent semantic category for CHI participants, responses invoked three trains of thought regarding speaker appeal/reputation. The first revolves around "judgment [of] the speaker's capacity/potential" (participants no. 37, 39, 43, 46) or "judgment of the substance of a speaker's mind" (40, 43, 53), that is the actual ability of the speaker to judge and produce correct content. The second involves the attractiveness of the speaker in terms of how "captivating"(participants no. 37, 39, 43, 46) or "important" (participant no. 42) a speaker may be. The third relates to the character of the speaker, with responses citing words like "trustworthy" (participants no. 33, 41, 47) and "credible"(53, 55). All three of these categories involve judgment from the listener, thus pulling the focus away from the speaker and revealing the *listener* as the determiner of speaker confidence.

In an other-relational collectivist society like China, how one appears to others may involve far more nuanced layers than in an individualist society like the US. Accordingly, the lack of US references to speaker appeal/reputation may indicate a lack of focus on other's opinions, whilst the majority of CHI responses falling under this category, combined with the high frequency of CHI responses in the listener belief category above, seems to strongly correlate with the other-centric notion of a collectivist sociolinguistic environment with relational self-construal (Markus & Kitayama, 1991; Kitayama & Uskul, 2011). That is, in a society where identity is relationally constructed through the

eyes of others, the appeal or reputation of a speaker may be crucially linked to how another appraises and defines the speaker.

Comfortableness of speaker to share statement: It is interesting to note that while 6 US responses cited this semantic category, no CHI speakers cited this category at all, indicating that CHI speakers may not consider comfortableness of the speaker to be relevant to speaker confidence. The absence of CHI responses in this category may once again coincide with the other-centric focus of a relational-collectivist society, whereas the US focus on the speaker once more mirrors the self-centric focus of an individualistic society as determined above.

As the third most frequently cited semantic category for US participants, responses ranged from general notions of "how comfortable the speaker is in expressing their opinion" (participant no. 3, 18) and "self-esteem" (participant no. 7) to more specific ideas of overcoming apprehension, such as being "comfortable enough to risk being wrong" (participant no. 5) and "not being afraid to say so" (participant no. 13). This self-assured US notion parallels the individualistic right of self-expression without fear of judgment due to higher levels of interpersonal trust within societies with independent self-construal, as expressed by Ingelhart and Welzel (2010) and Hofstede (1981/2011).

Consequently, a lack of annotations may indicate that for CHI speakers, a high level of speaker comfort is not relevant to speaker confidence the way it may be to US speakers. Recall that in a relational-collectivist society where one's identity is relationally constructed; comfortableness to share, a solely personal and internally seated emotion,

does not extend to others. Thus, this category may not be as relevant to CHI individuals as it is to the internally self-defined individuals of the US.

Moreover, the absence of comfortableness to share one's opinion may coincide with Chinese culture's notion of saving face and reluctance to commit to a statement, for fear of limiting any manoeuvring of future language and perhaps a malleable, changeable disposition of self (Norenzayan, Nisbett & Smith, 1998). By defining one's attitude or belief toward one's own statement, one would be taking responsibility for it (Kärkkäinen, 2003); thus, individuals in a collectivist society may be less comfortable in their desire to share their opinion for fear of being proven wrong in the future (Zou et all, 2010; Fox, 2002). These results further correlate with Fong's (1998) finding that CHI individuals give fewer, less encouraging and less strong compliments than their US counterparts, whereas compliments were an essentially expected response in US culture. That is, in a culture like China where positive or encouraging reactions are fewer and less explicit, it follows logically then that comfortableness is less taken into consideration.

In sum, US individuals' definitions had a significantly higher frequency for Direct/firm tone of language, Speaker belief toward statement, and Comfortableness to share statement, all categories in which the speaker and the speaker's self-expression are emphasized. These findings thereby coincide with notions that Americans place more value on the individual self when ranking themselves higher in self-esteem ratings than Chinese individuals, paralleling individualistic notions of an independent self-construal (Sedikides, 2013; Hofstede, 1980, Markus & Kitayama, 1991). Conversely, Chinese individuals' definitions had a significantly higher frequency for Listener belief toward

statement and Appeal/Reputation of Speaker, both semantic categories in which focus on the listener and the listener's judgment is significant, as is to be expected within a relational-collectivist culture with interdependent self-construal (Yum, 1984 Hofstede, 1980, Markus & Kitayama, 1991).

Furthermore, peak and lowermost frequencies within categories appeared to have an inverse relationship between US and Chinese participants: Peak frequencies for US annotations (Speaker belief toward statement, Directness/tone of language), appear in the same category where CHI annotations exhibit lower frequencies. In contrast, peak frequencies for CHI annotations (Appeal/Reputation of speaker, Listener belief of speaker) appear in the category with the lowest frequencies for US annotations. This inverse dichotomy may be related to the inverse relationship between individualistic and collectivistic mindsets (and thereby correlating with independent-interdependent self-construals), especially with regard to self-expression and identity. A society with individualistic tendencies as the US may promote speakers to express confidence via directness and assertiveness with confidence being related to willingness to share, as well as being related to the speaker's own belief of his utterance, given that identity is self-determined by (and thus only limited by) the individual. Conversely, in a society with collectivistic tendencies like China, individuals may be less likely to express oneself using a direct or assertive tone in order to lessen speaker commitment and save face, while at the same time the other-dependent identity of such a culture may nuance individuals to think that confidence of a speaker is determined more by his relation to the listener (the other) and how the speaker's reputation appeals to the listener.

Thus, in response to RQ 2, results from all five semantic categories appear to support this thesis' hypothesis that sociolinguistic background affects pragmatic attitudes toward speaker confidence in a manner where US individuals will mirror the mindset of an individualistic society with independent self-construal, whilst Chinese individuals will mirror that of a collectivist society with interdependent self-construal. Hence for the former, speaker confidence is appraised and determined by the individual self, or the speaker, whereas for the latter, it is evaluated and determined by the other, or the listener.

Limitations, recommendations for future research and implications of Phase 1's findings will be presented alongside those of Phase 2 in Chapter 9, after presentation and interpretation of Phase 2's findings. The following chapter examines Phase 2's results on how [I + verb] usage may vary as a function of sociolinguistic background and discourse context.

7. PHASE 2 RESULTS

Results for Phase 2 are reported in this chapter in the order that Phase 2 analyses were presented in section 4.3.4. Length of audio recording is examined first, followed by analyses for frequency of [I + verb]. Next is an in-depth sub-analysis for EFL individuals in order to investigate the effect of L2 proficiency on frequency of [I + verb], followed by a further in-depth sub-analysis on EFL US group to examine the effect of Time immersed in a new sociolinguistic environment on frequency of [I + verb]. Significance was set at the .05 value (Berry, 1997:260).

7.1. Length of audio recording

Length of Audio recording (as per total utterances produced) was compared between-groups to identify whether variability in mean difference of total utterances between groups could influence frequency of [I + verb]. Incomplete responses to prompts (i.e. responses stopping short of the 15 second mark), which were excluded from analysis, accounted for .67% of responses in the Self-opinion prompt set (6 out of 900), and 1.11% of responses in the Opinion of others prompt set (10 out of 900). Moreover, incomplete responses did not appear to represent a bias in participants' responses in certain categories; instead, such responses were spread across nearly all prompts and occurred across all groups. Thus, given such distribution and the tiny amount of incomplete data in relation to the total data set, data exclusion should not compromise subsequent analyses. Furthermore, recall that measuring length of audio recording via mean utterances

produced per prompt for each participant alleviates any difficulties that could arise from an uneven number of responses between participants.

Results are accordingly presented below for both prompt sets of the open production task.

Part 1: Self-opinion

Part 1 of the production task consisted of 9 open-ended direct opinion prompts, asking participants to offer their opinion regarding self-preference and beliefs.

Means and SDs for utterances per prompt are presented in Table 7.1.

Table 7.1: Means and SDs for utterances per prompt for Self-opinion set

Group	Mean utterances/prompt	SD
Native US	5.17	.57
EFL US	5.07	.71
EFL CHI	4.81	.74
Native CHI	4.86	.53
Total (average)	4.98	.65

Levene's test suggested homogeneity of variance ($p = .62$). A one-way ANOVA showed no significant effect of group on mean utterances per prompt ($F(1,3) = 1.80$, $p = .15$). Additionally, post-hoc paired t-tests showed no significant difference between any groups ($p > .05$). Thus, any variability in length of data between participants should not influence frequency results.

Part 2: Opinion of others

The second prompt set of the production task consisted of 3 sets of short stories wherein an abstract character makes an ambiguous mistake, followed by 3 opinion questions each, asking the participant to offer their opinion regarding social judgment of the character in the story.

Means and SDs for utterances per prompt are presented in Table 7.2.

Table 7.2: Means and SDs for utterances per prompt for Opinion of others set

Group	Mean utterances/prompt	Standard deviation
Native US	4.55	.62
EFL US	4.61	.51
EFL CHI	4.41	.54
Native CHI	4.60	.95
Total (average)	4.54	.55

Levene's test suggested homogeneity of variance ($p = .80$). A one-way ANOVA showed no significant effect of group on mean utterances per prompt ($F(1,3) = .69$, $p = .56$). Additionally, post-hoc t-testing did not reveal any significant differences between any groups. Thus, any variability in length of data between participants should not influence frequency results.

7.2. Frequency of functional [I + verb] usage

Three analyses were carried out on the data from the production task (as presented in section 4.3.4.) to identify whether frequency of functional [I + verb] usage varies as a function of sociolinguistic background, and whether this usage varies between prompt sets regarding Self-opinion (personal preferences/beliefs) vs. Opinion of others (appraisals/judgments) (RQ 3).

As outlined in section 4.3.4, the first two analyses differ only in their dependent variable (Percentage of pragmatic function of [I + verb] to total utterances produced vs. Percentage of pragmatic function of [I + verb] to total [I + verb] constructs used). The third analysis uses a calculated dependent variable, in which average confidence ratings are tabulated for each subject based on percentage of each [I + verb] condition out of total constructs used, multiplied by confidence rating values for each construct variation from Phase 1. For each analysis, results are subsequently presented for each prompt set of the production task.

7.2.1. ANALYSIS 1 - Percentage of pragmatic function of [I + verb] to total utterances produced

Results are presented below for both prompt sets of the open production task. Descriptive statistics are given first, followed by presentation of inferential statistics.

Part 1: Self-opinion

Means for Percentage of pragmatic function to total utterances are presented in Table 7.3[50].

Table 7.3: Means and SDs for Percentage of pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	Native US	25.27	9.93
	EFL US	14.18	5.62
	EFL CHI	7.40	5.35
	Native CHI	3.43	2.17
	Total	12.48	10.28
Uncertain (U)	Native US	20.47	10.25
	EFL US	13.65	10.00
	EFL CHI	6.77	6.91
	Native CHI	5.74	4.05
	Total	11.66	10.17
Reluctant (R)	Native US	4.47	3.74
	EFL US	10.71	4.76
	EFL CHI	19.41	7.13
	Native CHI	15.32	9.68
	Total	12.48	8.94

Shapiro-Wilk testing indicated normality of data for all groups at Deliberate condition and half of the groups at Uncertain (Native US, Native CHI) and Reluctant (EFL US,

[50] Note that all means given are analyzed with Age as covariate, as previously stated in the analysis chapter.

EFL CHI) conditions respectively (see Table 7.4). Whilst data distribution exhibited positive skew at these four conditions, there were no ceiling or floor effects present. Additionally, no data points were identified outside of 3 standard deviations from the mean (Howell, 1998; refer to pg. 140 for reference), and thereby there should be no outliers affecting normality. Given that data distribution appeared overall normal, parametric ANCOVA testing was accordingly conducted.

Table 7.4: Shapiro-Wilk test results for Phase 2 prompt set 1

Pragmatic Attitude	Group	DF	S-W Statistic	Significance
Deliberate	**Native US**	25	**.94**	**.13**
	EFL US		**.96**	**.36**
	EFL CHI		**.93**	**.08**
	Native CHI		**.93**	**.07**
Uncertain	**Native US**		**.98**	**.82**
	EFL US		.91	.03
	EFL CHI		.83	<.01
	Native CHI		**.96**	**.40**
Reluctant	Native US		.92	.05
	EFL US		**.94**	**.16**
	EFL CHI		**.95**	**.24**
	Native CHI		.90	.02

Results denoting normality of data distribution indicated in bold.

To investigate the frequency of pragmatic function of [I + verb] usage (RQ 3), a repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor

(sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 9.39, p = .01), results calculated with Mauchly's sphericity vs. Greenhouse-Geiser (.91) or Huynh-Feldt measurements (.97) do not differ regarding significance. Since the Greenhouse-Geiser measurement did not fall below .75, the Hunyh-Feldt estimate was chosen as a correction for sphericity given that Greenhouse-Geiser measurements may be too conservative (Hyunh & Feldt, 1976, Field, 2013:1778). Positive results for near-sphericity are furthermore signified by the close proximity of the Hunyh-Feldt measurement to 1 vs. the lower-bound estimate (1/([# conditions] – 1) = .5) (Field, 2013:1777). Levene's test did not indicate homogeneity of variance ($p < .01$ for Deliberate and Uncertain conditions; P = .01 for Reluctant condition); therefore, non-parametric Kruskal-Wallis test results are also included in Appendix 15 for reference.

ANCOVA results revealed a between-group effect of sociolinguistic background ($F(1,3) = 22.26$, $p < .001$, $\eta_p^2 = .41$). Additionally, there was a significant within-group effect of pragmatic function ($F(1,2) = 3.10$, $p = .05$, $\eta_p^2 = .03$). Moreover, results also showed a large interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 30.80$, $p < .001$, $\eta_p^2 = .49$). There did not appear to be a statistically significant effect of age ($F(1,1) = 1.97$, $p = .16$, $\eta_p^2 = .02$) or interactional effect between pragmatic function and age ($F(1,2) = 2.67$, $p = .07$, $\eta_p^2 = .03$).

The interaction between sociolinguistic background and pragmatic function was further investigated through post-hoc paired t-tests with Bonferroni adjustment for

multiple comparisons on all three pragmatic function conditions between-groups. These mean differences may be observed in Figure 7.1 below. Results are presented in Table 7.5, with significant mean differences in bold.

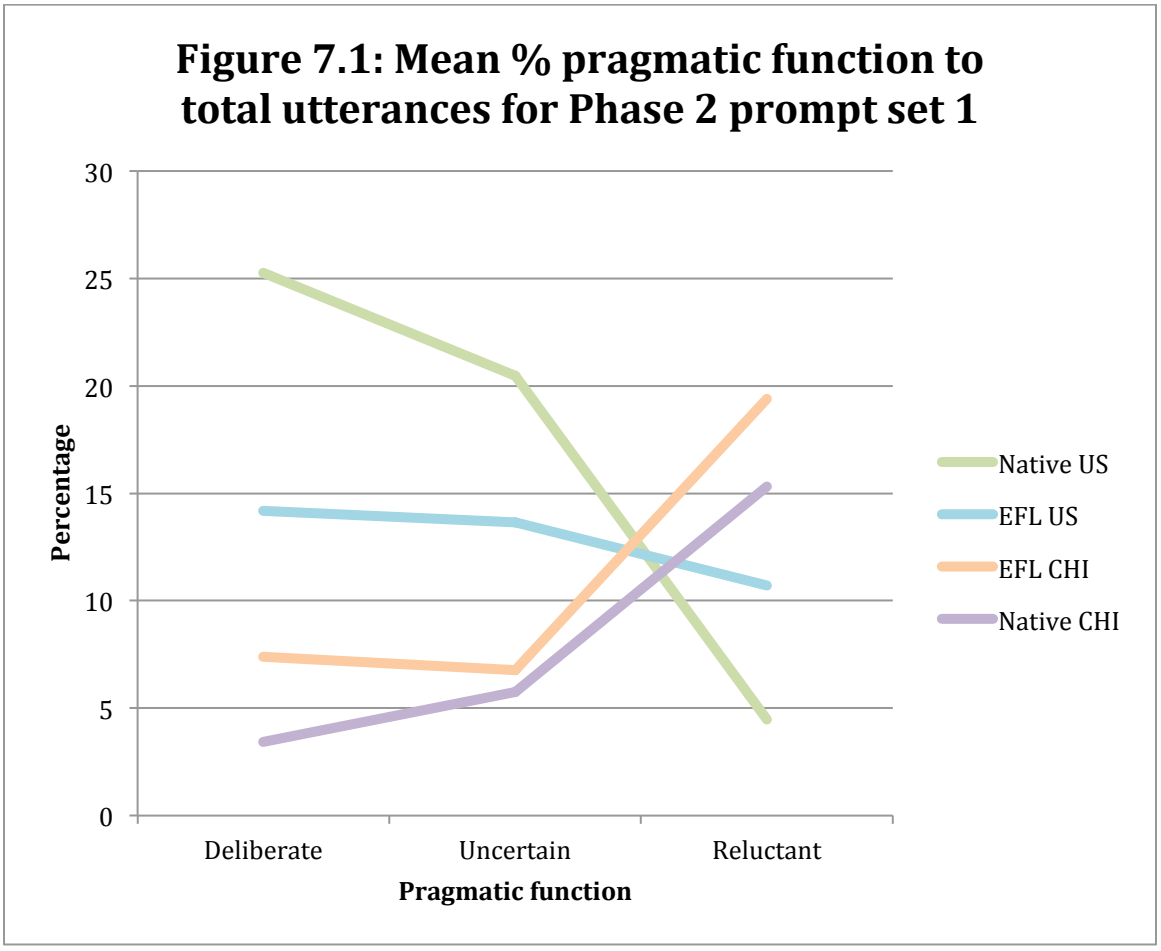


Table 7.5: Post-hoc paired t-tests, Percentage of pragmatic function to total utterances for Self-opinion

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
Native US to EFL US	**Deliberate**	**11.09**	**<.001**	**1.84**
	Uncertain	**6.82**	**.03**	**2.38**
	Reluctant	**-6.24**	**.01**	**1.92**
Native US to EFL CHI	**Deliberate**	**17.87**	**<.001**	**1.84**
	Uncertain	**-13.70**	**<.001**	**2.39**
	Reluctant	**-14.94**	**<.001**	**2.32**
Native US to Native CHI	**Deliberate**	**-21.84**	**<.001**	**1.79**
	Uncertain	**14.73**	**<.001**	**2.32**
	Reluctant	**-10.86**	**<.001**	**1.88**
EFL US to EFL CHI	**Deliberate**	**6.78**	**.01**	**1.98**
	Uncertain	**6.88**	**.05**	**2.57**
	Reluctant	**-8.70**	**<.001**	**2.07**
EFL US to Native CHI	**Deliberate**	**10.75**	**<.001**	**1.87**
	Uncertain	**7.91**	**..01**	**2.42**
	Reluctant	-4.62	.12	1.95
EFL CHI to Native CHI	Deliberate	3.90	.19	1.82
	Uncertain	1.03	1.00	2.36
	Reluctant	4.09	.21	1.91

Notice the trend within each condition: Deliberate condition percentage decreases from Native US to EFL US to EFL CHI/CHI, with higher levels of US sociolinguistic background associated with higher Deliberate condition percentage and higher levels of CHI sociolinguistic background associated with lower Deliberate condition percentage. The data for Uncertain condition percentage exhibits this trend as well. For Reluctant

condition percentage, the trend appears to be opposite, with higher levels of US sociolinguistic background associated with lower Reluctant condition percentage and higher levels of CHI sociolinguistic background associated with higher Reluctant condition percentage.

Additionally, performance of EFL speakers in the US (EFL US) mirrors that of native monolingual US speakers (Native US) more than the performance of EFL speakers in China (EFL CHI) does, whilst performance of EFL CHI mirrors that of Mandarin native monolingual speakers (Native CHI). It may be noted that whilst there are no significant mean differences at any levels between EFL CHI and Native CHI, the mean differences between Native US and EFL US are significant, and the Reluctant condition showed no significant mean difference between EFL US and Native CHI.

Overall, it appears that US sociolinguistic background leads to more assertive and confident usages of [I + verb] constructs. Conversely, CHI sociolinguistic background appears to lead to less confident usage of [I +verb].

Part 2: Opinion of others

Means for Percentage of pragmatic function to total utterances are presented in Table 7.6.

Table 7.6: Means and SDs for Percentage of pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	Native US	5.36	3.06
	EFL US	11.12	4.63
	EFL CHI	27.50	11.08
	Native CHI	27.23	8.23
	Total	17.80	12.38
Uncertain (U)	Native US	11.56	8.65
	EFL US	13.20	8.04
	EFL CHI	9.25	8.86
	Native CHI	5.28	4.72
	Total	9.82	8.20
Reluctant (R)	Native US	7.00	5.51
	EFL US	7.27	4.86
	EFL CHI	8.48	5.31
	Native CHI	4.88	3.80
	Total	6.91	5.01

Shapiro-Wilk testing indicated normality of data for all groups at all conditions except EFL CHI and Native CHI groups at Uncertain condition and Native US group at Reluctant condition (see Table 7.7). While data distribution at these three conditions exhibited positive skew, there were no ceiling or floor effects. Additionally, no data points were identified outside of 3 standard deviations from the mean, and thereby there should be no outliers affecting normality. Thus, given that data distribution overall appeared normal, parametric testing was conducted.

Table 7.7: Shapiro-Wilk results for Phase 2 set 2, % construct to total utterances

Pragmatic Attitude	Group	DF	S-W Statistic	Significance
Deliberate	**Native US**	25	**.93**	**.08**
	EFL US		**.99**	**.98**
	EFL CHI		**.96**	**.43**
	Native CHI		**.97**	**.64**
Uncertain	**Native US**		**.93**	**.08**
	EFL US		**.93**	**.08**
	EFL CHI		.87	<.01
	Native CHI		.89	.01
Reluctant	Native US		.89	.01
	EFL US		**.95**	**.29**
	EFL CHI		**.95**	**.30**
	Native CHI		**.93**	**.09**

Results denoting normality of data distribution indicated in bold.

To investigate the frequency of pragmatic function of [I + verb] usage (RQ 3), a repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed ($X^2(1,2) = 27.67$, $p < .001$), results calculated with Mauchly's sphercity vs. Huynh-Feldt measurements (.84) do not differ. Additionally, Levene's test indicated homogeneity of variance at all levels ($p = .08$ for Uncertain condition, $p = .25$ for Reluctant condition) except Deliberate condition ($p < .001$). Thus, ANCOVA testing was conducted.

Results revealed a small within-subject effect of pragmatic function ($F(1,2) = 7.92$, p

$< .01, \eta_p^2 = .08$) as well as a large between-subject effect of sociolinguistic background $F(1,3) = 17.53, p < .001, \eta_p^2 = .36$). Furthermore, there appeared a large interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 24.93, p < .001, \eta_p^2 = .44$). While results indicated a small effect of age ($F(1,1) = 17.53, p < .001, \eta_p^2 < .01$), there did not appear to be an interactional effect between age and pragmatic function ($F(1,2) = .15, p = .83, \eta_p^2 < .01$).

To investigate the interactional effect between pragmatic function and sociolinguistic background, post-hoc paired t-testing was conducted between-groups on all three pragmatic function conditions. These mean differences may be observed in Figure 7.2 below. Results are presented in Table 7.8, with significant mean differences in bold.

Figure 7.2: Mean % pragmatic function to total utterances for Phase 2 prompt set 2

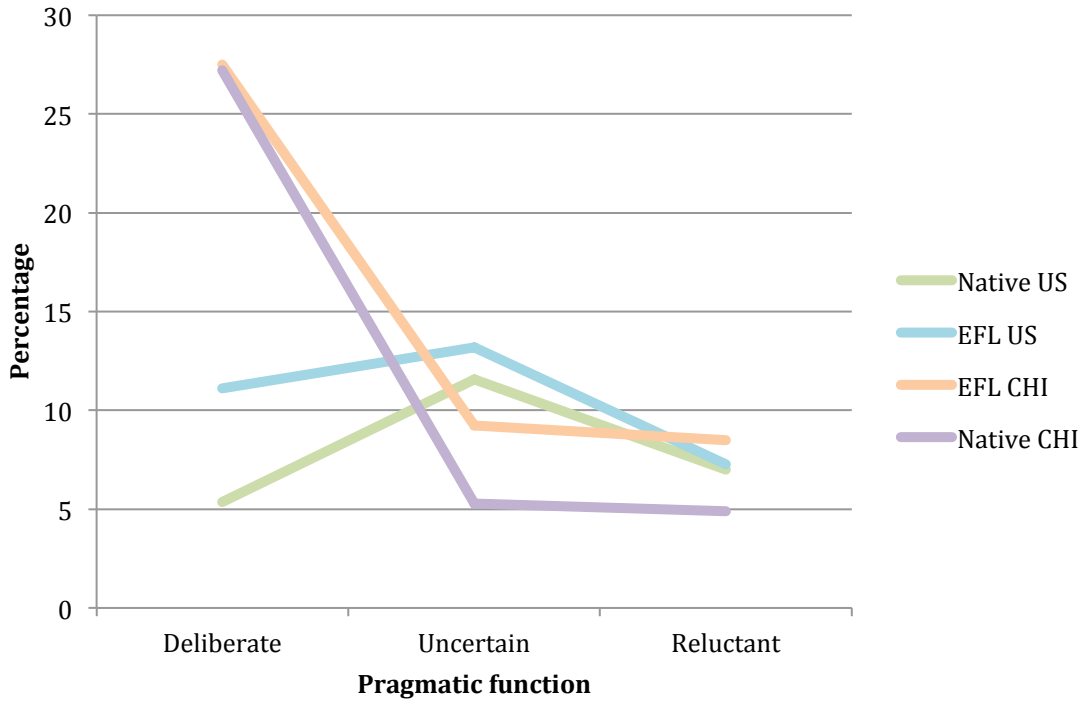


Table 7.8: Post-hoc paired t-tests for Percentage of pragmatic function to total utterances for prompt set 2

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
Native US to EFL US	Deliberate	-5.76	.06	2.17
	Uncertain	-1.64	1.00	2.26
	Reluctant	-.27	1.00	1.43
Native US to EFL CHI	**Deliberate**	**-22.15**	**<.001**	**2.17**
	Uncertain	2.31	1.00	2.27
	Reluctant	-1.48	1.00	1.44
Native US to Native CHI	**Deliberate**	**-21.87**	**<.001**	**2.12**
	Uncertain	**6.23**	**.03**	**2.21**
	Reluctant	2.13	.79	1.40
EFL US to EFL CHI	**Deliberate**	**-16.38**	**<.001**	**2.34**
	Uncertain	3.95	.65	2.44
	Reluctant	-1.21	1.00	1.55
EFL US to Native CHI	**Deliberate**	**-16.11**	**<.001**	**2.20**
	Uncertain	**7.92**	**.01**	**2.30**
	Reluctant	2.40	.62	1.46
EFL CHI to Native CHI	Deliberate	.28	1.00	2.15
	Uncertain	3.97	.48	2.24
	Reluctant	3.60	.08	1.42

Notice that the inverse in trend of Percentage of pragmatic function of [I + verb] to total utterances for Deliberate and Reluctant conditions between part 1 (Self-opinion) and part 2 (Opinion of others) of the production task. For part 2, percentage Deliberate condition deceases with more US and less CHI sociolinguistic background. For Reluctant condition, Percentage is lowest for Native US group and highest for EFL US group, but with little mean difference between groups. Trend for Percentage Uncertain condition

still mirrors that of the Self-opinion prompt set, with percentage decreasing with more US and less CHI sociolinguistic background; however, mean differences between groups were not as large compared with results of the first prompt set.

Once again, EFL US's performance mirrors that of Native US whilst EFL CHI mirrors that of Native CHI. It is interesting to note that in this prompt set, there were no significant mean differences between Native US and EFL US. Additionally, similarly to the first prompt set there were no significant mean differences between EFL CHI and Native CHI groups, who both derive from a Chinese sociolinguistic background. For the second prompt set, there also appeared less significant variation between the two EFL groups, who differed in the addition of US sociolinguistic background for EFL US group. Whilst there were significant mean differences for Uncertain and Reluctant conditions, there were no significant mean differences for Deliberate condition. Significant mean differences also appeared for Deliberate conditions between Native US vs. EFL CHI and Native CHI groups, as well as between EFL US and Native CHI. For Uncertain condition, significant mean differences were found between Native US and Native CHI, and between EFL US and Native CHI. Results for Reluctant condition did not show significant mean differences between any groups.

7.2.2. ANALYSIS 2 - Percentage of pragmatic function of [I + verb] construct to total constructs used

To further explore the frequency of pragmatic function of [I + verb] usage (RQ 3), a second analysis was run, differing from Analysis 1 only in its dependent variable, to

examine whether percentage of pragmatic function of [I + verb] constructs to total constructs used also indicated the same results as Analysis 1 regarding influence of sociolinguistic background. Results are presented below for both prompt sets of the open production task. Descriptive statistics are given first, followed by presentation of inferential statistics.

Part 1: Self-opinion

Means for Percentage of pragmatic function to total utterances are presented in Table 7.9.

Table 7.9: Means and SDs for % pragmatic function to total constructs

Pragmatic function	Group	Mean	SD
Deliberate (D)	Native US	50.86	19.00
	EFL US	37.79	14.86
	EFL CHI	21.98	13.35
	Native CHI	18.29	12.77
	Total	32.23	19.74
Uncertain (U)	Native US	40.03	15.75
	EFL US	34.19	18.80
	EFL CHI	20.01	17.94
	Native CHI	21.46	12.54
	Total	28.92	18.79
Reluctant (R)	Native US	9.11	7.65
	EFL US	28.02	10.67
	EFL CHI	58.01	14.93
	Native CHI	60.25	14.30
	Total	38.85	25.19

Shapiro-Wilk testing indicated normality of data for all groups at all conditions except EFL CHI group at Uncertain condition and Native US group at Reluctant condition (see Table 7.10). Further exploratory analysis within these two distributions showed positive skew, albeit with no ceiling or floor effects present. Additionally, no outliers were present outside of 3 standard deviations from the mean and thereby normality should not be affected. Thus, given that data distribution overall appeared normal, parametric testing was conducted.

Table 7.10: Shapiro-Wilk test results for % pragmatic function to total constructs, prompt set 1

Pragmatic Attitude	Group	DF	S-W Statistic	Significance
Deliberate	**Native US**	25	**.95**	**.25**
	EFL US		**.96**	**.43**
	EFL CHI		**.96**	**.45**
	Native CHI		**.96**	**.39**
Uncertain	**Native US**		**.97**	**.59**
	EFL US		**.96**	**.32**
	EFL CHI		.85	.00
	Native CHI		**.94**	**.17**
Reluctant	Native US		.92	.04
	EFL US		**.94**	**.18**
	EFL CHI		**.97**	**.64**
	Native CHI		**.95**	**.19**

Results denoting normality of data distribution indicated in bold.

A repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI), with age as a covariate, and with a DV of % of Percentage of pragmatic function of [I + verb] to total [I + verb] constructs produced.

While sphericity cannot necessarily be assumed ($X^2(1,2) = 15.22$, $p < .001$), results calculated with Mauchly's sphercity vs. Huynh-Feldt measurements (.92) do not differ. Additionally, Levene's test indicated homogeneity of variance at all levels ($p = .10$ for Uncertain condition, $p = .06$ for Reluctant condition) except Reluctant condition ($p < .001$). Thus, ANCOVA testing was conducted.

Results revealed a small within-subject effect of pragmatic function ($F(1,2) = 5.14$, $p = .01$, $\eta_p^2 = .05$), with no significant between-subject effect of sociolinguistic background $F(1,3) = <.001$, $p = 1.00$, $\eta_p^2 < .001$). Nonetheless, there was a large interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 33.92$, $p < .001$, $\eta_p^2 = .52$). Additionally, the effect of age ($F(1,1) < .001$, $p = 1.00$, $\eta_p^2 <.001$) and interactional effect between age and pragmatic function ($F(1,2) = 2.30$, $p = .11$, $\eta_p^2 < .05$) was not found to be statistically significant.

To investigate the interactional effect between pragmatic function and sociolinguistic background, post-hoc paired t-tests were conducted between-groups on all three pragmatic function conditions. These mean differences may be observed in Figure 7.3 below. Results are presented in Table 7.11, with significant mean differences in bold.

Figure 7.3: Mean % pragmatic function to total constructs for Phase 2 prompt set 1

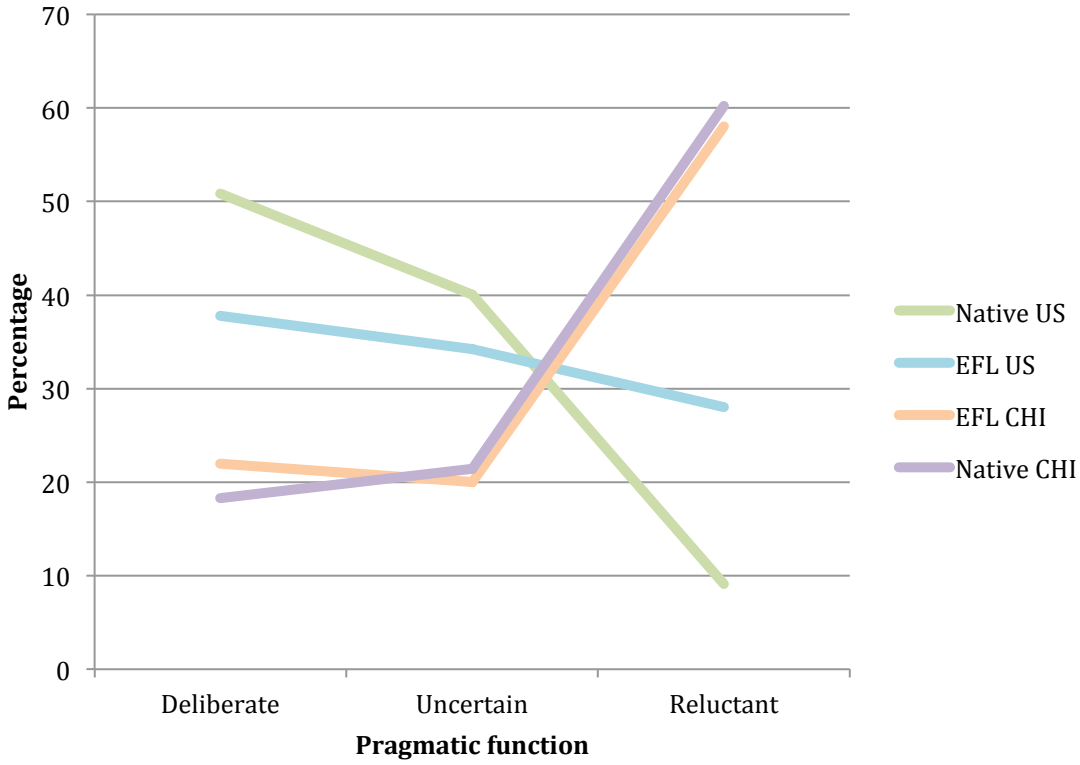


Table 7.11: Post-hoc testing for % pragmatic function to total constructs, prompt set 1

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
Native US to EFL US	**Deliberate**	**13.07**	**.02**	**4.43**
	Uncertain	5.83	1.00	4.71
	Reluctant	**-18.91**	**<.001**	**3.51**
Native US to EFL CHI	**Deliberate**	**28.88**	**<.001**	**4.44**
	Uncertain	**20.02**	**<.001**	**4.73**
	Reluctant	**-48.90**	**<.001**	**3.52**
Native US to Native CHI	**Deliberate**	**32.57**	**<.001**	**4.32**
	Uncertain	**18.57**	**<01**	**24.60**
	Reluctant	**51.14**	**343**	**<001**
EFL US to EFL CHI	**Deliberate**	**15.81**	**.01**	**4.77**
	Uncertain	**14.18**	**.04**	**5.08**
	Reluctant	**-29.99**	**<.001**	**3.78**
EFL US to Native CHI	**Deliberate**	**19.50**	**<.001**	**4.49**
	Uncertain	12.74	.06	4.78
	Reluctant	**-32.23**	**<.001**	**3.57**
EFL CHI to Native CHI	Deliberate	3.69	1.00	4.38
	Uncertain	-1.45	1.00	4.67
	Reluctant	-2.24	1.00	3.48

Notice a decreasing trend in both Deliberate and Uncertain condition percentages as US sociolinguistic background increases and CHI sociolinguistic background decreases; following predictions, this trend matches the trend for these conditions in the first analysis. Similarly, the increasing trend in Reluctant condition percentage as CHI sociolinguistic background increases and US sociolinguistic background decreases also

mirrors the results for this condition in the first analysis. Once again, Native US and EFL CHI/Native CHI groups appear to have inversely correlated trends regarding confidence of pragmatic function. As predicted, EFL US's performance appears more closely correlated with Native US group's performance vs. Native CHI's performance, whilst EFL CHI's performance appears virtually parallel to that of Native CHI.

It is interesting to note that within the Uncertain condition, there is no significant mean difference between the two US-residing groups, but a significant mean difference exists between the US groups and EFL CHI group. However, there does not seem to be a significant mean difference between EFL US and Native CHI at Uncertain condition. Furthermore, following with predictions, there is no mean difference at any condition between the two CHI-residing groups.

Part 2: Opinion of others

Means for Percentage of pragmatic function to total constructs are presented in Table 7.12.

Table 7.12: Means/SDs for % pragmatic function to total utterances, prompt set 2

Pragmatic function	Group	Mean	SD
Deliberate (D)	Native US	26.63	16.67
	EFL US	34.79	12.83
	EFL CHI	62.85	21.26
	Native CHI	73.62	17.01
	Total	49.71	25.55
Uncertain (U)	Native US	43.64	24.21
	EFL US	43.61	19.33
	EFL CHI	18.21	18.10
	Native CHI	13.78	12.32
	Total	29.81	22.86
Reluctant (R)	Native US	29.73	18.27
	EFL US	21.61	14.01
	EFL CHI	18.95	10.06
	Native CHI	12.61	8.41
	Total	20.72	16.67

Shapiro-Wilk testing indicated normality of data for all groups at all conditions except EFL CHI/Native CHI at Uncertain condition (see Table 7.13). Further exploratory analysis of distribution within these two conditions showed positive skew, but no ceiling or floor effects were present. Additionally, no data points were present outside of 3 standard deviations from the mean, and thereby there should be no outliers affecting normality. Thus, given that data distribution overall appeared normal, parametric testing was conducted.

Table 7.13: Shapiro-Wilk results for part 2, % pragmatic function to total constructs

Pragmatic Attitude	Group	DF	S-W Statistic	Significance
Deliberate	**Native US**	25	**.93**	**.08**
	EFL US		**.99**	**.98**
	EFL CHI		**.96**	**.43**
	Native CHI		**.97**	**.64**
Uncertain	**Total**		**.98**	**.90**
	Native US		**.98**	**.86**
	EFL US		.89	.01
	EFL CHI		.92	.04
Reluctant	**Native CHI**		**.93**	**.09**
	Total		**.93**	**.10**
	Native US		**.93**	**.09**
	EFL US		**.96**	**.43**

Results denoting normality of data distribution indicated in bold.

A repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total [I + verb] constructs produced.

Mauchly's test indicated sphericity (p = .12) and Levene's test indicated homogeneity of variance in all levels (p = .96 for Deliberate, p = .33 for Uncertain) except Reluctant (p = .02). Given these results, ANCOVA testing was conducted.

Results revealed a small within-subject effect of pragmatic function ($F(1,2) = 5.26$, $p < .01$, $\eta_p^2 = .05$). However, there was no significant between-group effect of sociolinguistic background ($F(1,3) = 0$, $p = 1.00$, $\eta_p^2 < .001$). There was an interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 22.39$, $p <$

.001, η_p^2 = .41). Additionally, the effect of age (F(1,1) < .001, p = 1.00, η_p^2 <.001) and interactional effect between age and pragmatic function (F(1,2) = .93, p = .40, η_p^2 < .01) did not appear to be statistically significant. It must be noted that observed power for this interaction (.21) is below the .8 mark (Cohen, 1988), indicating that this particular significance may not necessarily be reliable, and thus effect size should be considered.

To investigate the interactional effect between pragmatic function and sociolinguistic background, post-hoc paired t-tests were conducted between-groups on all three pragmatic function conditions. These mean differences may be observed in Figure 7.4 below. Results are presented in Table 7.14, with significant mean differences in bold.

Figure 7.4: Mean % pragmatic function to total constructs for Phase 2 prompt set 2

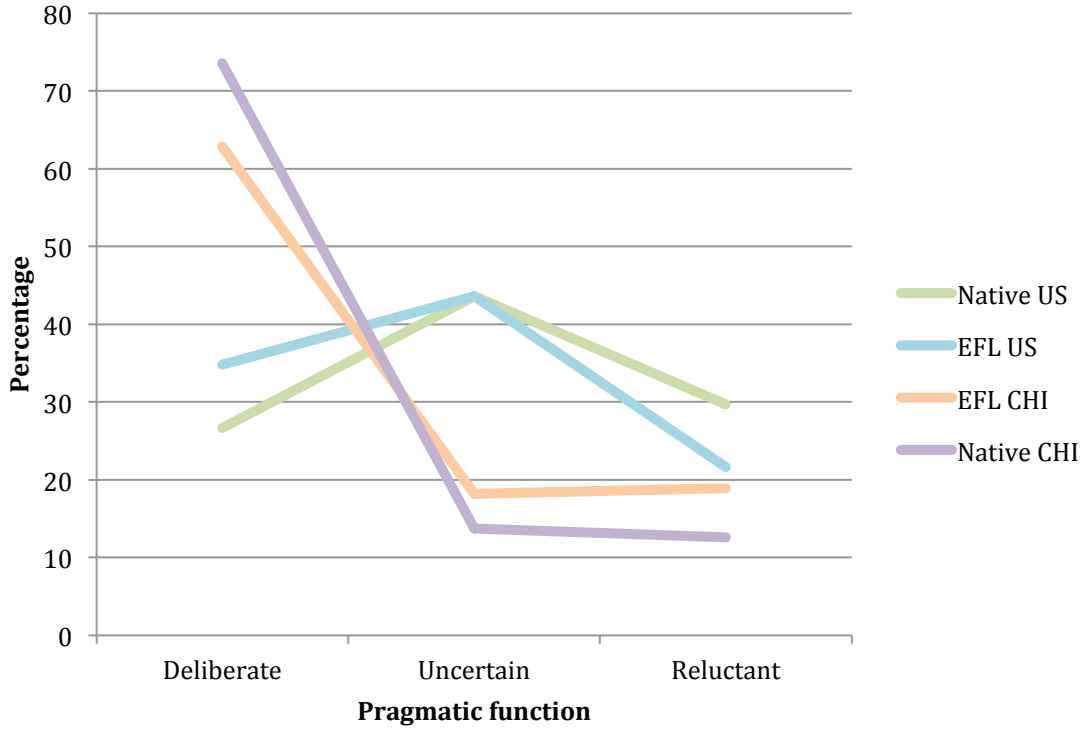


Table 7.14: Post-hoc paired t-tests for % pragmatic function to total utterances for prompt set 2

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
Native US to EFL US	Deliberate	-8.16	.64	5.01
	Uncertain	.03	1.00	5.49
	Reluctant	8.13	.22	3.85
Native US to EFL CHI	**Deliberate**	**-36.22**	**<.001**	**5.02**
	Uncertain	**25.43**	**<.001**	**5.51**
	Reluctant	**10.79**	**.04**	**3.86**
Native US to Native CHI	**Deliberate**	**-46.99**	**<.001**	**4.89**
	Uncertain	**29.86**	**<.001**	**5.36**
	Reluctant	**17.13**	**<.001**	**3.76**
EFL US to EFL CHI	**Deliberate**	**-28.06**	**<.001**	**5.40**
	Uncertain	**25.40**	**<.001**	**5.93**
	Reluctant	2.66	1.00	4.15
EFL US to Native CHI	**Deliberate**	**-38.83**	**<.001**	**5.09**
	Uncertain	**29.83**	**<.001**	**5.58**
	Reluctant	9.00	.14	3.91
EFL CHI to Native CHI	Deliberate	-10.77	.20	4.96
	Uncertain	4.43	1.00	5.44
	Reluctant	6.34	.60	3.81

For this Opinion of others prompt set, trends for all three conditions match the trends for these conditions in Analysis 1 (Percentage construct to total utterances). Interestingly, the trends for the first prompt set (Self-opinion) vs. second prompt set (Opinion of others) appear to have an inverse relation, with US sociolinguistic background associated with higher confidence of [I + verb] usage in the first prompt set but lower confidence in the second prompt set, and Chinese sociolinguistic background associated with lower confidence in the first prompt set and higher confidence in the second.

In sum, there were no significant mean differences in frequency of [I + verb] at any levels between the US-residing groups. Once again, the CHI-residing groups' performances show no significant mean differences as predicted by their common sociolinguistic background. All other between-group mean differences appeared to be significant, except Reluctant condition between EFL US vs. CHI-residing groups.

7.2.3. ANALYSIS 3 - Calculated CR of speaker confidence

Confidence rating (CR) values were calculated for each participant based on CR values of prosodic-syntactic variations from Proof of Concept results. Each prosodic-syntactic variation was multiplied by the percentage of total constructs comprised of that particular prosodic variation for each participant, resulting in an assigned average CR for speaker confidence for each participant. Results are presented below for both prompt sets of the open production task. Descriptive statistics are given first, followed by presentation of inferential statistics.

Part 1: Self-opinion

Means for CR of each group are presented below in Table 7.15.

Table 7.15: Means and SDs for CR of speaker confidence for Phase 2 prompt set 1

Group	Mean	SD
Native US	5.30	.33
EFL US	4.93	.23
EFL CHI	4.43	.25

Native CHI	4.36	.34
Total	4.75	.49

Shapiro-Wilk testing indicated normality of data for all groups (see Table 7.16). Additionally, no data points were identified outside of 3 standard deviations from the mean, and thereby there should be no outliers affecting normality. Thus, given that data distribution overall appeared normal, parametric testing was conducted.

Table 7.16: Shapiro-Wilk results for average CR of speaker confidence, prompt set 1

Group	Shapiro-Wilk statistic	Significance
Native US	**.95**	**.22**
EFL US	**.95**	**.26**
EFL CHI	**.97**	**.68**
Native CHI	**.98**	**.77**

Results denoting normality of data distribution indicated in bold.

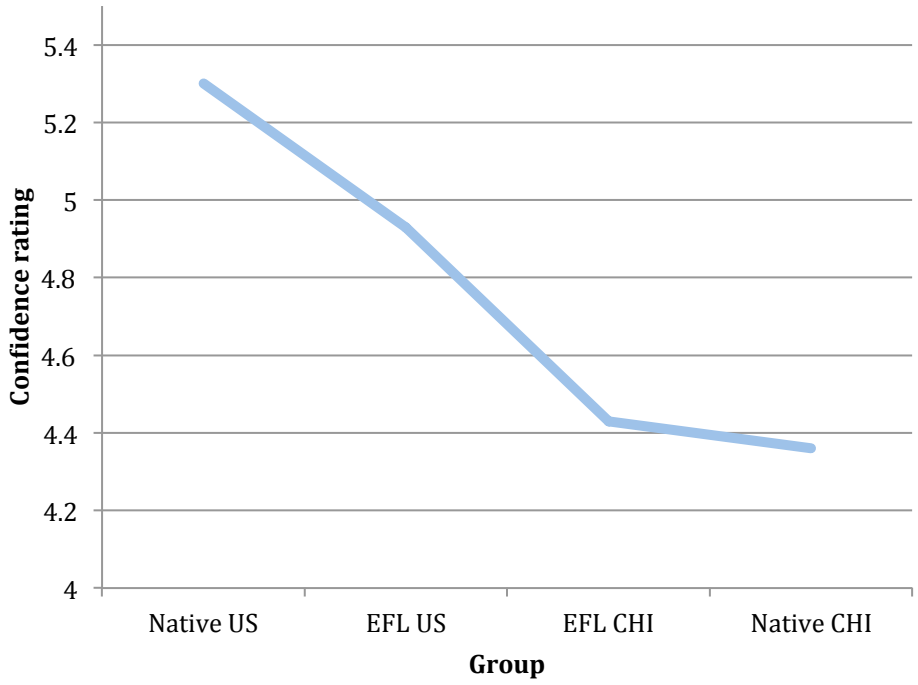
Levene's test suggested homogeneity of variance ($p = .09$), indicating that ANCOVA may be conducted. To investigate the effect of sociolinguistic background on CR values of speaker confidence, a one-way ANCOVA was run with one between-group factor (sociolinguistic background) measured at two levels (US, CHI), with age as a covariate, and with average CR as the dependent variable.

ANCOVA results revealed a large significant between-group effect of sociolinguistic background ($F(1,3) = 54.28$, $p < .001$, $\eta_p^2 = .63$). The covariate of age did not appear to have a statistically significant effect ($F(1,1) = .20$, $p = .65$, $\eta_p^2 = < .01$).

To further investigate the effect of sociolinguistic background, post-hoc paired t-tests

were conducted. Results showed significant mean differences between all groups ($p < .001$) except between EFL CHI/Native CHI groups ($p = 1.00$). As predicted, CR increased with more US sociolinguistic background and decreases with more CHI sociolinguistic background; a high-to-low CR trend appears from US sociolinguistic to CHI sociolinguistic background. The lack of a significant mean difference between EFL CHI and Native CHI groups is also coherent with this thesis' predictions, given that participants of both groups reside in China and thus possess Chinese sociolinguistic background. This trend may be observed in Figure 7.5.

Figure 7.5: Mean CR of speaker confidence for Phase 2 prompt set 1



Part 2: Opinion of others

Means for CR of each group are presented below in Table 7.17.

Table 7.17: Means and SDs for CR of speaker confidence

Group	Mean	SD
Native US	4.70	.32
EFL US	4.90	.25
EFL CHI	5.31	.37
Native CHI	5.61	.34
Total	5.13	.48

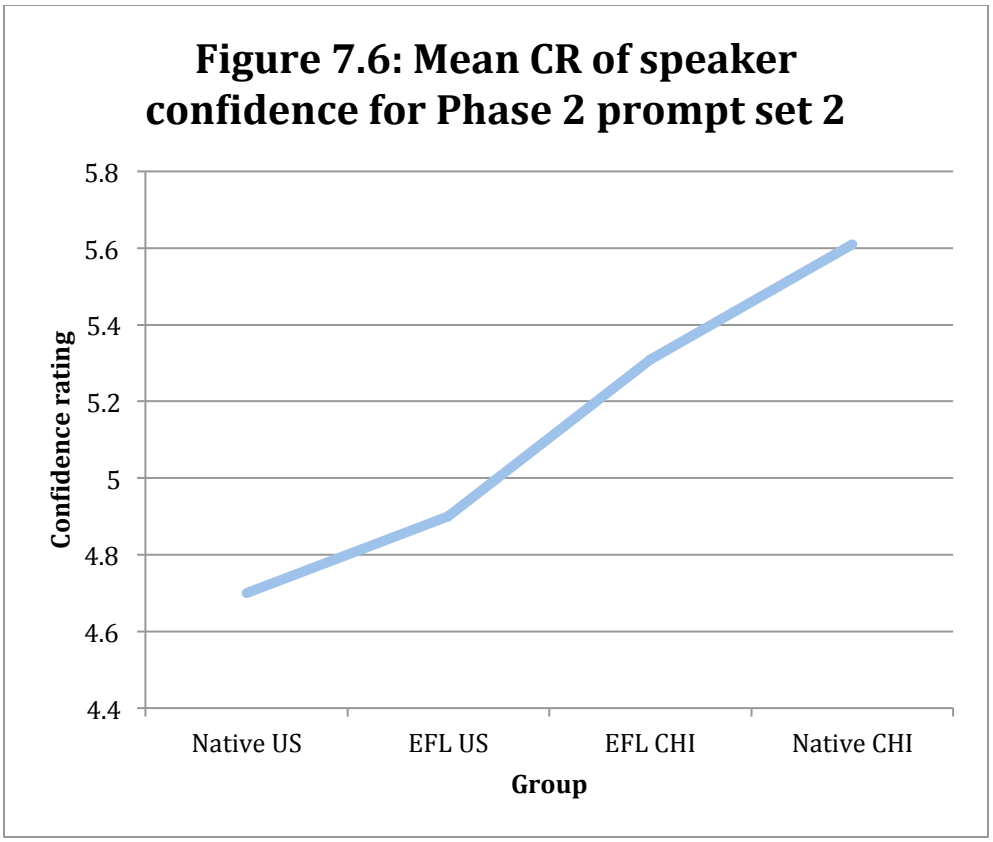
Shapiro-Wilk testing indicated normality of data for all groups (see Table 7.18). Additionally, no data points were identified outside of 3 standard deviations from the mean, and thereby there should be no outliers affecting normality. Thus, given that data distribution overall appeared normal, parametric testing was conducted.

Table 7.18: Shapiro-Wilk results for average CR of speaker confidence for Phase 2 prompt set 2

Group	Shapiro-Wilk statistic	Significance
Native US	**.95**	**.30**
EFL US	**.97**	**.58**
EFL CHI	**.94**	**.11**
Native CHI	**.98**	**.74**

Results denoting normality of data distribution indicated in bold.

Levene's test suggested homogeneity of variance ($p = .32$), indicating that ANCOVA may be conducted. ANCOVA results revealed a large significant between-group effect of sociolinguistic background ($F(1,3) = 36.92$, $p < .001$, $\eta_p^2 = .53$). The covariate of age did not appear to have a statistically significant effect ($F(1,1) = .10$, $p = .75$, $\eta_p^2 = < .01$). To further investigate the effect of sociolinguistic background, post-hoc paired t-tests were conducted. Results showed significant mean differences between all groups ($p < .01$ between groups EFL US and EFL CHI groups, $p = .01$ between EFL CHI and Native CHI groups, $p < .001$ between rest) except between Native US and EFL US groups ($p = .20$). As predicted, CR increased with more US sociolinguistic background and decreases with more CHI sociolinguistic background; a high-to-low CR trend appears from US sociolinguistic background to CHI sociolinguistic background. The lack of a significant mean difference between EFL US and Native US groups is also coherent with this thesis' predictions, given that participants of both groups reside in America and thus possess US sociolinguistic background, with performance of EFL US individuals appearing to conform to the performance of Native US speakers. This trend may be observed in Figure 7.6.



7.2.4. Summary of trends and significance

Table 7.19 below outlines the significant variables and corresponding trends for each ANCOVA of the main analysis of the production task. Bold font indicates same results between ANCOVA 1 and corresponding analysis of ANCOVA 2. Note that where interactional effects occur, significance of related direct effects are consequently irrelevant and therefore not listed.

Table 7.19: ANCOVA trends and significant effects from Phase 2 main phase testing

Analysis	Prompt set	Trend	Significant effects
ANCOVA 1: % pragmatic function of [I + verb] to total utterances produced	**Part 1: Self-opinion**	**Perceived confidence increases for D and U and decreases for R as US sociolinguistic background increases.**	**Significant interactional effect between pragmatic function and sociolinguistic background.**
	Part 2: Opinion of others	**Perceived confidence decreases for D and U as US linguistic/cultural influence increases; no significant trend for R.**	**Significant interactional effect between pragmatic function and sociolinguistic background.**
ANCOVA 2: % pragmatic function of [I + verb] construct to total constructs used	**Part 1: Self-opinion**	**Perceived confidence increases for D and U and decreases for R as US linguistic/cultural influence increases.**	**Significant interactional effect between pragmatic function and sociolinguistic background.**
	Part 2: opinion of others	**Perceived confidence decreases for D and U as US linguistic/cultural influence increases; no significant trend for R.**	**Significant interactional effect between pragmatic function and sociolinguistic background.**
ANCOVA 3: Calculated CR of speaker confidence	Part 1: Self-opinion	CR trends from high-to-low from Native US group, to EFL US group, to EFL CHI/Native CHI groups.	Significant between-subject effect of sociolinguistic background.
	Part 2: opinion of others	CR trends from high-to-low from Native CHI group, to EFL CHI group, to EFL US/Native US groups.	Significant between-subject effect of sociolinguistic background.

As expected, results showed an overall relationship between ANCOVA 1 and ANCOVA 2, with the same trends and significant effects in each ANCOVA for both Self-opinion and Opinion of others prompt sets. Thus given this parallel of findings, further sub-analyses will be conducted on ANCOVA 1 for clarity of one set of statistics and avoidance of redundancy.

For ANCOVA 3, the calculated CR of speaker confidence coincides with the trends exhibited in ANCOVA 1 and 2. In sum, it appears that findings follow predictions of more confident [I + verb] usage being associated with higher levels of US sociolinguistic background and less confident [I + verb] usage being associated with higher levels of CHI sociolinguistic background for the Self-opinion prompt set, whilst for the Opinion of others prompt set more confident [I + verb] was associated with higher levels of CHI sociolinguistic background and less confident [I + verb] being associated with higher levels of US sociolinguistic background. These trends will be further discussed in the next chapter.

7.3. Comparison of Self-opinion vs. Opinion of others prompt set

In order to identify whether frequency of functional [I + verb] usage varies across groups regarding Self-opinion vs. Opinion of others (RQ 3), a dependent samples t-test was conducted for each of the four groups, comparing each participant's Deliberate, Uncertain, and Reluctant [I + verb] usage frequency between Self-opinion and Opinion of others prompt sets. These comparisons were conduced for each dependent variable in the analyses above (Percentage of pragmatic function of [I +verb] to total utterances,

Percentage of pragmatic function of [I +verb] to total constructs, Calculated confidence rating). Patterns of mean differences may be viewed in Figures 7.7 and 7.8 below; paired t-test results are presented in Table 7.19, with dependent variables being grouped in parallel with the above analyses.

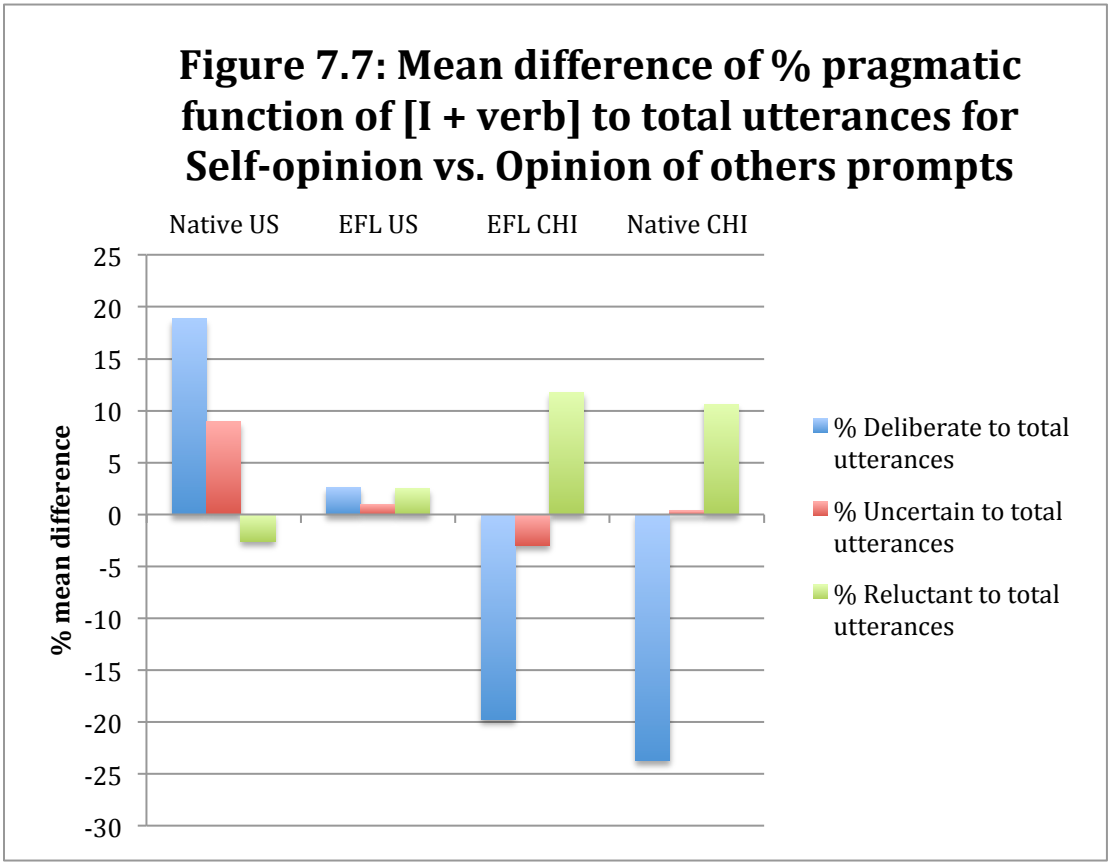
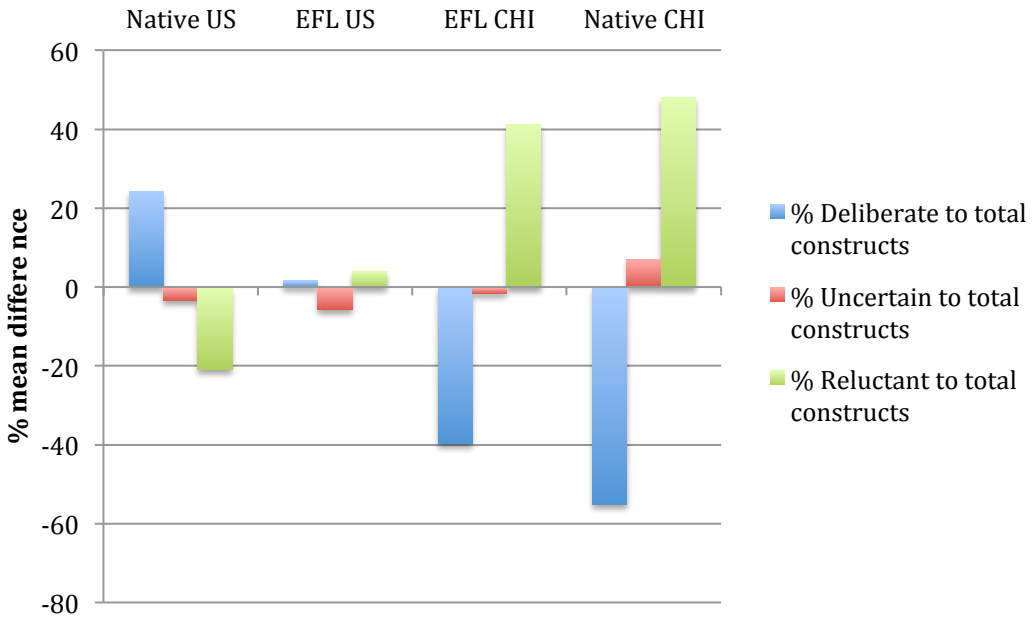


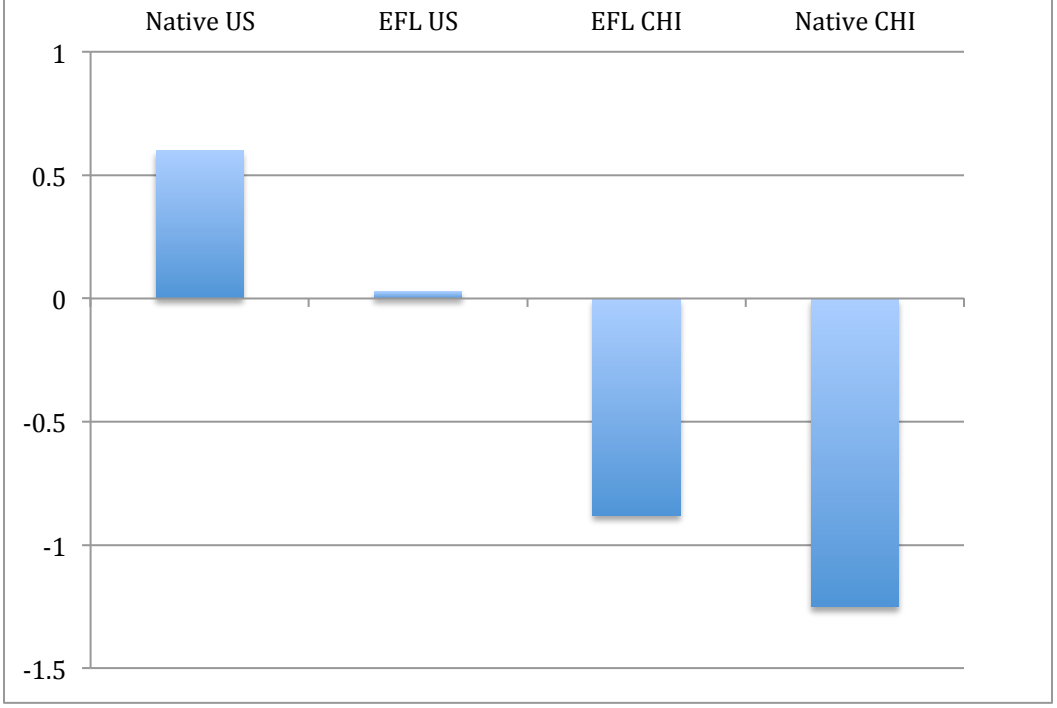
Figure 7.8: Mean difference of % pragmatic function of [I + verb] to total constructs for Self-opinion vs. Opinion of others prompts



Inspecting the above figures, there appeared a congruent trend between dependent variables (Percentage of pragmatic function of [I +verb] to total utterances produced vs. Percentage of pragmatic function of [I +verb] to total constructs used) within the majority of conditions: Mean difference of Percentage Deliberate condition was highest for Native US group and declines with increase of Chinese sociolinguistic background, being lowest for Native CHI; conversely, Mean difference for Percentage Reluctant condition was lowest for Native US group and increases with increase of Chinese sociolinguistic background, being highest for Native CHI. Mean difference for Percentage Uncertain

condition did not have as large a mean difference as these other two conditions nor as observable a trend. This will be further investigated in Table 7.20's results below; Figure 7.8 examines the mean difference of calculated confidence rating between Self-opinion and Opinion of others prompt sets.

Figure 7.8: Mean difference of Calculated confidence rating for Self-opinion vs. Opinion of others prompt sets



In terms of mean difference of Calculated CR between prompt sets, there once again appears an observable trend: Mean difference of Calculated CR is highest for Native US group, decreasing with the increase of Chinese sociolinguistic background so that mean difference is lowest for Native CHI group.

Table 7.20: Paired t-test results between Self-opinion vs. Opinion of others prompt set

		Native US	EFL US	EFL CHI	Native CHI
% Deliberate to total utterances	Mean difference (SD)	**19.88 (10.55)**	**2.62 (6.17)**	**-19.71 (8.77)**	**-23.72 (9.62)**
	T(1,24) value, significance	**9.42, p < .001**	**2.12, p = .04**	**-11.23, p < .001**	**-12.33, p < .001**
% Uncertain to total utterances	Mean difference (SD)	**8.95 (10.86)**	1.01 (8.21)	**-2.99 (6.25)**	.36 (4.54)
	T(1,24) value, significance	**4.12, p < .001**	.62, p = .54	**-2.40, p = .03**	.40, p = .69
% Reluctant to total utterances	Mean difference (SD)	**-2.60 (4.26)**	**2.50 (6.02)**	**11.77 (7.49)**	**10.60 (7.32)**
	T(1,24) value, significance	**-3.05, p = .01**	**2.08, p = .05**	**7.86, p < .001**	**7.24, p < .001**

		Native US	EFL US	EFL CHI	Native CHI
% Deliberate to total constructs	Mean difference (SD)	**24.13 (25.16)**	1.63 (12.26)	**-39.63 (16.59)**	**-55.09 (16.87)**
	T(1,24) value, significance	**4.80, p < .001**	.67, p = .51	**-11.95, p < .001**	**-16.32, p < .001**
% Uncertain to total constructs	Mean difference (SD)	-3.34 (28.84)	-5.69 (18.84)	-1.57 (11.61)	**7.05 (15.44)**
	T(1,24) value, significance	-.58, p = .57	-1.51, p = .14	-.68, p = .51	**2.28, p = .03**
% Reluctant to total constructs	Mean difference (SD)	**-20.80 (17.10)**	4.05 (15.40)	**41.20 (14.96)**	**48.04 (13.17)**
	T(1,24) value, significance	**-6.08, p < .001**	1.32, p = .20	**13.77, p < .001**	**18.24, p < .001**

		Native US	EFL US	EFL CHI	Native CHI
Calculated CR	Mean difference (SD)	**.60 (.42)**	.03 (.25)	**-.88 (.35)**	**-1.25 (.37)**
	T(1,24) value, significance	**7.13, p < .001**	.58, p = .57	**-12.65, p < .001**	**-16.91, p < .001**

Significant mean differences are presented in bold.

Paired t-test results highlight patterns observed in the figures above: For the dependent variable of Percentage of pragmatic function of [I +verb] to total utterances, there were significant mean differences between prompt sets between all groups at Deliberate and Reluctant conditions. Similarly in an almost identical result, results for the dependent variable of Percentage of pragmatic function of [I +verb] to total constructs used showed significant mean differences between prompt sets at Deliberate and Reluctant conditions for all groups except EFL US. Mean differences at Uncertain condition, which did not exhibit as clear a trend as in figures above, were significant for Native US and EFL CHI groups for Percentage of pragmatic function of [I +verb] to total utterances, and significant for Native CHI group for Percentage of pragmatic function of [I +verb] to total constructs used.

In terms of Calculated CR, Mean differences between prompt sets were found between all groups except EFL US. However, it is interesting to note that while EFL US did not show a significant mean difference, it is roughly half way between the positive mean difference of .60 for Native US group on one hand and the negative mean difference of -.88 for EFL CHI group on the other hand.

To summarize, it appears that there is a significant mean difference between prompt sets for all groups at Deliberate and Reluctant conditions, with increased US sociolinguistic background leading to more Deliberate [I + verb] usage in the Self-opinion prompt set, and more Reluctant [I + verb] usage in the Opinion of other prompt set. Conversely, increased CHI sociolinguistic background may lead to more Reluctant [I + verb] usage in the Self-opinion prompt set, and more Deliberate [I + verb] usage in Opinion of other prompt sets. These findings were congruent with Mean difference of

Calculated CR results, which showed that increased US sociolinguistic background was associated with higher Calculated CRs for Self-opinion prompts and lower Calculated CRs for Opinion of others prompts, whilst increased CHI sociolinguistic background on the other hand was associated with lower Calculated CRs for Self-opinion prompts but higher Calculated CRs for Opinion-of-other prompts.

7.4. Effect of L2 proficiency on frequency of [I + verb]:

An in-depth sub-analysis was carried out on the two EFL groups to identify whether frequency of DM [I + verb] usage varied as a function of L2 proficiency, as identified by the Lex30 score (RQ 6). Because RQ 6 only concerns itself with DM variation (Reluctant condition) of [I + verb] (see section 3.1.3 for overview of relevant literature), this in-depth sub-analysis involves the DV of Reluctant condition only. However, in the interest of research, an analysis conducted on all [I + verb] conditions is included in Appendix 16 for reference, as this analysis was somewhat tangential to addressing the research questions.

For Analysis 1 above, the same analysis was run but with EFL US and EFL CHI participants only (N = 50) and with Reluctant condition of each prompt set as DVs, subsequently followed by a similar analysis with the addition of the between-subject factor of Lex30 score, as well as the between-subject factors of Age of L2 acquisition and Native language of L2 instructor(s)[51]. The first N = 50 analysis serves to act as a

[51] As previously mentioned in Section 4.3.4, Language spoken at school was also considered as a between-subject factor. However, lack of variance (N = 47 Only Mandarin, N = 1 English, N = 2 Both Mandarin/English) meant a comparison could not be successfully conducted. Thus, this variable was

comparison model for the second analysis, which acts to answer RQ 6 as well as to control for potential influence from such variables.

To begin, this section presents the descriptives for these between-subject factors. Following these, the outcomes of inferential statistic analyses are provided. Results for the first N = 50 analysis (no covariates added) is subsequently presented. Following are results from the second N = 50 analysis (with addition of between-subject factors), with comparisons to the first N = 50 analysis included.

7.4.1. Descriptives for EFL-specific between-subject factors

LEX30: Means, SDs, and Minimum/Maximum values for Lex30 vocabulary score are presented below in Table 7.21. Shapiro-Wilk testing indicated normality of distribution for both EFL US (p = .75) and EFL CHI (p = .25) groups. Additionally, no outliers were present.

Table 7.21: Descriptives for Lex30 vocabulary score

Group	Mean	SD	Min - Max
EFL US	76.52	12.50	51 - 98
EFL CHI	79.24	9.06	64 - 94

dropped from analysis. Similarly, Preferred language for socialization was also measured, but once again lack of variance within each group (N = 2 English, N = 23 Mandarin, N = 0 Both Mandarin/English, for EFL US group; N = 0 English, N = 23 Mandarin, N = 2 Both Mandarin/English) meant a comparison could not be successfully conducted and subsequently resulted in exclusion from analysis as well.

Native Language of L2 Instructor: Nominal counts per group are presented in Table 7.22 below.

Table 7.22: Descriptives for Native Language of L2 Instructor

Group	Teacher's native language	N
EFL US	US/Non-CHI	3
	Only CHI	22
EFL CHI	US/Non-CHI	12
	Only CHI	13

Age of L2 acquisition: Means, SDs, and Minimum/Maximum values for Age of L2 acquisition are presented below in Table 7.23. Shapiro-Wilk testing indicated normality of distribution for both EFL US ($p = .26$) and EFL CHI ($p = .22$) groups. While analysis did flag one potential outlier in EFL US group, removing this data point only led to more potential outliers. Moreover, this data point was less than 2 SDs away from the mean and thereby under the 3 SD threshold for identifying outliers affecting normality. Thus, this case was deemed not be an outlier and was included in analysis.

Table 7.23: Descriptives for Age of L2 Acquisition

Group	Mean	SD	Min - Max
EFL US	12.52	4.09	4 - 20
EFL CHI	10.40	3.27	3 - 16

7.4.2. ANALYSIS 1a - Percentage of pragmatic function of [I + verb] to total utterances produced for N = 50

Results are presented below with the same repeated measures ANCOVAs serving to include both prompt sets of the open production task. For each prompt set, findings from the sub-analysis N = 50 results (with no added covariates) are first presented, followed by a comparison with N = 50 ANCOVA with inclusion of covariates.

ORIGINAL MODEL: To act as a comparison model for the second set (which serves to answer RQ 6 and control for external influences), a repeated measures ANCOVA of N = 50 was run with one within-group factor (pragmatic function) measured at two levels (Reluctant condition for Self-opinion, Reluctant condition for Opinion of others) and one between-group factor (sociolinguistic background) measured at two levels (EFL US, EFL CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

Because there were only two levels of repeated measures, sphericity was not calculated. However, Levene's test showed homogeneity of variance at all levels ($p = .10$ for Reluctant condition of Self-opinion, $p = .24$ for Reluctant condition of Opinion of others, and parametric testing thereby proceeded.

ANCOVA results revealed a significant between-group effect of sociolinguistic background ($F(1,1) = 12.00$, $p = .001$, $\eta_p^2 = .20$, OP = .92). Additionally, there was a significant within-group effect of pragmatic function ($F(1,1) = 17.32$, $p < .001$, $\eta_p^2 = .27$, OP = .98). Moreover, results also showed a large interactional effect between pragmatic function and sociolinguistic background ($F(1,1) = 7.18$, $p < .001$, $\eta_p^2 = .13$, OP = .75).

There did not appear to be a statistically significant effect of age ($F(1,1) = .11$, $p = .74$, $\eta_p^2 = < .01$, OP = .06), but results showed an interactional effect between pragmatic function and age ($F(1,2) = 4.42$, $p = .04$, $\eta_p^2 = .09$, OP = .54).

Note that, whilst the observed power threshold for reliability is .80 (Cohen, 1988), with the reduced sample size of N = 50 some variables did not reach this threshold (OP = .75 for the interaction between pragmatic function and sociolinguistic background; OP = 74 for the interaction between pragmatic function and age), meaning that these significance values the may not be as trustworthy. Subsequently, an observation of effect size may prove more telling for this N = 50 sub-analysis.

In terms of the interactional effect between pragmatic function and sociolinguistic background, post-hoc paired t-testing showed a significant mean difference of 8.63 between EFL US and EFL CHI groups at Reluctant condition for Self-opinion prompts ($p < .001$) and no significant mean difference for Opinion of others prompts ($p = .19$), coinciding with results of previous analyses.

ADDITION OF COVARIATES: To investigate whether frequency of pragmatic function of [I + verb] usage varies as a function of L2 proficiency (RQ 6) as well as to account for possible confounding variables, a comparable repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at two levels (Reluctant condition for Self-opinion, Reluctant condition for Opinion of others) and one between-group factor (sociolinguistic background) measured at two levels (EFL US, EFL CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

Because there were only two levels of repeated measures, sphericity was not calculated. However, Levene's test showed homogeneity of variance at all levels (p = .23 for Reluctant condition of Self-opinion, p = .12 for Reluctant condition of Opinion of others, and parametric testing thereby proceeded.

Table 7.24 below presents ANCOVA results in terms of significance values (F included if significant), effect size, and observed power. Significant p values with sufficient observed power (> .80) are presented in bold; for all other cases, given the addition of these variables, observed power may be too low for an actual significant effect to exhibit a significant p value (i.e. Type II error), and thus measures of effect size should be taken into consideration instead.

Table 7.24: ANCOVA results of significance values, effect size and observed power

Within-subject effect	Significance	η_p^2	Observed power
Pragmatic function	.04 (F(1, 1) = 4.55)	.10	.55
Pragmatic function * Age	.12	.06	.34
Pragmatic function * Age L2 acquisition	.70	<.01	.07
Pragmatic function * Lex30	.35	.02	.15
Pragmatic function * Sociolinguistic background	.10	.06	.38
Pragmatic function * Native language of L2 instructor	.86	.01	.07
Pragmatic function * sociolinguistic background * Native language of L2 instructor	.22	.04	.23

Between-subject effect	Significance	η_p^2	Observed power
Age	.65	.01	.07
Age L2 acquisition	.63	.01	.08
Lex30 score	.34	.02	.16
Sociolinguistic background	.01 (F(1,1) = 7.15	.05	.74
Native language of L2 instructor	.70	.02	.10
Sociolinguistic background * Native language of L2 instructor	.56	.01	.09

The within-subject effect of pragmatic function was the only variable that showed a significant effect (although low observed power due to additional covariates may have influenced results). Regarding the significant effect of pragmatic function, post-hoc paired t-tests showed a significant mean difference of 7.14 between Reluctant condition for Self-opinion vs. Opinion of others prompt sets ($p < .001$), coinciding with results from section 7.3.

In terms of effect sizes, pragmatic function accounted for the largest percentage of variation at 10%, mirroring results of previous analyses. The variables that accounted for the second highest amount of variance were the interaction between pragmatic function and sociolinguistic background and pragmatic function and age at 6%, followed by the between-subject effect of sociolinguistic background at 5%. All other variables, including Lex30 score and any interactions with it, accounted for less than 5% of total variability each.

Below in Table 7.25 is a comparison of effect sizes between N = 50 original ANCOVA and N = 50 with addition of covariates ANCOVA.

Table 7.25: Comparison of effect sizes between original N = 50 ANCOVA and N = 50 with addition of covariates ANCOVA

Within-subject effect	Original η_p^2	Additional covariates added η_p^2
Pragmatic function	.27	.10
Pragmatic function * Age	.09	.06
Pragmatic function * Sociolinguistic background	.13	.06

Between-subject effect	Original η_p^2	Additional covariates added η_p^2
Age	<.01	.01
Sociolinguistic background	.20	.05

When compared to one another, it appears that the variable explaining the largest amount of total variance, pragmatic function, remained the same, maintaining nearly twice the total amount of variability explained after the addition of covariates. The interactional effects of pragmatic function * age and pragmatic function * sociolinguistic background were both decreased to 5%, with the biggest change being the effect of sociolinguistic background decreasing by quadruple fold from 20% to 5% with the addition of covariates added to analysis. Such decrease in effect sizes were to be expected given the addition of extra variables into the second ANCOVA. Additionally, effect size of age remained low, accounting for the smallest amount of total variance in both analyses.

In summary, it appears that for N = 50 ANCOVA, frequency of DM [I + verb] usage (as determined by percentage of pragmatic [I + verb] construct out of total utterances) variec mainly as a factor of the pragmatic function, with not much variation being a function of L2 proficiency or other confounding variables analyzed.

7.5. Effect of Time immersed in new sociolinguistic environment on frequency of [I + verb]

A further in-depth sub-analysis was carried out on the EFL US group to identify whether frequency of functional [I + verb] usage varied as a function of Time immersed in a new sociolinguistic environment (RQ 4). For Analysis 1 above, the same analysis was run but with additional covariates of Lex30 score, Age of L2 acquisition, and Native language of L2 instructor and with EFL US participants only (N = 25, meaning that the between-subject factor of sociolinguistic environment was removed due to there only being one group involved); this was subsequently followed by a similar analysis with the addition of the between-subject factor of Time immersed in new sociolinguistic environment. The first N = 25 analysis serves to act as a comparison model for the second analysis, which acts to answer RQ 4 as well as to control for potential confounds.

Firstly, descriptives for the EFL US group-specific within-subject factor of Time immersed in new sociolinguistic environment are presented. Next, drawing additionally from the descriptive statistics in section 7.4.1. above, inferential statistics are given below per analysis. The first set of analyses was identical to Analysis 1a with additional covariates above, but with N = 25, and are presented first with differences in relation to the same analysis but with N = 50 (refer to Appendix 16) being subsequently reported.

Following are results from the second N = 25 analysis (with addition of between-subject factors), with comparisons to the first N = 25 analysis included. Results are subsequently presented for each prompt set of the production task.

7.5.1. Descriptives for EFL US-specific between-subject factors

Time immersed in new sociolinguistic environment: Means, SDs, and Minimum/Maximum values for lex30 vocabulary score in Table 7.26. While Shapiro-Wilk testing did not necessarily indicate normality of distribution ($p = .01$), examination of the histogram showed only slight negative skew. Additionally, no outliers were present.

Table 7.26: Descriptives for Time immersed in new sociolinguistic environment

Group	Mean (years)	SD	Min - Max
EFL US	18.72	7.23	5 - 30

7.5.2. ANALYSIS 1b - Percentage of pragmatic function of [I + verb] to total utterances produced for N = 25

In order to address RQ 4, the following sub-analysis was run on the EFL US group (who immigrated from their Native L1 China to the US) to investigate the effect of Time immersed in a new sociolinguistic environment. Results are presented below for both prompt sets of the open production task (Self-opinion, then Opinion of others). For each prompt set, differences between the N = 50 ANCOVA with additional covariates and

sub-analysis N = 25 results are first presented, followed by a comparison with N = 25 ANCOVA with inclusion of between-subject factor of Time immersed in new sociolinguistic environment.

Part 1: Self-opinion

ORIGINAL MODEL: To act as a comparison model for the second set (which serves to answer RQ3b) as well as to control for confounds, a repeated measures ANCOVA was first run at N = 50 with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor of Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with Age, Age of L2 acquisition, and Lex30 score as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced (same as Analysis 1a above). While sphericity cannot necessarily be assumed (X^2 (1, 2) = 13.19, p < .01), results calculated with Mauchly's sphericity vs. Greenhouse-Geisser measurements (.67) do not differ regarding significance. Levene's test indicated homogeneity of variance at all levels (p = .59 Deliberate; p = .69 Uncertain, p = .50 Reluctant).

ANCOVA results revealed no significant within-group effect of pragmatic function (F(1,2) = 2.78, p = .10, η_p^2 = .12, OP =.41). Additionally, there did not appear to be a statistically significant between-subjects effect of age (F(1,1) = 1.08, p =.31, η_p^2 = .05, OP = .17), Lex30 score (F(1,1) = .06, p = .80, η_p^2 < .01, OP = .06), Age of L2 acquisition (F(1,1) = 1.32, p = .26, η_p^2 = .06, OP = .20), or Native language of L2 instructor (F1,1) = .04, p = .85, η_p^2 < .01, OP = .05). Moreover, results also showed no interactional effects between pragmatic function and Lex30 score (F (1,1.33) = 3.21, p = .07, η_p^2 = .14, OP =

.47), pragmatic function and Age of L2 acquisition (F (1,1.33) = 1.68, p = .37, η_p^2 = .04, OP = .17), pragmatic function and Native language of L2 instructor (F (1,1.33) = 1.79, p = .42, η_p^2 = .04, OP = .15), or pragmatic function and age (F (1,1.33) = .99, p = .35, η_p^2 = .05, OP = .18). This differs from the N= 50 ANCOVA only in that no significant effect of age was found. For all cases, given the addition of these variables, observed power may be too low for an actual significant effect to exhibit a significant p value, and thus effect size should be taken into consideration instead. Table 7.27 below compares effect sizes of N = 25 to the effect sizes of the same analysis run with N = 50.

Table 7.27: Effect size results of N = 25 ANCOVA vs. N = 50 ANCOVA

Within-subject effect	N = 50 η_p^2	N = 25 η_p^2
Pragmatic function	.02	.12
Pragmatic function * Age	.05	.05
Pragmatic function * Age L2 acquisition	.01	.04
Pragmatic function * lex30	.03	.14
Pragmatic function * Native language of L2 instructor	.01	.04

Between-subject effect	N = 50 η_p^2	N = 25 η_p^2
Age	.01	.05
Age L2 acquisition	.01	.06
Lex30 score	.01	< .01
Native language of L2 instructor	.001	< .01

Interestingly for the N = 25 ANCOVA, the variable that accounted for most variance was Lex30 score at 14%, vs. merely 3% for the comparable N = 50 ANCOVA. The variable that explained the second most variance was pragmatic function, again rising from 2% to 12%. This was followed by Age of L2 acquisition at 6%, then age and pragmatic function * age at 5%. Direct effects of Lex30 score and Native language of L2 instructor did not account for nearly any variance at <1%, although their interactional effects with pragmatic function did account for 14% and 4% respectively.

ADDITION OF TIME IMMERSED IN NEW SOCIOLINGUISTIC ENVIRONMENT: To investigate whether frequency of pragmatic function of [I + verb] usage varies as a function of Time immersed in a new sociolinguistic environment, a comparable repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factors of Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with age, Age of L2 acquisition, Lex30 score, and Time immersed in new sociolinguistic environment as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 12.72, p < .01), results calculated with Mauchly's sphericity vs. Greenhouse-Geisser measurements (.66) do not differ regarding significance. Levene's test indicated homogeneity of variance at all levels (p = .52 Deliberate; p = .60 Uncertain, p = .73 Reluctant), mirroring results of the N = 25 ANCOVA above.

Table 7.28 below presents ANCOVA results in terms of significance values (F included if significant), effect size, and observed power. For all cases, given the addition of these variables, observed power may be too low for an actual significant effect to exhibit a significant p value, and thus effect size should be taken into consideration instead.

Table 7.28: ANCOVA results of significance values, effect size and observed power

Within-subject effect	Significance	η_p^2	Observed power
Pragmatic function	.24	.07	.25
Pragmatic function * Age	.57	.02	.10
Pragmatic function * Age L2 acquisition	.40	.04	.16
Pragmatic function * lex30	.12	.11	.37
Pragmatic function * Native language of L2 instructor	.51	.03	.12
Pragmatic function * Time immersed in new sociolinguistic environment	.88	<.01	.06

Between-subject effect	Significance	η_p^2	Observed power
Age	.60	.02	.08
Age L2 acquisition	.30	.06	.18
Lex30 score	.89	< .01	.05
Native language of L2 instructor	.78	< .01	.06
Time immersed in new sociolinguistic environment	.82	<.01	.06

For this N = 25 ANCOVA with addition of Time immersed in new sociolinguistic environment, the variable that accounted for most variance was the interaction between Lex30 score and pragmatic function at 11%. This was followed by pragmatic function at 7%, then Age of L2 acquisition at 6%. Note that Time immersed in a new sociolinguistic environment, both as a direct effect and an interactional effect with pragmatic function, seemed to account for less than 1% of total variance.

Table 7.29 below compares effect sizes of N = 25 with addition of Time immersed in new sociolinguistic environment to the effect sizes of the same N = 25 analysis without inclusion of this factor.

Table 7.29: Effect size results N = 25 ANCOVA without vs. with additional covariate

Within-subject effect	N = 25 η_p^2	N = 25 with added factor η_p^2
Pragmatic function	.12	.07
Pragmatic function * Age	.05	.02
Pragmatic function * Age L2 acquisition	.04	.04
Pragmatic function * lex30	.14	.11
Pragmatic function * Native language of L2 instructor	.04	.03

Between-subject effect	N = 25 η_p^2	N = 25 with added factor η_p^2
Age	.05	.02
Age L2 acquisition	.06	.05
Lex30 score	< .01	.01
Native language of L2 instructor	< .01	.03

From this comparison, it appears that Time immersed in new sociolinguistic environment, which explains less than 1% of variance, lowered the interactional effect and direct effect of age by 1.5 times the original effect size, as well as lowering effect of pragmatic function by almost twofold. Thus, results suggest that in answer to RQ 4, frequency of functional [I + verb] usage did not vary as a function Time immersed in a new sociolinguistic environment. Furthermore, Lex30 score appeared to explain a relatively large amount of variance for EFL US group with no other variable accounting for as much of the total variance; however, while frequency of [I + verb] seemed to vary as a function of L2 proficiency for EFL US group, given results from Analysis 1a above it does not appear to influence frequency count for EFL CHI group. Additionally, age of L2 acquisition may also play more of a role when looking specifically at EFL US group.

Part 2: Opinion of others

ORIGINAL MODEL: To act as a comparison model for the second prompt set (which serves to answer RQ 4) as well as to control for confounds, a repeated measures ANCOVA was run at N = 50 with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor of Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with age, Age of L2 acquisition, and Lex30 score as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced (same as Analysis 1a above).

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 8.75, p = .01), results calculated with Mauchly's sphericity vs. Greenhouse-Geisser measurements (.73) do not

differ regarding significance. Additionally, Levene's test indicated homogeneity of variance at all levels (p = .07 Deliberate; p = .31 Uncertain, p = .14 Reluctant). Thus, parametric testing was carried out.

ANCOVA results revealed no significant within-group effect of pragmatic function ($F(1,2) = 2.32$, $p = .13$, $\eta_p^2 = .10$, OP =.37). Additionally, there did not appear to be a statistically significant between-subjects effect of Age ($F(1,1) = 3.79$, $p = .07$, $\eta_p^2 = .16$, OP = .46), Lex30 score ($F(1,1) = .2.85$, $p = .11$, $\eta_p^2 = .13$, OP = .36), Age of L2 acquisition ($F(1,1) = .17$, $p = .68$, $\eta_p^2 = .01$, OP = .07), or Native language of L2 instructor ($F1,1) = 2.00$, $p = .17$, $\eta_p^2 < .09$ OP = .27). Moreover, results also showed no interactional effects between pragmatic function and Lex30 score ($F (1,1.33) = 1.26$, $p = .29$, $\eta_p^2 = .06$, OP = .22), pragmatic function and Age of L2 acquisition ($F (1,1.33) = 1.62$, $p = .22$, $\eta_p^2 = .08$, OP = .27), pragmatic function and Native language of L2 instructor ($F (1,1.33) = 1.81$, $p = .19$, $\eta_p^2 = .08$, OP = .30), or pragmatic function and Age ($F (1,1.33) = .56$, $p = .52$, $\eta_p^2 = .03$, OP = .12). Significance level results did not differ from the original N= 50 ANCOVA with additional covariates. For all cases, however, observed power may be too low to depend on significance values for such a sample size of 25, and thus effect size should be taken into consideration instead. Table 7.30 below compares effect sizes of N = 25 to the effect sizes of the same analysis run with N = 50.

Table 7.30: Effect size results of N = 25 ANCOVA vs. N = 50 ANCOVA

Within-subject effect	N = 50 η_p^2	N = 25 η_p^2
Pragmatic function	.06	.10
Pragmatic function * Age	.08	.03
Pragmatic function * Age L2 acquisition	< .01	.08
Pragmatic function * lex30	.01	.06
Pragmatic function * Native language of L2 instructor	.02	.08

Between-subject effect	N = 50 η_p^2	N = 25 η_p^2
Age	.02	.16
Age L2 acquisition	<.001	.01
Lex30 score	.01	.13
Native language of L2 instructor	.01	.09

Interestingly for this N = 25 ANCOVA, the variable that accounted for most variance was the direct effect of age at 16%, vs. merely 2% for the comparable N = 50 ANCOVA. The variable that explained the second most variance—almost as much variance as age—was the direct effect of Lex30 score, again rising from only 1% up to 13%. This was followed by pragmatic function at 10% and Native language of L2 instructor at 9%, then interactional effect of pragmatic function * Native language of L2 instructor as well as pragmatic function * age at 8%, which was up from 2% in N = 50 ANCOVA.

ADDITION OF TIME IMMERSED IN NEW SOCIOLINGUISTIC ENVIRONMENT: To investigate whether frequency of pragmatic function of [I + verb] usage varies as a function of Time immersed in a new sociolinguistic environment, a comparable repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factors of Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with Age, Age of L2 acquisition, Lex30 score, and Time immersed in new sociolinguistic environment as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 8.37, p = .02), results calculated with Mauchly's sphericity vs. Greenhouse-Geisser measurements (.73) do not differ regarding significance. Levene's test indicated homogeneity of variance at all levels (p = .08 Deliberate; p = .39 Uncertain, p = .12 Reluctant), mirroring results of the N = 25 ANCOVA above.

Table 7.31 below presents ANCOVA results in terms of significance values (F included if significant), effect size, and observed power. For all cases, observed power may be too low to depend on significance values, and thus effect size should be taken into consideration instead.

Table 7.31: ANCOVA results of significance values, effect size and observed power

Within-subject effect	Significance	η_p^2	Observed power
Pragmatic function	.28	.07	.23
Pragmatic function * Age	.92	< .01	.06
Pragmatic function * Age L2 acquisition	.24	.07	.25
Pragmatic function * lex30	.34	.05	.19
Pragmatic function * Native language of L2 instructor	.41	.04	.16
Pragmatic function * Time immersed in new sociolinguistic environment	.92	<.01	.05

Between-subject effect	Significance	η_p^2	Observed power
Age	.82	< .01	.06
Age L2 acquisition	.75	.01	.06
Lex30 score	.10	.14	.37
Native language of L2 instructor	.49	.03	.10
Time immersed in new sociolinguistic environment	.62	.01	.08

For this Opinion of others ANCOVA of N = 25 with addition of Time immersed in new sociolinguistic environment, the variable that accounted for most variance was Lex30 score at 14%, mirroring results of the comparable Self-opinion ANCOVA above. This was followed by pragmatic function as well as pragmatic function * Age L2 acquisition at 7%. The variance accounted for by the latter is paralleled by a direct effect of Age of L2 acquisition in the previous analysis of Self-opinion prompts. The variable

that accounted for the next largest amount of variance was the interactional effect between pragmatic function and Lex30 score at 5%, resulting in a combined 19% of variance being related to L2 proficiency. Note that Time immersed in sociolinguistic environment, both as a direct effect and an interactional effect with pragmatic function, appeared to account for no more than 1% of total variance respectively.

Table 7.32 below compares effect sizes of N = 25 with addition of Time immersed in new sociolinguistic environment to the effect sizes of the same N = 25 analysis without inclusion of this factor.

Table 7.32: Effect size results of N = 25 ANCOVA vs. N = 50 ANCOVA

Within-subject effect	N = 25 η_p^2	N = 25 with added factor η_p^2
Pragmatic function	.10	.07
Pragmatic function * Age	.03	< .01
Pragmatic function * Age L2 acquisition	.08	.07
Pragmatic function * lex30	.06	.05
Pragmatic function * Native language of L2 instructor	.08	.04

Between-subject effect	N = 25 η_p^2	N = 25 with added factor η_p^2
Age	.16	< .01
Age L2 acquisition	.01	.01
Lex30 score	.13	.14
Native language of L2 instructor	.09	.03

From this comparison, it appears that the addition of Time immersed in new sociolinguistic environment resulted in the largest change in effect size for the factors of age and pragmatic function * age, dropping both to an effect size that explains less than 1% of variance. Additionally, effect size of Native language of L2 instructor decreased by twofold for its interactional effect and by three-fold for its direct effect. Nonetheless, findings from this analysis indicate that in answer to RQ 4, frequency of functional I + verb] usage did not vary as a function Time immersed in a new sociolinguistic environment. Additionally, results suggest that Lex30 may also play more of a role when looking specifically at EFL US group, although given results from sub-analysis 1a above, it appears that this factor does not affect Reluctant condition of [I + verb] and thereby does not appear relevant to RQ 6.

7.6. Expressions of negative appraisal in the Opinion of others prompt set

To address RQ 3a, a chi-squared test of independence was performed to investigate the relation between sociolinguistic background and expression of negative appraisal when asked for an impression of the characters in the Opinion of others prompt set. This analysis was conducted for each of the three stories on the third prompt in the story set, which asked the participant, "What is your impression of [X]?" where X is the abstract fictional character in the story. This test was run with one between-subject factor (sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI) and one within-subject factor (appraisal) measured at two levels (non-negative, negative).

Descriptive statistics are first presented in Table 7.33 (along with expected count as determined by the chi-square test for reference), with inferential results from the Chi-squared test presented below.

Table 7.33: Post-hoc paired chi-squared test results

Story	Group	Appraisal count	
		Negative	Non-negative
Charlotte	Native US	4	21
	EFL US	11	14
	EFL CHI	16	9
	Native CHI	18	7
	Expected Count	12.3	12.8
Gary	Native US	7	18
	EFL US	15	10
	EFL CHI	10	15
	Native CHI	12	13
	Expected Count	11	14
Jimmy	Native US	1	24
	EFL US	1	24
	EFL CHI	2	23
	Native CHI	7	18
	Expected Count	2.8	22.3

Because the "Jimmy" condition did not meet the expected cell count of "5" for every cell, a Fisher's exact test was calculated to determine significance, and is in place of the Pearson Chi-Square significance value when reported. All other conditions met the assumption of an expected cell count of "5" or more.

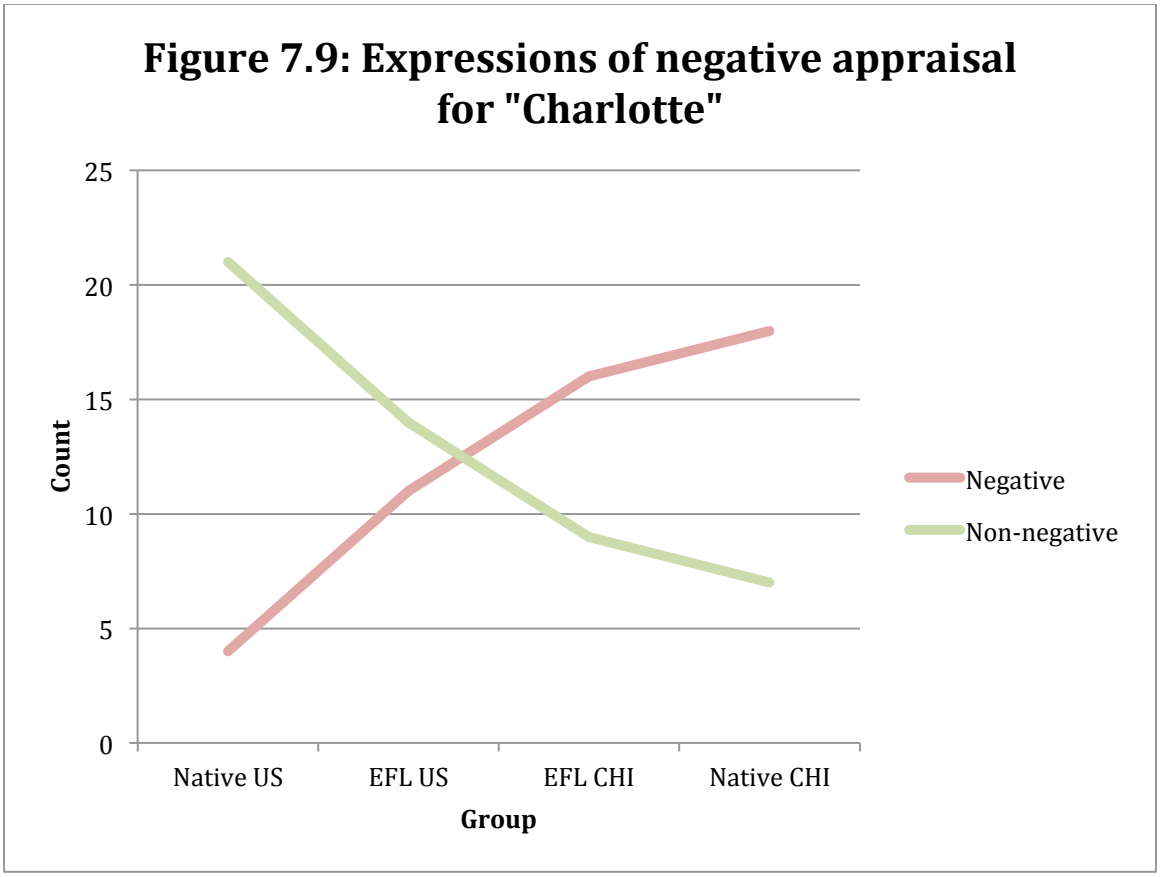
Results showed a statistically significant association between sociolinguistic background and expression of negative appraisal in the Charlotte story (X^2 (3) = 18.69, p < .001) and the Jimmy story (X^2 (3) = 10.11, p = .03), but there was no statistically significant association in the Gary story (X^2 (3) = 5.52, p = .14).

Post-hoc paired chi-squared tests were carried out for the significant results of "Jimmy" and "Charlotte" between each group, resulting in 12 comparisons. Consequently, given the correlated nature of these Chi-square tests, Keppel's Modified Bonferroni correction (Keppel, 1991) to reduce the likelihood of a Type II error whilst still controlling for a Type I error. For these post-hoc chi-squared tests, the significance value was thereby set to .03. Results are presented in Table 7.34. Mean differences are observable in Figures 7.9, 7.10 and 7.11, and will be discussed below.

Table 7.34: Post-hoc paired chi squared tests for expressions of appraisal

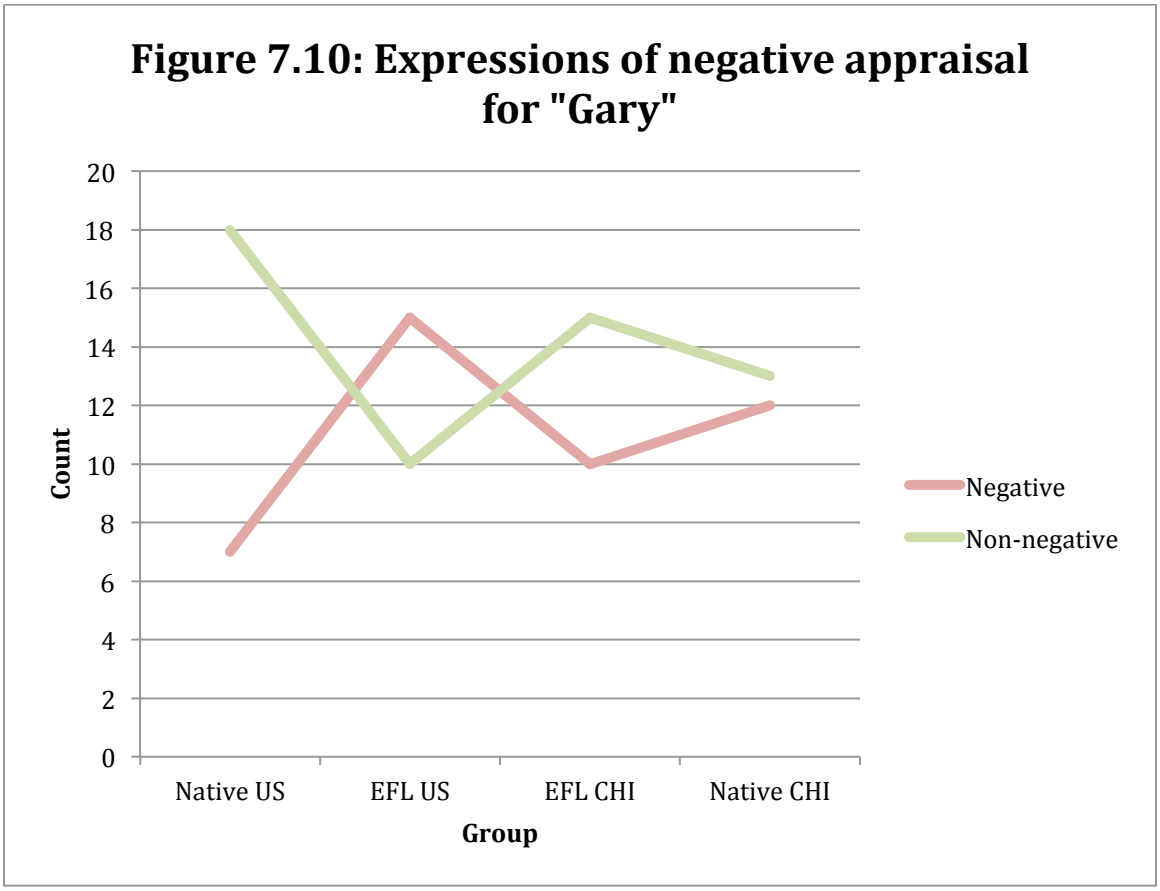
Group comparison	Story	X^2 (1) =	Significance
Native US to EFL US	**Charlotte**	**4.67**	**.03**
	Jimmy	0	.76
Native US to EFL CHI	**Charlotte**	**12.00**	**.001**
	Jimmy	.36	.50
Native US to Native CHI	**Charlotte**	**15.91**	**<.001**
	Jimmy	**5.36**	**.02**
EFL US to EFL CHI	Charlotte	2.01	.16
	Jimmy	.36	.50
EFL US to Native CHI	Charlotte	4.02	.05
	Jimmy	**5.36**	**.02**
EFL CHI to Native CHI	Charlotte	.37	.54
	Jimmy	3.39	.07

Significant results are indicated in bold.

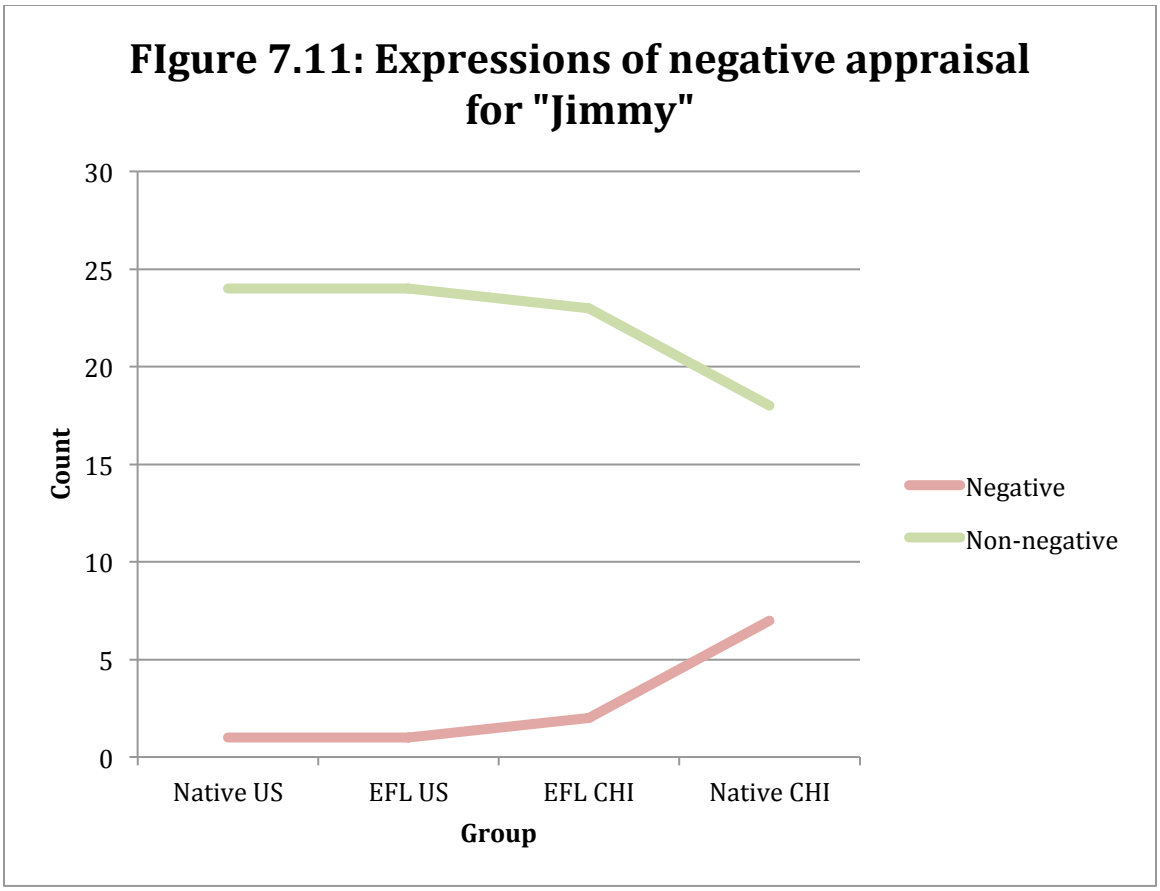


For "Charlotte," there is a visible trend of between expressions of negative appraisals and sociolinguistic background, indicating increased negative appraisal as levels of CHI sociolinguistic background increases and US sociolinguistic background decreases. Post-hoc paired t-test results indicated significant mean differences between Native US group and all other groups, with Native US participants expressing the fewest negative appraisals, followed by EFL US with the second most, EFL CHI wit the third most, and Native CHI with the most negative appraisals. It is interesting to note that the mean difference between EFL US, the group with the next highest level of US sociolinguistic background, and Native CHI would have been significant without the modified Bonferroni correction, furthermore supporting the

predicted hypothesis regarding sociolinguistic background's influence on expression of appraisal.



Results for "Gary" do not show as clear of a trend, with mixed levels of appraisals across groups, and is consistent with the lack of a significant association between the variables in analysis. Non-negative appraisals were expressed more often than negative appraisals for Native CHI, EFL CHI, and CHI groups, with the least negative appraisals being expressed by Native US group, and EFL US group exhibited more instances of negative appraisals than non-negative appraisals. Given the lack of significant relationship between variables, post-hoc testing was not carried out on this condition.



The same predicted trend seen for "Charlotte" is observable for "Jimmy", although overall appraisals are more positive for the latter than the former (potential reasons for the difference in overall appraisal levels for different characters are examined in the discussion chapter). Once more, negative appraisals decreased with higher levels of US sociolinguistic background and increased with higher levels of CHI sociolinguistic background, with the two US-residing groups each exhibiting only expression of negative appraisal, followed next by the EFL CHI group with 2 negative appraisals, and finally the Native CHI group having the highest amount at 7 instances of negative appraisal. Furthermore, post-hoc paired t-test results showed significant mean differences in appraisals between Native CHI group vs. Native US and EFL US groups.

In sum, out of the three stories in the Opinion of others prompt set, the predicted association between sociolinguistic background and expression of negative appraisal was observed in two conditions. In both the "Charlotte" and "Jimmy" stories, significant mean differences were shown between participants of US and CHI sociolinguistic background, with increased expression negative appraisal being associated with increased levels of CHI vs. US sociolinguistic background. The next section discusses these results as well as the rest of Phase 2's findings.

8. PHASE 2 DISCUSSION

This chapter offers a detailed discussion of the crucial research findings for Phase 2 results presented in Chapter 7, with reference to each of the relevant RQs. Firstly, frequency of functional [I + verb] results from the production task is examined and interpreted, with hypotheses being supported by results from the main analysis of N = 100. Results of the in-depth sub-analysis on two EFL groups and the further in-depth sub-analysis on the EFL US group are subsequently discussed, followed by interpretation of findings regarding expression of negative appraisal in the Opinion-of-others prompt set.

8.1. Frequency of functional [I + verb] usage

The main quantitative analysis of Phase 2's open-ended production task served to answer RQ 3 and by examining individuals' frequency of different functional [I+ verb] variations in Self-opinion prompts (personal preference/belief) and Opinion of others prompts (appraisal/judgment). Qualitative analysis of the production task served to answer RQ 3. RQ 6 was addressed by the two in-depth sub-analyses (N = 25 and N = 50), and RQ 5 by the second further in-depth sub-analysis (N = 25).

Each of these RQs and their hypotheses is reiterated below, with the hypotheses supported by results labeled in bold and followed by discussion of interpretation of data.

RQ 3: Does sociolinguistic background influence frequency of functional [I + verb] usage when expressing Self-opinion vs. Opinion of others?

H1: Sociolinguistic background affects the frequency of functional [1 + verb] usage. Specifically, it is predicted that higher levels of US sociolinguistic background will lead to using more assertive and less reluctant forms of [I + verb] when expressing Self-opinion, whilst higher levels of CHI sociolinguistic background will lead to using more assertive and less reluctant forms of [I + verb] when expressing Opinion of others.

H0: Sociolinguistic background does not affect frequency of functional [I + verb] usage.

Evidence from all three ANCOVA main analyses on both prompt sets support the hypothesis that sociolinguistic background does influence frequency of functional [I + verb] usage when expressing an opinion, both in terms of Self-opinion and opinion-of others. For an overview of result significance, refer to Table 7.19, which outlines the significant variables and corresponding trends for each ANCOVA of the main analysis (N = 100). Because results of ANCOVA 1 and ANCOVA 2 are similar in terms of significance, they will be discussed together as "Percentage of Deliberate/Uncertain/Reluctant [I + verb] condition;" ANCOVA 3 will be discussed as "Calculated CR of speaker confidence." Note that calculated CR results from ANCOVA 3 correspond with trends from the previous two ANCOVAs, thereby further strengthening the notion of sociolinguistic background's effect on speaker attitude and consequent pragmatic use of [I + verb]. Interpretations of each of these dependent variable measurements and trends are discussed for each prompt set (Self-opinion, Opinion of others) below.

Self-opinion prompts: To reiterate from the previous chapter, overall it appears that US

sociolinguistic background[52] is associated with more assertive and confident usages of [I + verb] constructs. That is, Native US speakers, with the highest level of US sociolinguistic background, used significantly more deliberate [I + verb] variations than EFL US speakers, who in turn used significantly more deliberate [I + verb] variations than those with Chinese sociolinguistic background (EFL CHI; Native CHI). Conversely, CHI sociolinguistic background appears to be associated with less confident usage of [I +verb], with EFL CHI and Native CHI groups using significantly more Reluctant [I + verb] variations than EFL US group, who used significantly more Reluctant [I + verb] than Native US group. Uncertain [I + verb] condition, which based on Phase 1 findings resides in the middle of the confidence rating scale, was not as similar between analyses: Whilst this [I + verb] condition was used significantly more by the Native US group than the Native CHI group in both analyses, Analysis 1 found additional significant mean differences wherein the Native US group used it more frequently than the EFL US group, who used it more frequently than the Native CHI group. Although at first these results may appear incongruous, between-group mean differences appears to nonetheless indicate an overall trend of performance, with US groups using Uncertain [I + verb] more often than CHI groups. Thus, the overall trend seems to be that higher levels of US sociolinguistic background are associated with increased Uncertain [I + verb] usage, with higher levels of CHI sociolinguistic background associated with decreased Uncertain [I + verb] usage.

For Analysis 3's calculated confidence ratings (CR), CR increased with higher levels of US sociolinguistic background and decreased with higher levels of CHI sociolinguistic background. Specifically, Native US speakers had higher calculated CRs than EFL US

[52] Refer to section 4.3.1 for distribution of sociolinguistic background across participant groups.

speakers, who had higher calculated CRs than EFL CHI and Native CHI speakers; no significant mean difference appeared between the CHI sociolinguistic groups. Thus, as predicted, a high-to-low CR trend appeared from US sociolinguistic to CHI sociolinguistic background. These findings will be interpreted after the following examination of findings for the Opinion of others prompt set.

Opinion of others prompts: To summarize, the frequency trend for this prompt set appears to juxtapose results from the previous prompt set. CHI sociolinguistic background is associated with more assertive and confident usages of [I + verb] constructs. In other words, Native CHI and EFL CHI group used significantly more Deliberate [I + verb] variations than EFL US and Native US groups. Conversely, overall it appears that US sociolinguistic background is associated with less confident usage of [I +verb], with Native US and EFL US groups using significantly more Reluctant [I + verb] variations than EFL CHI and Native CHI groups did in Analysis 2, although this mean difference was not found to be significant in Analysis 1. In terms of the Uncertain condition, trend of usage was not as clear cut as with the other conditions, with Native US and EFL US groups showing significantly higher usage than EFL CHI group in Analysis 2 but not in Analysis 1, but with Native US and EFL US groups showing significantly higher usage than Native CHI groups in both analyses. Regardless of this significance variation between analyses, there still appeared an observable trend of higher percentage of [I + verb] usage at Uncertain condition for the US-residing groups, and lower frequency for the CHI-residing groups, the difference being that EFL CHI's performance had a slightly higher percentage in relation to Native CHI's performance in Analysis 1 vs. Analysis 2. Nonetheless, the trend exhibited still remains the same across

both analyses, and this trend parallels findings for Uncertain condition in the Self-opinion prompt set, albeit being used more frequently by those with US sociolinguistic background in the Self-opinion vs. Opinion of others prompt set.

For calculated confidence ratings (Analysis 3), CR increased with higher levels of CHI sociolinguistic background and decreased with higher levels of US sociolinguistic background. Specifically, Native CHI speakers had higher calculated confidence ratings than EFL CHI speakers, who had higher confidence ratings than EFL US and Native US speakers; no significant mean difference appeared between the US sociolinguistic groups. Thus, as predicted, a high-to-low CR trend appeared from CHI sociolinguistic to US sociolinguistic background.

Furthermore, significant mean differences in calculated confidence rating were found between prompt sets as a function of sociolinguistic background, with higher confidence in Self-opinion vs. opinion-of-other prompt sets being associated with increased US sociolinguistic background, and increased mean CR difference in opinion-of-other prompt set being associated with CHI sociolinguistic background. These results mirror ANCOVA results above, suggesting that US sociolinguistic tendencies may lead to more confidence during self-expression of one's own personal preferences, whilst CHI sociolinguistic tendencies may lead to more confidence when passing social judgment on others.

Moreover, the trends in significant mean differences for performance at Deliberate and Reluctant conditions may be telling of each culture's attitude toward expressing one's own opinion vs. passing social judgment on others. Given that, in the Opinion of others prompt set, the characters under appraisal are fictional and abstract entities, it follows then that they would be processed by participants as out-group members since they are in no way connected

to them in reality. Recall that in a society like China where one's identity is relationally construed, individuals carry great consideration and empathy for in-group members, but may not share these same values when interacting with the out-group, with out-group members being held in lower esteem (Gudykunst, 1997). As a result of this lack of consideration, Chinese speakers may use more confident language in expressing their opinion of another when the other is an out-group member, especially when the other has made a socially unacceptable mistake as the characters in the prompts have done. Indeed, these findings of more Deliberate [I + verb] usage and less Reluctant [I + verb] usage associated with Chinese sociolinguistic background coincide with Brewer and Chen's (2007) claim that the traditional "indirectness" to save face that is associated with Chinese culture may only be applicable to in-group communication. Furthermore, results correlate with Bond and Wang's (1983) finding that whilst in-group relationships were dictated by reciprocal harmony and contained virtually no violence, there did appear to exist multiple forms of physical aggression toward out-group members in Chinese society, since relationships with out-group members appear to be temporary and anonymous and thus lacking concern and consideration. Without the normative consideration given when thinking of in-group members, it follows then that communication regarding potential negative appraisal of out-group members would be more aggressive and less polite, and thereby less indirect.

Alternatively, in the US where one's identity is self-dependent and the individual self is highly valued (Markus & Kitayama, 1991; Sedikedes, 2013), speakers may feel less inclined to have the right to judge others and may therefore use less direct language when passing judgment on another's mistake. In other words, when negatively appraising others they may tend to use less deliberate and more reluctant [I + verb] constructs as compared to those with

more Chinese sociolinguistic background, as results have shown. Such results coincide with Ingelhart and Welzel's (2010) and Hofstede's (1981/2011) notion that the US's high self-expression and uniqueness fosters a high level of interpersonal trust, wherein an atmosphere is created in which no opinion is put down merely for being different, and thereby US individuals would be more inclined to use more polite and therefore more indirect language when negatively appraising another. Additional support is offered by Katriel and Philipsen's (1981) study in which they found the opposite of positive evaluation in the US to not be negative evaluation, but instead neutral evaluation. It then follows that given its lack of strong stance, a neutral appraisal would be associated with less assertive, less confident language usage. Combined with Heine et al.'s (1999:780) assertion that Americans dispel negative information about the self[53] in order to maintain positive self-regard since negative self-esteem contradicts the individualistic notion of independence and self-valuation, these findings may offer an explanation to why results indicate that US sociolinguistic background appears to lead to more hesitant and less assertive [I + verb] constructs when appraising others.

Conversely, it is also this self-defined, high positive self regard and absence of fear for negative appraisal that may have led those with increased US sociolinguistic background to use more confident [I + verb] constructs when expressing their own personal opinions. Such confident [I + verb] usage from US participants corresponds to Fong's (1998) findings that US individuals give compliments more strongly and frequently than CHI individuals, thereby fostering a comfortable and forgiving conversational floor. Additionally, further support comes from Katriel and Philipsen's (1981) study, in which a US participant expresses the

[53] See Heine et al. (1999:780) for a review of studies regarding Americans' maintenance of positive self-regard, including Steele's (1988) Self-Affirmation Theory.

assumption of receiving positive evaluation when expressing oneself.

Results appear to further coincide with theories of self-construal put forth by Markus and Kitayama (1991). An independently defined self in an individualistic society such as the US seems to cultivate a higher level of self-worth and value on the individual self (Yum, 1998). This notion is further collaborated by Sedikede's (2013) finding that compared to Chinese individuals, US individuals ranked themselves in higher self-esteem and thereby possessing a higher level of significance on a conversational floor. Moreover, this self-confidence in expressing one's personal preference coincides with Choi, Nisbett and Norenzayan's (1999) and Norenzayan, Nisbett, and Smith's (1998) studies regarding US individuals viewing disposition as internal whilst Chinese individuals viewed it as malleable and situation-dependent. In other words, the internally-defined self of US culture subsequently leads to one's thoughts and opinions being defined internally by the individual as well, meaning that one's own opinion is steadfast to one's own beliefs, and thus not as easily swayed by the beliefs of others. Thus, it follows that for US speakers, an internally defined disposition would lead to more confident and direct communication style when expressing one's personal preferences. Hence, it appears that a highly tolerant conversational floor, combined with a high self-regard and focus placed on what Katriel and Philipsen (1981:304) deem the "unique and precious individual," foster a direct, assertive and confident mode of self-expression when offering one's own opinion.

This view is contrary to Chinese culture's notion of relational self-construal, or that one's individual identity is determined by the judgment and perspective of those around them (Markus & Kitayama, 1991, Sedikedes, 2013). Indeed, for Self-opinion prompts, Chinese sociolinguistic background's correlations with lower frequency of assertive [I + verb]

constructs and higher frequency of reluctant [I + verb] constructs parallel Zou et al's (2010) claim that in a collective society with less perceived tolerance for individual opinions, reluctance to share one's personal opinion may arise from fear of negative appraisal (as shown above in the discussion of opinion-of-other prompts). This negative appraisal would thereby arise when presenting an opinion that does not comply with that of the cultural norm, and accordingly individuals may develop a more careful and thus more indirect communication style, as results have shown. This tendency toward low modality markers corresponds to what Kärkkäinen (2003) calls a display of a *perceived* degree of uncertainty toward one's statement, in order to avoid speaker commitment whilst preserving face through politeness and hedging. Results further correspond with Yum's (1988) assertion of a listener-focused communicative perspective of Chinese culture, wherein the listener has the ability to cast negative evaluation on the speaker's statement. Therefore, results seem to correlate with the notion that it is the speaker's responsibility to make sure not to offend another in-group member when expressing one's own opinion; such consideration may inevitably lead to more caution and less assertion when attempting to determine how another may react.

In terms of Uncertain condition, which was rated in Phase 1 as being near the middle of the confidence rating scale (see Table 4,12), whilst there did appear somewhat of a trend in frequency in that US sociolinguistic background seemed to be associated with more Uncertain [I + verb] usage, this trend was not consistently significant between all groups in both Analyses 1 and 2. However, frequency for Uncertain condition did not appear to affect overall CR trends as calculated in Analysis 3, with a clear trend of higher confidence linked to US sociolinguistic background for Self-opinion prompts and higher confidence linked to CHI sociolinguistic background for Opinion of others prompts. Thus, since Uncertain [I +

verb] usage did not affect results matching predicted trends, combined with the notion that RQ 3 only concerns assertive and reluctant [I + verb] usage frequency, it appears that less clear results for Uncertain [I + verb] frequency do not contradict any hypotheses.

Given that Uncertain (EM) [I + verb] is associated with a moderate confidence rating, its usage subsequently should not impact levels of speaker confidence as much as usage of more assertive (i.e. DIR, PRN) or less assertive (i.e. DM) [I + verb] variations. It is therefore possible that usage of this [I + verb] variation may not be as affected by context of discourse. Indeed, it appears that results for mean difference of Uncertain condition between prompt sets possess the least mean difference across prompt sets of all [I + verb] conditions (refer to Table 7.20), with no significant mean difference for EFL US and Native CHI groups in Analysis 1 and no significant mean difference for Native US, EFL US and EFL CHI groups in Analysis 2.

Perhaps an overall higher usage of Uncertain (EM) [I + verb] for US groups vs. CHI groups may instead be a result of cultural norms transmitted through US culture, coinciding with Hofstede's (1980) theory regarding culture as being derived from social environment, and subsequently corresponding with Lakoff and Johnson's (1980) view that communicative discourse is also influenced by these socio-cultural norms. The finding that EFL US's overall performance was higher than that of the CHI groups provides evidence for Kitayama and Uskul's (2011) neuro-cultural interaction model, resulting in the input of US sociolinguistic behavioral norms shaping a speaker's neural patterns and thereby influencing speaker style of the EFL US individuals in a way that distances behavioral norms from that of CHI sociolinguistic speakers.

In all analyses across both prompt sets (aside from calculated CR of Self-opinion

prompts), the lack of a significant mean difference between frequency counts at all levels for EFL CHI and Native CHI groups follows predictions given that they share the same Chinese sociolinguistic background, with the difference between groups being merely EFL proficiency for EFL CHI individuals. For the Opinion of others prompt set, it is interesting to note that whilst both Analysis 1 and Analysis 2 note no significant mean differences in frequency at any level, the mean difference of calculated CR of EFL CHI participants was significantly higher than that of Native CHI participants by .3. This slightly higher calculated CR for EFL CHI participants may be associated with cultural priming from participants responding in English, coinciding with Ross, Xu, and Wilson's (2002) and Briley, Morris, and Simonson's (2005) findings that using English vs. Chinese primed participants to adopt a more individualistic vs. collectivistic mindset, respectively. A potential language priming effect, whilst enough to create a significant mean difference in calculated CR between EFL CHI and CHI groups, was not however strong enough to motivate EFL CHI's performance to the same level of confidence as the corresponding EFL group residing in the US, with EFL US participants still possessing a significantly higher mean calculated CR than EFL CHI participants. Thus, it appears that while language priming may potentially lead to some shift in mindset in expressing Opinion of others, it is not as influential as cultural immersion in a new sociolinguistic environment. In other words, these results appear to indicate that sociolinguistic environment acts as a much stronger influence on communication style than language priming, which is to be expected given the total immersion involved in a shift in sociolinguistic environment.

Moreover, in all analyses for the Opinion of others prompt set, the lack of a significant mean difference between frequency counts at all levels for Native US and EFL US groups

may indicate a shift in cultural mindset for the EFL US group. Given that participants of both groups reside in America and thus possess US sociolinguistic background, it appears that EFL US speakers have acquired target-like behavior in functional [I + verb] usage when expressing Opinion of others. However, significant mean differences in frequency counts for the Self-opinion prompt set suggests that target-like behavior has not yet been achieved in this particular conversational context. Nonetheless, the fact that frequency counts for EFL US group is not only significantly different to that of EFL CHI, but are also much more closely aligned than EFL CHI to frequency counts of Native US group, provides evidence that acculturation of L2 pragmatic norms has occurred to a significant extent. In other words, results suggest that for expressions of Self-opinion, these EFL US individuals are nonetheless modifying their pragmatic speaking style to mirror that of US native speakers, whilst at the same time distancing their performance from that of CHI speakers in their old sociolinguistic environment. Perhaps the reason that target-like norms have already been reached in Opinion of others prompts is due to less of a mean difference between Native US and Native CHI/EFL CHI prompts to begin with, meaning less of a shift in pragmatic behavior is necessary to achieve native-like behavior in Opinion-of-other prompts as compared to Self-opinion prompts. Nevertheless, these results suggest a noteworthy influence of sociolinguistic background on frequency of functional [I + verb] usage.

In summary, these findings support the predicted hypothesis for RQ 1 in that both Analysis 1 and Analysis 2 found a significant interactional effect between sociolinguistic background and pragmatic function. Specifically, results supported the prediction that sociolinguistic background does influence the frequency of functional [I + verb] usage so that US sociolinguistic background corresponds to higher confidence in terms of [I + verb] usage

in expressing Self-opinion, whilst CHI sociolinguistic background corresponds to higher confidence in terms of [I + verb] usage in expressing Opinion of others.

8.2. Expressions of negative appraisal

RQ 3A: For the "Opinion of others" prompt set, does sociolinguistic background influence whether participants express a negative appraisal when asked for their impression of the character in the story?

H1: Sociolinguistic background affects the expression of negative appraisals in the Opinion of others prompt set. Specifically, it is predicted that frequency of negative appraisals increase as the level of CHI sociolinguistic background increases and the level of US sociolinguistic background decreases.

H0: There is no effect of sociolinguistic background on the expression of negative appraisals in the Opinion of others prompt set.

Out of the three story conditions in the Opinion of others prompt set, the two which showed a significant association between sociolinguistic background and expression of negative appraisal both exhibited the same trend as predicted by this thesis' hypothesis, although it can be noted that the proportion of negative to non-negative appraisals was higher for "Charlotte" condition than for "Jimmy" condition across all groups. In sum, results from both significant conditions indicate that frequency of negative appraisals increases as the

level of CHI sociolinguistic background increases and the level of US sociolinguistic background decreases.

In terms of the "Gary" condition, perhaps the reason results here did not show a similar significant association is due to the open-ended nature of the prompts. Because participants were purposefully not given more information on the characters in the story so as to not create personal bias toward the characters, this resulted in the possibility for participants to make (or not make) their own assumptions regarding the scenario. Indeed, some posited whether the mistake for Gary missing his breakfast date was due to various situations from the night before, and whether those situations were justifiable. Others posited what he would do after (e.g. make amends, prevent future occurrences), and based their appraisals on these assumptions. It may be unwise to attempt to empirically evaluate what percentage of people made certain assumptions regarding the scenario vs. others, for just because a participant did not explicitly state an assumption regarding the scenario does not mean that no assumptions were made—simply they weren't said, but appraisals may still have been directed by these assumed variables. Thus, perhaps this prompt may benefit from slightly more clarification about the scenario at hand, such as stating explicitly why the mistake occurred (i.e. he forgot to set the alarm clock) whilst still maintaining an abstract, unbiased character persona in the condition.

Perhaps the reason for the higher frequency of negative appraisals for "Charlotte" vs. "Jimmy" is due to the nature of the storyline in each condition. In order to avoid repetition, each condition had a slightly different, yet very simple storyline (see Appendix 10 for reference). Because participants were purposefully not given more information on the characters in the story so as to not create personal bias toward the characters, this resulted in

the possibility for participants to make (or not make) their own assumptions regarding the characters. In the "Jimmy" condition, the character's father tells him that it's raining because there's clouds in the sky, and Jimmy subsequently makes an incorrect logical jump that every time it's cloudy, it must rain. Given this pretense, many participants assumed that Jimmy was a "child" (22 participants), calling him "young" (32 participants) or "little" (9 participants), or referring to him as a "kid" (35 participants). Responses regarding age included expressions of using age as an excuse, such as, "Don't place too many expectations on a child" (participant no. 86) and "[He's] just a kid so I think he's very smart," (participant no. 75), exemplifying the presence of lower expectations when appraising a child vs. an adult.

A chi-squared test showed no significant association between sociolinguistic background and amount of participants assuming youth of the abstract character, so such an assumption should not affect results regarding an association between sociolinguistic background and expression of negative appraisal. However, whilst 71% of participants' responses indicated an assumption of youth for "Jimmy", only 2% of participants made this assumption regarding "Charlotte," with only participants no. 55 and 86 calling her "young." This difference may be attributed to the storyline in "Charlotte" wherein she makes a mistake on an invitation for a party she is hosting, as the assumption would be made that someone old enough to host their own party is not a child. Thus, it is possible that the reason for the higher proportion of negative appraisals for "Charlotte" vs. "Jimmy" across all groups is due to a variation in assumption of age between differing storylines. Nonetheless, it must be noted that both conditions showed the same predicted trend, with higher frequency of negative appraisals being associated with higher levels of CHI sociolinguistic background.

Like the above discussion regarding Opinion of others prompts for RQ 3a, Phase 2's qualitative analysis shows trends that appear to mirror the cultural mindset of each sociolinguistic environment. These results coincide with Fong's (1998) findings that US individuals give compliments more strongly and frequently than CHI individuals, thereby fostering a more welcome and forgiving conversational floor when expressing one's own opinions; this correlates to results from the main analysis wherein US individuals used more confident and assertive [I + verb] forms to express personal opinions, whereas CHI individuals used more reluctant and less confident [I + verb] forms, with a lower calculated CR for Self-opinion prompts for CHI participants. Results furthermore coincide with Triandis' (1994) and Cross, Hardin and Gercek-Swing's (2011) work on cross-cultural differences in what comprises self esteem, with the CHI goal of "self-improvement" leading to expressions of negative appraisal being more socially acceptable in a CHI sociolinguistic environment, and the US goal of "self-enhancement" consequently viewing such negative appraisal as a threat to this goal and thereby deeming such behavior as less socially acceptable. Whilst a CHI individuals' offer of a negative appraisal may derive from a benevolent goal to facilitate self-correction and improvement in the person they are judging, cultural differences in a US sociolinguistic environment may lead to this action being deemed as hurtful or even threatening.

In sum, the two conditions which were successful in identifying an association between sociolinguistic background and expression of negative appraisal both indicated the same trend of increased frequency of negative appraisal being associated with increased CHI sociolinguistic background, thereby providing evidence for the validity of the predicted hypothesis.

8.3. Effect of Time immersed in new sociolinguistic environment

RQ 4: Does frequency of functional [I + verb] usage vary as a function of Time immersed in a new sociolinguistic environment?

H1: Frequency of functional [I + verb] usage varies as a function of Time immersed in a new sociolinguistic environment. Specifically, it is predicted that the longer one is immersed in a new sociolinguistic environment, the more a speaker's performance regarding frequency of functional [I + verb] usage will mirror a native speaker's frequencies.

H0: A speaker's performance regarding frequency of functional [I + verb] usage does not increase in similarity to a native speaker's performance the longer one is immersed in a new sociolinguistic environment.

This study's findings did not follow the prediction that a speaker's frequency of pragmatic [I + verb] usage varies as a function of Time immersed in a new L2 environment; on the contrary, results suggested that for the N = 25 sample size of EFL US participants, Time immersed in a new L2 environment did not have any significant effect on performance frequency of [I + verb], with the variable accounting for less than 1% of total variance in either prompt set. Thus, given that Time immersed in a new L2 environment did not appear to effect performance both in terms of significance and effect size, it can be safely concluded that for this data set there was no influence present, at least in terms of this particular measurement of new L2 environmental influence.

Perhaps the reason for a lack of effect lies in the nature of the participant pool: Recall that

at the time of the study, participants had been residing in the US for a range of 5 to 30 years. Such a range inevitably surpasses the amount of Time immersed in a new L2 environment Blum-Kulka and colleagues cite as needed to achieve native-like performance for the particular pragmatic speech acts they investigated, with Blum-Kulka (1986) observing native-like pragmatic behavior in request and apology acts in EFL Hebrew speakers after 5-7 years and Blum-Kulka and Olshtain (1984) after 8 years. Thus, whilst both studies observed native-like performance varying as a function of Time immersed in a new L2 environment, it is possible that the current findings do not mirror this result due to too high a range of time immersed. Indeed, it appears that for the Opinion of others prompt set, EFL US participants had successfully achieved native-like norms of pragmatic [I + verb] usage. However, what then of the Self-opinion prompt set, in which a significant mean difference in frequency was found between Native US and EFL US group's performance?

Given enough Time immersed in an L2 environment so as to acculturate and acquire native-like performance for essentially all participants, one would expect similar frequencies of [I + verb] for the Self-opinion prompt set as well. However, EFL US's significantly less confident [I + verb] usage indicates one of two possibilities: Given that the studies above investigated different pragmatic speech acts and different L1 languages than in this study, it is possible that different pragmatic tools require different lengths of Time immersed in an L2 environment to acquire native-like performance, and may differ between languages based on the salience of relationship between the L1 and L2. Nonetheless, this would not explain why native-like pragmatic performance occurred in the Opinion of others prompt set but not the Self-opinion prompt set, unless communicative context also played a large role in the acculturation process of the same [I + verb] speech act. Alternatively, another explanation

may be that there is some aspect of CHI culture that is more salient to the context regarding regarding social judgment of others than the self, which in the context of the latter would then be so prevalent that it dominates the parallel influences of US culture. For instance, coinciding with theories of RelSC found in a Relational-collectivist society like China, the goal of maintaining group harmony through reciprocity and conformity may be so imbedded in one's mindset that regardless of Time immersed in a new L2 environment a speaker may still be more reluctant to present one's own opinion lest it be construed by the listener as disagreeable or undesirable. In sum, regardless of a potentially more forgiving and positive conversational floor, and being surrounded by native speakers with high self-esteem and direct speaking style, when it comes to expressing one's own opinion EFL CHI speakers' cultural roots may run deep enough to prevent acculturation of native-like [I + verb] norms when expressing an opinion of self preference.

8.4. Effect of L2 Proficiency

RQ 6: For Mandarin EFL speakers, does frequency of functional [I + verb] usage in expressing an opinion in English vary as a function of L2 proficiency?

H1: For Mandarin EFL speakers, usage frequency of different prosodic [I + verb] belief constructs in expressing an opinion in English varies as a function of L2 proficiency. Specifically, it is predicted that DM use of [I + verb] constructions will increase as L2 proficiency decreases.

H0: For Mandarin EFL speakers, DM use of [I + verb] constructions does not increase as L2 proficiency decreases.

The results from the sub-analysis on the EFL groups showed only 2% of total variance was explained by the between-subject effect of Lex30 score and the interactional effect between Lex30 score and pragmatic function. It can therefore reasonably be concluded that L2 proficiency does not appear to influence frequency of DM [I + verb] usage, and consequently this frequency does not increase as L2 proficiency decreases. Thus, rather than supporting the predicted hypothesis of an increase in DM [I + verb] usage as L2 proficiency decreases, results support the null hypothesis. Nonetheless, with a total of N = 50 for this sub-analysis, it should be noted that observed power for the effect of Lex30 score (.16) and interactional effect between Lex30 score and pragmatic function (1.5) were much lower than the .80 threshold for reliability (Cohen, 1988), and results should therefore be interpreted with some caution.

In terms of why findings did not result in the predicted hypothesis, it is possible that a lack of significant finding may be rooted in the proficiency test chosen: While the Lex30 test and Phase 2's production task are both production tasks, Lex30 is a written production task measuring receptive knowledge, whereas Phase 2's production task is an oral production task. One may argue that perhaps a different type of production task may have yielded different results. However, given Meara and Fitzpatrick's (2000:20) argument for an implicit assumption of a correlative relationship between receptive and productive vocabulary knowledge, an individual's productive vocabulary score on the Lex30 should be implicitly correlated with active productive vocabulary proficiency in the Phase 2 oral production task.

Thus, the type production task chosen should not influence the significance of results.

Alternatively, it may more prove interesting to examine why findings did not coincide with the developmental pattern of L2 pragmatic acquisition in relation to L2 proficiency found by Trosberg (1987) in his study on Danish EFL learners' frequency of use for apology strategies, nor were they in line with Liu's (2013) findings that Mandarin EFL speakers overused the pause filler DM [I + think] form in comparison to native speakers. Perhaps the absence of increased DM use for Mandarin EFL speakers may be attributed to the range of proficiency levels of the participants. The oral DCT involved in this study inherently requires a minimum proficiency threshold in order to complete the task, given that participants had to not only comprehend the audio recordings presented to them but more importantly also be able to respond to the prompts with a substantially long response. Thus, participants with English proficiency lower than basic conversational level would be unable to complete the task and accordingly could not be included for analysis, subsequently resulting in all participants possessing a certain minimum level of L2 proficiency. Such a proficiency threshold is observable in the range of Lex30 scores, which start at 51 and reach 98 on a scale from 1 - 120. Noticeably, the bottom half of the L2 proficiency scale was not represented in this data set, but such low L2 proficiency speakers, being unable to complete the DCT, could consequently not be included, thereby suggesting that without a finer grained analysis of proficiency, it may not be possible to conclusively answer this question.

Results of a lack of effect of L2 proficiency may moreover be supported by the theory of "universal pragmatic knowledge" purported by Kasper and Rose's (1999) meta-analysis, wherein they noted that similar to this study's situation, not only did most studies analyze mostly intermediate to advanced EFL learners that were of a proficiency threshold to

complete elicitation tasks, but that all pragmatic speech acts were accessible as long as L2 proficiency reaches such a threshold for linguistic competence. Indeed, results do show that all functional forms of [I + verb] were used across all groups. It is therefore possible that the lack of a significant effect of L2 proficiency may be related to a high enough L2 linguistic competence so that all pragmatic speech acts are easy accessible for utilization. Indeed, instead of proficiency, it appears that sociolinguistic background is once again a much stronger predictor of frequency of different pragmatic [I + verb] constructs, including DM constructs of Reluctant condition.

Building on Phase 1 and 2's discussion chapters, the following chapter considers the limitations, recommendations for future research, and implications of Phase 1 and Phase 2 of this study.

9. CONCLUSION

This chapter begins with a summary of key findings from both Phase 1 and Phase 2 of this study. Limitations, future recommendations for research, and implications are subsequently presented, ending with concluding remarks regarding the study as a whole.

9.1. Summary of key findings

In sum, findings from Phase 1's novel Proof of Concept task indicated a cross-linguistic similarity in pragmatic-prosodic mapping of [I + verb] between Mandarin and English, with a resulting calculated confidence rating scheme for each [I + verb] variation being produced for use in Phase 2's analysis. As predicted, perceived confidence of [I + verb] constructs corresponded to alternative semantic interpretations created by Kontrast placement, with direct statements appearing most confident, followed by Main clause, Epistemic marker, and finally Discourse marker [I + verb] variations. Additional investigation of cross-linguistic attitudes toward speaker confidence definition showed a significant effect of sociolinguistic background on pragmatic attitude that reflected relational-collectivist vs. individualist cultural perceptions of each respective sociolinguistic environment, with US individuals relating speaker confidence to the speaker's internal state whilst Chinese individuals defined speaker confidence in terms of the listener.

Phase 2 results indicated a cross-cultural difference in frequency of functional [I + verb] usage in a manner that mirrors the sociolinguistic background of the speakers;

variations [I + verb] usage lead to different perceptions in speaker confidence, resulting in higher levels of US sociolinguistic background being associated with more confident speaker style when offering expressions of Self-opinion and less confident speaker style when offering expressions regarding Opinion of others, with the opposite trend in confidence being associated with higher levels of CHI sociolinguistic background. Frequency of [I + verb] was not found to be associated with L2 proficiency or Time immersed in a new sociolinguistic environment, but this may be perhaps more to do with proficiency and immersion time thresholds of the sample in the study. Mixed-methods examination of expression of negative appraisal further coincided with cultural perspectives of US vs. CHI sociolinguistic environment, with higher frequency of negative appraisals being associated with a self-improvement focused CHI sociolinguistic environment and lower frequency of negative approvals being associated with a self-enhancement focused US environment.

9.2. Limitations and recommendations for future research

Limitations and recommendations for future research are presented below for Phase 1 and Phase 2 of this study.

Phase 1: Proof of Concept perception task

Addition of another language: Since Phase 1 of this study only included analysis of two languages in the Proof of Concept perception task, it is limited in that no claims can be made about a definitive universality of form-function mapping for [I + verb]. Whilst

these results nonetheless seem a promising indicator of a potential universality for [I + verb] mapping especially given that they are consistent with theories of alternative semantics (see section 3.1.2) and focus-marking (see section 2.2), the addition of another language in the perception task would further solidify a notion of universality, as opposed to mere cross-cultural comparison. The choice to include a typologically different language from English and Mandarin would therefore strengthen this concept of a universal pragmatic-prosodic mapping for [I + verb], so long as the language is a) grammatically capable of [I + verb] usage (i.e. contains a first person singular pronoun marker), and b) contains socio-culturally programmed norms of [I + verb] usage (i.e. behavioral norms would lead to usage of this pragmatic speech act in conversational discourse). Given the notion of focus-marking as a cue for prominence, combined with fruitful results of cross-linguistic similarity from this study, it is predicted that another language, while typologically different, should also possess a congruent pragmatic-prosodic mapping to that of English and Mandarin.

Perception task of other cognitive verbs: Whilst due to pragmatic constraints of length Phase 1's perception task was limited in its use of only the most common cognitive verb, "think," in the [I + verb] perception task, there should be no reason to think that using another cognitive belief verb (e.g. feel, believe, suppose) should illicit different results, so long as the cognitive verb in question produces the same semantic connotations in alternative semantic theory (i.e. focus-marking the cognitive verb still results in an uncertainty stance of "I am not sure/I do not know for certain").

Nonetheless, further research may benefit from using a variety of cognitive verbs (e.g. believe, feel, etc.). Although inclusion of multiple verbs would require a larger and therefore longer task, research would benefit from observing whether the pragmatic-prosodic mapping differed across [I + verb] forms with other cognitive verbs. Such a study would therefore decrease ambiguity about the universality of a form-function mapping for [I + verb].

Alternatively, future research could also choose to conduct the perception task using the strong verdical belief verb of "know," which differs from the rest of the cognitive belief verbs in terms of the deliberate semantic connotation created through focus-marking the verb of the [I + verb] (refer to section 3.1.2.). Based on alternative semantics and focus-marking theories, a verbally focus-marked variation (EM) of "know," being deliberate in stance, would then behave more similarly in terms of speaker confidence to the pronoun focus-marked variation (MC) and may therefore illicit a higher perceived CR from participants than was rated for other verbs' EM variations. Whilst Phase 2 of this study was limited in assigning the MC's calculated CR to the EM variation of "know," if pragmatic constraints permit the additional perception testing of "know" for future research then those results may act to offer confirmation of a perceived CR for this specific [I + verb] case.

Analysis of effect of gender on confidence ratings of [I + verb] variations: Given the scope of this study, the variable of gender was not included in analysis (see section 4.2.1). However, future research including this variable may prove fruitful given the burgeoning cross-cultural work being done on this area, with some suggesting that gender

may play a role in self vs. relational construal and how this may relate to in-group vs. out-group interactions in Western (US) culture (see Cross & Madson, 1997).

However, even more interesting may be the interaction between gender and cultures, with Cross and Madson (1997) emphasizing that this gender divide may not necessarily assume a similar construct in other cultures. In sum, it may prove noteworthy to examine how gender and sociolinguistic background interact with the pragmatic-prosodic interface of [I + verb] constructs and the findings regarding speaker confidence definitions of Phase 1, as well as with the frequency of functional [I + verb] constructs in the production task of Phase 2.

Phase 2: Cross-cultural production task

Measurements of self-construal: Whilst this study based its classification of self-construal on the independence vs. interdependence of the participant's sociolinguistic background, future research may benefit from inclusion of a self-construal task (e.g. Singelis, 1994) in order to establish the degree of independent, relational, or collectivist self-construal. Without such a task, assumptions had to be made in this study about a participant's sociolinguistic background based on theoretical implications of a homogenous cultural environment as opposed to empirically established notions of self that are specific to each individual. Establishing such a measure of self-construal in future experiments would allow a quantitative basis for analyzing participants' cultural mindsets at an individual level, as well as allowing for a within-group comparison of how self-construal measurements vary within each sociolinguistic background. However, Kitayama et al. (2009) warn that given the explicit self-report nature of such self-construal tasks, participants may complete the task

guided by habitual psychological tendencies learned from cultural norms, instead of tapping into an innate measure of one's self-construal. Nonetheless, such a measurement of self-construal may prove useful in providing an empirical basis of measuring cultural processes, placing a quantitative inquiry onto such a foundation of sociocultural theory.

Opinion of close-others vs. strangers: Although this study included two separate prompt sets–Self-opinion and Opinion of others—in order to examine how participants may respond differently based on who their opinion was evaluating, the Opinion of others prompt set was limited in that it only included stories regarding strangers (i.e. out-group others) and thereby signaling high social distance between the participant and the character (Brown & Levinson, 1987). Currently, the Opinion of others prompt set assumes that since the characters in the prompts are fictional, being only identified by name, participants will treat these characters as strangers and therefore out-group members. A point of improvement for future studies could involve explicitly specifying these characters as strangers, such as by saying, "A girl working at the shop you've entered" or "A man passing by on the street" to ensure high social distance.

Furthermore, future research may benefit from including a third prompt set regarding opinion of *close*-others so as to observe how lower social distance may influence speaker style. The inclusion of a third prompt set regarding appraisal of close in-group members, whilst lengthening the time needed to complete the task, would provide useful data regarding whether relational-collectivist participants actually behaved differently toward appraising in-group vs. out-group members and whether individualistic participants performed similarly toward both in-group and out-group others. That is, an investigation could thereby be

conducted on whether there is an interaction between sociolinguistic background and social distance in terms of speaker style. Such a prompt set might appear similar to the Opinion of others prompt set, only instead of providing the name of a fictional, abstract character, prompts may describe the character as "Your mother" or "A close friend" in order to establish a close relationship with the participant.

Empirical justification of a cross-cultural design: Recall that given the dichotomous evaluation between US and Chinese culture, it is not possible to make any claims regarding overall differences between Independent and Relational-Collectivist cultures. However, if future research were to include more than one culture under each term respectively, then there could be a within-cultures comparison between independent and relational-collectivist cultures to assess if they behave similarly in terms of frequency on the production task (as well as in terms of cultural variation in speaker definition of confidence from Phase 1). That is, the addition of other sociolinguistic backgrounds may allow for less ambiguity in claims of whether such concepts of individualism-collectivism and self-construal are indeed the influencing causes behind behavioral variation in findings.

Inclusion of adolescent participants regarding analysis of Time immersed in new sociolinguistic environment: Whilst this study only recruited adult participants, it may be interesting to see if such results would be replicated in a younger sample pool. Although many of the participants in this data set have spent decades in the US and many began learning English as youths, only two actually immigrated before adulthood, at the ages of 5 and 6 respectively, which may perhaps already be old enough to affect neuroplasticity. Thus,

the current study can only make claims regarding the effect of Time immersed in a new sociolinguistic environment on frequency of [I + verb] for adult participants. It is possible that with adolescents or children, native-like performance could be acquired for expressions of Self-opinion as well as Opinion of others. Subsequently, further research may choose to include a wider age range to incorporate participants whose neuroplasticity may be more malleable than that of adults, and to include participants who immigrated to their new L2 environment whilst still children.

Increase reliability of findings for analysis of L2 proficiency: Given the internal reliability and pragmatic ease of administering the Lex30 proficiency test, the only recommendation for analyzing this variable would be to increase the sample size of each group. Whilst a power effect size was calculated before data collection (see section 4.3.1), this was calculated based on the original N = 100 analyses as well as taking into consideration the pragmatic constraints of data collection for this thesis. However, with the addition of covariates for the sub-analyses, observed power was lowered and could no longer match the reliability threshold of .80, subsequently weakening the potential significance of results. Ideally, given more time and participant incentive, a larger sample could be collected so as to raise the reliability of results when covariates such as Lex30 are incorporated into analysis. If then findings still indicate no significant effect of L2 proficiency on DM [I + verb] usage, then this result, going against previous literature, would prove even more interesting and dependable.

9.3. Implications

With no previous research having conducted such an investigation, results of Phase 1's Proof of Concept task provide evidence for a cross-linguistic similarity in pragmatic-prosodic mapping of [I + verb]. This finding offers a novel contribution to research involving the pragmatic-prosodic interface: Whilst previous research on [I + verb] has focused on theories regarding how focus-marking may pragmatically affect the meaning behind [I + verb] variations (e.g, Dehe & Wichmann, 2010; Wichmann, Dehe & Barth-Weingarten, 2009), research has not previously investigated the cognitive basis of these notions in an empirical manner. This study thereby provides novel evidence that such assertions regarding assertiveness of various [I + verb] conditions created through focus-marking, and the speaker attitude they invoke regarding confidence, have validity in the potential universality of focus-marking as a cue for prominence. Furthermore, due to the different natures of focus-marking in English vs. Mandarin, the results from this phase also serve to strengthen the notion that pitch and pitch range both function successfully as a cue of prominence.

Phase 1's analysis of "speaker confidence definition" further contributes toward the field of cross-cultural pragmatics. In a research area with numerous qualitative theories (e.g. Hofstede, 1980, Triandis, 1994), this mixed methods analysis subsequently provides evidence of culture's impact on perception, with CHI individuals perceiving confidence in relation to listener perception and US individuals in terms of speaker perception. Pedagogical implications therefore include the necessity to teach L2 learners not only the fundamental linguistic bases of a language, but also the culturally programmed social perceptions behind the language use. In order to successfully express one's desired

speaker attitude and stance in each sociolinguistic environment, one must be taught the pragmatic patterns of thinking that underlie the process of conversational discourse.

Phase 1's perception task also benefits future research by providing calculated notions of confidence rating for each [I + verb] variation, which no previous studies have as of yet to produce. This can be used by future research to examine whether speaker confidence in terms of [I + verb] usage is different across sociolinguistic backgrounds or conversational contexts, in a manner that explores whether different individuals behave differently in various conversational or social situations. Indeed, this confidence rating scheme for [I + verb] variations is adopted and applied in Phase 2's analysis, allowing for the relationship between trends of confidence rating and frequency of functional [I + verb] usage to be better investigated.

Whilst previous research on [I + verb] has investigated focus-marked [I + verb] variation within one language (e.g. Dehe & Wichmann, 2010) or the particular DM [I + verb] across two languages (e.g. Huang, 2011), there has not been work conducted on the full range of prosodic [I + verb] variations in a cross-linguistic design. The production task of Phase 2 thereby contributes to the field of interlanguage pragmatics by offering a novel glance into how sociolinguistic background interacts with pragmatic [I + verb] use in English vs. Mandarin, and how this may reflect perceptions of speaker confidence.

Additionally, because findings indicating no effect of L2 proficiency for this particular pragmatic speech act do not correlate with previous research (e.g. Liu, 2013), the results from this study should prompt further investigation into the relationship between proficiency and [I + verb] gap filler use, and perhaps whether frequency of usage is more dependent on conversational context.

Furthermore, the significantly different results between frequency of functional [I + verb] variations in the Self-opinion prompt set and Opinion of others prompt set highlight the importance of communicative context when making generalizations about a specific sociolinguistic environment's behavior. Many cross-cultural theories make assumptions about homogeneity of language use across all situations, leading to such unanimous notions of "directness" and "high modality" being associated with individualistic IndSC societies like the US, (Singelis, 1994; Markus & Kitayama, 1991) and notions of "indirectness" and "low modality" being associated with collectivist RelSC societies like China (Lu, 2002; Wu, 1996). However, it is clear from this study's genre-specific results of the Opinion of others prompt set that such generalizations do not hold consistently across all contexts, and that given a particular situation US individuals may tend to adopt a more reluctant speaker style whereas CHI individuals may tend to adopt a more assertive speaker style. Thus, the implication for future research on theories of cultures should take into account not only the issue of homogenizing a culture of individuals, but also the issue of homogenizing a culture's behavior across different social contexts.

Even though Phase 1 results indicate a congruent pragmatic-prosodic mapping of perceived speaker confidence for [I + verb] between Mandarin and English, perception in a specific or isolated incident such in Phase 1's perception task may not necessarily reflect how speakers utilize such pragmatic speech-acts in context-specific situations. For instance, Mandarin speakers may utilize indirectness (i.e. Reluctant [I + verb]) to distance themselves from an opinion in order to save face or protect oneself from negative appraisal, but such indirectness may not necessarily be a reflection of the personality of the speaker; it is possible that culturally specific neural patterning may also cause a

speaker to speak less assertively due to a relational-collectivist notion of thinking about the listener's mindset and not wanting to offend (Yum, 1988). However, given cross-cultural differences in communicative norms, such indirectness may automatically be viewed unfavorably as nervousness or reticence in an individualistic sociolinguistic environment such as the US, where the focus is on the individual self and direct communication acts as a sign of high self-esteem. Given that language acts to present who we are and how we want to be seen (Paltridge & Phakiti, 2015:205), learners of an L2 language should be sure to familiarize themselves with how native speakers use and interpret pragmatic cues in conversation so as to be accurately evaluated by native speakers.

Thus, much like for Phase 1's speaker confidence analysis, results for Phase 2 indicate the need to teach sociocultural norms of communication in the L2 classroom. Not only should L2 learners receive knowledge on when to use certain speech acts, but they should also be taught the sociocognitive processes underlying the use of such speech acts. Given that focus-marking appears to act as a universal cue for prominence, it may not be necessary to instruct learners on now to utilize focus-marking to create differences in pragmatic intent, since this process should be relatively analogous across languages. However, that is not to say that they will automatically know the programmed cultural norms of an L2 sociolinguistic environment: When said with a certain pitch marking or intonation, a speech act may carry a culturally-specific connotation, pragmatic knowledge that an L2 learner must acquire through being taught since it has not been transmitted through habitual patterns in the L2 environment over time. L2 pragmatic fluency may thereby be further developed through consciousness-raising techniques to make speakers

aware of such cross-linguistic, cross-cultural pragmatic differences, such as through providing L2 learners with explicit metapragmatic information (House, 1996:250).

The same pedagogical implication for L2 classroom teaching of pragmatic knowledge is confirmed from Phase 2's results on expression of negative appraisals. Given the behaviorally transmitted sociocultural norms behind communicative intent, what may be acceptable or appropriate in one sociolinguistic context may not necessary be so in another (Cross, Hardin & Gercek-Swing, 2011). These findings are another example of the necessity to include a pragmatic cultural component in the L2 classroom, not merely one that addresses the traditions and habits of a culture but one that also invites the learner to explore the cultural mindset of an L2 environment.

9.4. Concluding remarks

In sum, the two phases of this cross-linguistic, cross-cultural study of US vs. CHI native and EFL speakers served to investigate the link between language, culture, and cognition in terms of pragmatic use of [I + verb], with findings that offer novel contributions in cross-cultural pragmatics and interlanguage pragmatics. Whilst, like in any study, limitations were still present, analyses nonetheless indicated significant and interesting results in both phases of research.

Whereas Phase 1's Proof of Concept perception task provided evidence for a cross-linguistic similarity in pragmatic-prosodic mapping of perceived confidence rating of [I + verb] constructs, Phase 2's findings revealed that the particular cultural convention in which these pragmatic speech acts are used nevertheless differs between US and CHI

sociolinguistic environments. That is, even if a speech act is used in both an L1 and L2 language, the implicature of such a pragmatic tool in a particular social context, and how this may reflect upon the speaker, may nonetheless differ depending on cultural mindset.

Indeed, cross-cultural differences in frequency of functional [I + verb] usage found in Phase 2, combined with analogous culturally derived differences in pragmatic perception found in Phase 1, indicate the need to emphasize L2 pragmatic fluency when acquiring another language and especially when entering a new sociolinguistic environment. Whilst native speakers of a particular sociolinguistic environment acquire such pragmatic knowledge subconsciously through behavioral patterns over time, L2 learners outside of this environment must be taught such knowledge in order to avoid having their communicative intent be misconstrued by native speakers. Such fluency includes the knowledge that speaker style and intent may vary across different social situations, and future research should make a point of choosing a particular genre of communicative context when investigating a cross-cultural design. Through learning the culturally specific sociocognitive processes behind language use in different environments, as well as how pragmatic language use may vary between different communicative contexts, the L2 learner can further their path toward acquiring native-like fluency, so as to better understand the mindset of a native speaker and minimize potential cross-cultural miscommunication.

REFERENCES

Aijmer, K. (1997). *I think* - an English modal particle. In: T. Swan and O. J. Westvik (Eds.), *Modality in Germanic Languages: Historical and Comparative Perspectives*. Berlin: Mouton de Gruyter.

Alford, D. K. H. (1978). Demise of the Whorf Hypothesis. *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society,* 485-499.

Anthony, L. (2011). *AntConc (Version 3.2. 2)*[Computer Software]. Tokyo, Japan: Waseda University.

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole, & J. Hualde (Eds.), *Laboratory Phonology (9),* 547–576.

Aslan, E. (2015). When the Native Is Also a Non-native: "Retrodicting" the Complexity of Language Teacher Cognition. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 71(3),* 244–269.

Bardovi-Harlig, K. & Hartford, B. S. (2005). Institutional Discourse and Interlanguage Pragmatics Research. In K. Bardovi-Harlig & B.S. Hartford. (Eds.), *Interlanguage Pragmatics*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Baumgarten, N. & House, J. (2010). I think and I don't know in English as lingua franca and native English discourse. *Journal of Pragmatics, 42,* 1184-1200.

Beckman, M. & Hirschberg, J. (1999). *The ToBI Annotation Conventions*. Retrieved from http://www.ling. ohio-state.edu/~tobi/ame_tobi/ annotation_conventions.html.

Blum-Kulka, S., & Olshtain, E. (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics, 5,* 196-213.

Blum-Kulka, S., & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Journal of Pragmatics, 8,* 47-61.

Boersma, P. & Weenink, D. (2016). Praat: Doing phonetics by computer. [Computer program]. Version 6.0.21.

Bond, M., & Wang, S. (1983). Aggressive behavior in Chinese society: The problem of maintaining order and harmony. *Bulletin of the Hong Kong Psychological Society*, *8*, 5–25.

Breen, M., Fedorenko, E., Wagner, M. & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes,* 25, 1044-1098.

Briley, D. A., Morris, M. W. & Simonson, I. (2005). Cultural chameleons: Biculturals, conformity motives, and decision making. *Journal of Consumer Psychology, 15(4),* 351-362.

Brown, G & Levinson, S. (1987). *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press.

Cai, L. & Wang, Y. (2013). Interlanguage pragmatics in SLA. *Theory and Practice in Language Studies, 3(1),* 142-147.

Calhoun, S. (2007). What Makes a Word Contrastive? Prosodic, Semantic and Pragmatic Perspectives. Presented to the panel on 'Prosody and Pragmatics in Spoken Language Corpora', *10th International Pragmatics Conference*, Gothenburg.

Cao, J. (1998). Some aspects on Chinese intonation. *Proceedings of the Conference on Phonetics of the Language in China,* Hong Kong, May 28-30.

Cao, J. (2004) Intonation Structure of Spoken Chinese: Universality and Characteristics. In *From Traditional Phonology To Modern Speech Processing*, 31-38.

Cappelli, G. (2007). "*I Reckon I Know how Leonardo Da Vinci Must Have Felt...*" *Epistemicity, evidentiality and English verbs of cognitive attitude*. Pari, Italy: Pari Publishing.

Chai, E. & Fontaine, G. (2007). Context preference shifts in the communicative behavior of Caucasian students in Hawaii. *Intercultural communication studies, 16(3),* 179-191.

Chen, S., Chen, S., & He, T. (2012). Surface cues and pragmatic interpretation of given/new in Mandarin Chinese and English: A comparative study. *Journal of Pragmatics, 44(4),* 490-507.

Chen, Y., Chen, X., & Portnoy, R. (2009). To whom do positive norm and negative norm of reciprocity apply? Effects of inequitable offer, relationship, and relational-self orientation. *Journal of Experimental Social Psychology, 45,* 24-34.

Chen, Y. & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics, 36(2008),* 724–746.

Choi, I., & Markus, H. R. (1998). *Implicit theories and causal attribution East and West*. Unpublished manuscript, Ann Arbor: University of Michigan.

Choi, I., Nisbett, R.E. & Norenzayan, A. (1999). Causal attribution across cultures: variation and universality. *Psychological Bulletin, 125,* 47–63.

Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive Psychology, 11,* 430-477.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*, 5-37.

Cutler, A., Dahan, D. & Donselaar, W.V. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech 40 (2),* 141–201.

Dehe, N. & Wichmann, A. (2010). Sentence-initial I think (that) and I believe (that): Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in Language, 34(1),* 36-74.

Diessel, H. & Tomasello. M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics, 12(2),* 97–141.

Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Fetzer, A. (2011). I think this is I mean perhaps this is too erm too tough a view of the world but I often think... Redundancy as a contextualization device. *Language Sciences (33)*, 255-267.

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*. London: Sage.

Fitzpatrick, T., & Meara, P. (2004). Exploring the validity of a test of productive vocabulary. *VIAL: Vigo International Journal of Applied Linguistics*, *1*, 55–73.

Flemming, E. (2008). The role of pitch range in focus marking. In *Slides from a talk given at the Workshop on Information Structure and Prosody, Studiecentrum Soeterbeeck.*

Fong, M. (1998). Chinese immigrants' perception s of semantic dimensions of direct/indirect communication in inter cultural compliment inter actions with North Americans. *The Howard Journal of Communication, 9,* 245–262.

Fox, B. A. (2001). Evidentiality: Authority, Responsibility, and Entitlement in English Conversation. *Journal of Linguistic Anthropology, 11(2),* 167-192.

Fox Tree, J. E., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics, 34,* 727-747.

Fus, G. (2008). Pragmatic markers in Chinese. *Journal of Pragmatics, 40*, 1687-1718.

Gao, G. (1998). "Don't Take My Word for It." Understanding Chinese Speaking Practices *International Journal of Intercultural Relations, 22(2),* 163-186.

Glucksberg, S. & McGlone, M. (2001). *Understanding figurative language: from metaphors to idioms*. Oxford: Oxford Univ. Press.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics,* 41-58. New York: Academic Press.

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Gudykunst, W. B. (1997). Cultural variability in communication. *Communication Research, 24(4),* 327-348.

Gumperz, J. J. and Cook-Gumperz, J. (2008), Studying language, culture, and society: Sociolinguistics or linguistic anthropology? *Journal of Sociolinguistics, 12*, 532–545.

Gumperz, J. J. & Levinson, S. C. (1991). Rethinking linguistic relativity. *Current Anthropology, 32(5),* 613-623.

Gwet, K. L. (2010). *Handbook of Inter-Rater Reliability (2nd Edition)*. Gaithersburg, MD: Advanced Analytics, LLC.

Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review, 11,* 4-27.

Heine, S.J., Lehman, D.R., Markus, H.R. & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psycho. Review, 106*, 766–94.

Heine, S. J., & Lehman, D. R. (2004). Move the body, change the self: Acculturative effects on the self-concept. In M. Schaller & C. Crandall (Eds.), *Psychological foundations of culture*. Mahwah, NJ: Lawrence Erlbaum.

Hill, T. (1997). *The development of pragmatic competence in an EFL context*. Ph.D. Thesis, Tokyo: Temple University Japan.

Hofstede, G. (1980). Culture's Consequences: International Differences in Work-Related Values. Beverly Hills, CA: Sage Publications.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind.* 3rd Edition, McGraw-Hill USA.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture, 2(1)*. 3-26.

Hong, Y., Chiu, C. & Kung, T. M. (1997). Bringing culture out in front: Effects of cultural meaning system activation on social cognition. In K. Leung, U. Kim, S. Yamaguchi, & Y. Kashima (Eds.), *Progress in Asian social psychology,* 135-146. Singapore: Wiley.

Howell, D. C. (1998). *Statistical methods in human sciences*. New York: Wadsworth.

House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition, 18,* 225-252.

Huang, L. (2011). *Discourse markers in spoken English: a corpus study of native speakers and Chinese non-native speakers (Ch. 8: Analysis of "I think")*. Ph.D. Thesis, Birmingham: University of Birmingham.

Hui, C.H., & Triandis, H.C. (1986). Individualism-collectivism: A study of cross-cultural researches. *Journal of Cross-cultural Psychology, 17,* 225-248.

Huynh, H. & Feldt, L.S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and plit-plot designs. *Journal of Educational Statistics, 1(1),* 69-82.

Ingelhart, R. & Welzel, C. (2010). "Changing Mass Priorities: The Link Between Modernization and Democracy." *Perspectives on Politics, 8(2),* 110-119.

Jin, S. (1996) *An Acoustic Study of Sentence Stress in Mandarin Chinese*. Ph.D. dissertation, Columbus, Ohio: Ohio State University.

Kaltenbock, G. (2009). *Pragmatic functions of parenthetical I think*. In Paper Presented at the 11th International Pragmatics Conference, Melbourne.

Kärkkäinen, E. (2003). *Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on I think*. Amsterdam/Philadelphia: John Benjamins.

Kasper, G. & Rose, K. (1999). Pragmatics and SLA. *Annual Review of Applied Linguistics, 19,* 81-104.

Kashima, Y. (2008). A social psychology of Cultural Dynamics: Examining how cultures are formed, maintained, and transformed. *Social and Personality Psychology Compass, 2(1),* 107-120.

Katriel, T., & Philipsen, G. (1981). "What we need is communication": "Communication" as a cultural category in some American speech. *Communication Monographs, 48,* 301-317.

Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Englewood Cliffs, N.J: Prentice Hall.

Kerekes, J. (1992). *Development in nonnative speakers' use and perception of assertiveness and supportiveness in mixed-sex conversations*. Honolulu, HI: University of Hawai'i at Manoa, Department of English as a Second Language. Occasional Paper #21.

Khojastehrad, S. (2012). Distribution of Hesitation Discourse Markers Used by Iranian EFL Learners during an Oral L2 Test. *International Education Studies, 5(4),* 179-187.

Kim, H. & Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity: A cultural analysis. *Journal of Personality and Social Psychology, 77,* 785-800.

Kim, M.-S., Sharkey, W. F., & Singelis, T. M. (1994). The relationships between individuals' self-construals and perceived importance of interactive constraints. *International Journal of Intercultural Relations, 18,* 117-140.

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology, 97*, 236-255.

Kitayama, S. & Uskul, A. K. (2011). Culture, mind, and the brain: Current evidence and further directions. *Annual review of psychology, 62,* 419-449.

Kruskal, J. B. (1983). An overview of sequence comparison. In D. Sankoff & J.B. Kruskal (Eds.), *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*. Reading, MA: Addison-Wesley.

Kurumada, C., Brown, M., & Tanenhaus, M. K. (2012). Pragmatic interpretation of speech: It looks *like* speech adaptation. *The proceedings of the 35th annual meeting of the Cognitive Science Society*, Sapporo, Japan, August.

Ladd, R. D. (1996). *Intonational Phonology*. Cambridge, MA: Cambridge University Press.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press: Ann Arbor.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar, Vol. 1: Theoretical perspectives*. Stanford: Stanford University Press.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Vol II: Descriptive Applications*. Stanford: Stanford University Press.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Levenshtein, V. L. (1996). Binary codes capable of correcting deletions, insertions and reversals. *Proceedings of the Soviety Physics-Doklady, 10(8),* 707-710.

Lin, C. Y., Woodfield, H. & Ren, W. (2012). Compliments in Taiwan and Mainland Chinese: The influence of region and compliment topic. *Journal of Pragmatics, 44,* 1486-1504.

Liu, B. (2013). Effect of first language on the use of English discourse markers by L1 Chinese speakers of English. *Journal of Pragmatics, 45(1)*, 149-172.

Lu, D. (2002). Cultural features in speech acts: A Sino-American comparison. *Language, Culture and Curriculum, 14*(3), 214-223.

Ma, R. (1992). The role of unofficial intermediaries in interpersonal conflicts in the Chinese culture. *Communication Quarterly*, *40*, 269- 278.

Maass, A., Arcuri, L., & Suitner, C. (2014). Shaping intergroup relations through language. In T. Holtgraves (Ed.), *Oxford Handbook of Language and Social Psychology*. Oxford University Press.

Markus, H. R., & Kitayama, S. (1991). Culture and the self; Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.

Meara, P. & Fitzpatrick, T. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. *System, 28(1),* 19-30.

Medgyes, P. (1992). Native or non-native: Who's worth more? *ELT Journal, 46(4),* 340–349.

Medgyes, P. (1994). *The non-native teacher*. London, UK: Macmillan.

Meyer, R. & Mleinek, I. (2006). How prosody signals force and focus: A study of pitch accents in Russian yes–no questions. *Journal of Pragmatics, 38,* 1615–1635.

Miles, M. & Huberman, M. (2013). *Qualitative data analysis: An expanded sourcebook*. New York: Sage Publications, Inc.

Miyamoto, Y., & Eggen, A. T. (2013). Cultural perspectives. In *J. DeLamater & A. Ward (Eds.), Handbook of Social Psychology (2nd Ed.)*, pp. 595-624. New York: Springer.

Niemann, H., Noth, E., Batliner, A., Buckow, J, Gallwitz, F., Huber, R., Kießling, A., Kompe, R. & Warnke, V. (1998). Using prosodic cues in spoken dialogue systems. *Proceedings of Specom '98 Workshop, St. Petersburg,* 17-28.

Nikula, T. (1996). *Pragmatic Force Modifiers: A Study in Interlanguage Pragmatics*. Jyvaskyla: University of Jyvaskyla.

Norenzayan, A., Nisbett, R. E., & Smith, E. E. (1998). *Rule based vs. memory based category learning East and West*. Unpublished manuscript. Ann Arbor: University of Michigan.

Ouyang, I. C. & Kaiser, E. (2013). Prosody and information structure in a tone language: an investigation of Mandarin Chinese. Language and Cognitive Processes. *Language and cognitive processes,* 1-16.

Oyserman, D., Coon, H. M. & Kemmelmeiser, M. (2002). *Rethinking individualism and collectivism*: *Evaluation of theoretical assumptions and meta-analyses*. *Psychological Bulletin, 128,* 3-72.

Oyserman, D. & Lee, S.W.S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin, 134,* 311–42.

Palmer, F.R. (1986). *Mood and Modality*. Cambridge, MA: Cambridge University Press.

Paltridge, B. & Phakiti, A. (2015) *Research Methods in Applied Linguistics: a practical resource*. London: Bloomsbury.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Ph.D. Thesis, Cambridge, MA: MIT.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In: *Intentions in communication*, Philip R. Cohen, Jerry Morgan & Martha E. Pollack (Eds.), 271–311. Cambridge, MA: MIT Press.

Quirk, R, Greenbaum, S, Leech, G. & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow: Longman.

Rooth, M. (1985). Association with Focus. Ph.D. Dissertation, University of Massachusetts, Amherst. Graduate Linguistics Student Association (GLSA), Amherst/MA.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics 1,* 75–116.

Rooth, M. (1996). Focus. In S. Lappin (Ed.), *The Handbook of Contemporary Semantic Theory,* 271-297. Blackwell, Oxford.

Rose, K. R. (1998). *An exploratory cross-sectional study of interlanguage pragmatic development*. Unpublished manuscript.

Ross, M., Xun, W. Q., & Wilson, A. E. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin, 28(8),* 1040-1050.

Rubin, D.B. (1987). *Multiple imputation for non-response in surveys*. New York, NY: John Wiley.

Schack, K. (2000). Comparison of intonation patterns in Mandarin and English for a particular speaker. *University of Rochester Working Papers in the Language Sciences,* 1-29.

Schafer, J.L. & Olsen, M.K. (1998). Multiple imputation for multivariate missing data problems: a data analyst' perspective. *Multivariate Behavioral Research*, *33*, 545–571.

Schafer, J.L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research*, *8*, 3–15.

Sedikides, C. (2013). A Three-Tier Hierarchy of Self-Potency: Individual Self, Relational Self, Collective Self. *Advances in Experimental Social Psychology, 48*, 235-295.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin, 20,* 580 –591.

Spencer-Oatey, H. (Ed.). (2000). *Culturally speaking: Managing rapport through talk across cultures*. London: Continuum.

Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition*. 2nd edition. Oxford: Blackwell.

Steele, C.M. (1988). The psychology of self-affirmation: sustaining the integrity of the self. In L Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 261–302. Orlando, FL: Academic.

Stenström, A.B. (1994). *An Introduction to Spoken Interaction*. London: Longman.

Tench, P. (1990) *The Roles of Intonation in English Discourse*. Frankfurt, Germany: Peter Lang.

Tench, P. (1996) *The Intonation Systems of English*. London: Cassell Academic.

Triandis, H. C. (1994). *Culture and social behavior*. Boston: McGraw-Hill.

Triandis, H. C. (2001). Individualism-Collectivism and personality. *Journal of personality, 69(6),* 907-924.

Trosborg, A. (1987). Apology strategies in natives/non-natives. *Journal of Pragmatics. 11,* 147-167.

Van Ginkel, J.R. & Kroonenberg, P.M. (2014). Analysis of Variance of Multiply Imputed Data. *Multivariate Behavioral Research, 49(1),* 78-91.

Walters, J. (2012). Aspects of Validity of a Test of Productive Vocabulary: Lex30. *Language Assessment Quarterly, 9(2),* 172-185.

Wang, Q. (2008) L2 Stress Perception: The reliance on different acoustic cues. *Proceedings of Speech Prosody, Brazil,* 635-638.

Wang, V. (2011). *Making requests by Chinese EFL Learners*. Amsterdam: John Benjamins Publishing.

Wang, Z. (2014). Developing Accuracy and Fluency in Spoken English of Chinese EFL Learners. *English Language Teaching, 7(2)*, 110-118.

Whorf, B. L. (1956). A linguistic consideration of thinking in primitive communities. In J. B. Carroll (Ed.), *Language, Thought, and Reality: Selected Writings of Benjamin Lee,* 65-86. Cambridge: MIT Press.

White, I.R., Royston, P., & Wood, A.M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*, 377–399.

Wichmann, A. & Blakemore, D. (2005). The prosody-pragmatics interface. *Journal of Pragmatics, 38,* 1537-1541.

Wichmann, A., Dehé, N., & Barth-Weingarten, D. (2009). Where prosody meets pragmatics: research at the interface. In: D. Barth-Weingarten, N. Dehé & A.

Wichmann (Eds.), *Where Prosody meets Pragmatics (Studies in Pragmatics),* 1-20. Bingley: Emerald.

Wilson, D. & Wharton, T. (2006). Relevance and Prosody. *Journal of Pragmatics, 38,* 1559-1579.

Wong, R. Y.M. & Hong, Y. (2005). Dynamic Influences of Culture on Cooperation in the Prisoner's Dilemma. *Psychological Science, 16*, 429-434.

Wu, D. (1996) Chinese childhood socialization. In M.H. Bond (Ed.), *The Handbook of Chinese Psychology,* 143–54. Hong Kong: Oxford University Press.

Wu, Z. (1996b). A new method of intonation analysis for Standard Chinese: frequency transposition processing of phrasal contours in a sentence. In G. Fant et al. (Eds.). *Analysis, Perception and Processing of Spoken language*, Elsevier Science B.V.

Wu, Y., Wang, J. & Cai, Z. (2010). The use of 'I think' by Chinese EFL learners: a study revisited. *Chinese Journal of Applied Linguistics 33 (1),* 3–24.

Young, L. (1994). *Crosstalk and Culture in Sino-American Communication*. Cambridge: Cambridge University Press.

Yuan, Y. (2001). An inquiry into empirical pragmatics data-gathering methods: Written DCTs, oral DCTs, field notes and natural conversations. *Journal of Pragmatics, 33,* 271-292.

Yu, M. (1999). Universalistic and culture-specific perspectives on variation in the acquisition of pragmatic competence in a second language. *Pragmatics, 9(2),* 281-312.

Yum, J. O. (1988). The impact of Confucianism on interpersonal relationships and communication patterns in East Asia. *Communication Monographs, 55,* 374-38.

Zhao, L. (2013). *The effect of bilingualism on young children's metaphor comprehension*. Masters Thesis, London: University College London.

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *Neuroimage, 34,* 1310-1316.

Zimmermann, M. & Onea, E. (2011). Focus marking and focus interpretation. *Lingua, 121,* 1651-1670.

Zou, X., Bao, X., & Luo, L. (2010). Integration of Intonation in F0 Trajectory prediction using MSD-HMMs. In *Speech Prosody 2010-Fifth International Conference*.

Zou, X., Tam, K., Morris, W.M., Lee, L., Lau, I & Chiu, C.Y. (2010). Culture as Common Sense: Perceived Consensus vs. Personal Beliefs as Mechanisms of Cultural Influence. *Journal of Personality and Social Psychology. 97(4),* 579-597.

APPENDICES

Appendix 1: Information sheet (Phase 1)

[Group 1:]

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk



Participant Information Sheet

Title of Project: Cross-linguistic expressions of opinion
Researcher: Lucy Zhao, PRS DPhil Education

Hello, my name is Lucy Zhao, and I am a DPhil student at Oxford University under the supervision of Dr. Victoria Murphy. I would like to invite you to participate in this research project on the influence of culture and language on communication. You are being asked to participate due to your unique language skill set. Participation is completely voluntary; choosing not to take part will not be detrimental to you in any way. The survey involves voicing your thoughts on intriguing and sometimes quirky opinion questions.

Purpose of study: This research will explore how language knowledge and language environment affect self-expression and communication, particularly in expression of opinions. Participants in this study involve students from America and China.

Benefits of study: While participants do not directly benefit, this study does indirectly benefit participants through potential findings and implications for language learning and communication. This research aims shed light on how language acquisition and language-learning environment affects communication style and cross-cultural interaction, which has implications for improving language-teaching methods. Ultimately, I aim to improve cross-cultural communication between native and foreign language speakers.

Risks of study: There are no foreseeable or expected risks involved in participating in this project.

Description of procedure: If you agree to participate, I will ask you to fill out a brief (<1 minute) online questionnaire about your language background. You may then be asked to participate in the next phase, which consists of a short (~5 minute) online survey consisting of listening to audio recordings and answering questions relating to a short story.

Contact information: If you have any questions or concerns, or desire more information, please email the primary researcher, Lucy Zhao, at: lucy.zhao@education.ox.ac.uk; or the supervisor, Dr. Victoria Murphy, at: victoria.murphy@education.ox.ac.uk. I thank you for your time and look forward to your response!

[Group 2, 3, 4:]

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

参加调查信息说明书

题目：跨语言的观念表述(Cross-linguistic expressions of opinion)
研究者: Lucy Zhao, 教育系博士生 （PRS DPhil Education

我叫 Lucy Zhao, 我是牛津大学Victoria Murphy 导师的博士研究生。我选择请你们参加这项语言文化的研究课题，是因为你们具有特殊的语言能力。参加或是退出调查，是完全自愿的。调查题目也包括你对有些问题的看法。

研究目的: 这项研究是为了发现语言知识和语言环境是如何影响我们的语言表达,特别是如何影响我们对事物的看法。被调查者可以是美国学生或是中国学生。

研究获得的利益: 尽管参加调查的学生将不会直接受益，但是，他们可以间接获益于语言的学习和交流。这项研究着眼于找出语言环境是如何影响人们的表达方式和文化交流，从而改进语言教育方式。最终的目的是要改善在以讲英语为母语和讲外语的人们之间的文化交流。

研究的风险: 对参加调查者没有任何风险。

调查研究的程序: 如果你愿意参加调查，请你上网回答有关你的语言背景问题（约1分钟）。 然后，您可能会被要求参加下一阶段，它由一个短的（～5分钟），在线调查，包括听录音并回答有关短篇小说问题的。

联系方式: 如果你有任何问题, 请联系项目研究人员Lucy Zhao，电子邮箱: lucy.zhao@education.ox.ac.uk; 或者联系Victoria Murphy导师，电子邮箱: victoria.murphy@education.ox.ac.uk.
感谢你的时间和支持!

Appendix 2: Consent form (All Phases)

[Group 1:]

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk



Participant Consent Form

Title of Project: Cross-linguistic expressions of opinion
Researcher: Lucy Zhao, PRS DPhil Education

Declaration of Consent:

- I have read the participant information sheet and have had the opportunity to ask questions about the project and receive satisfactory answers to questions.
- I understand that at any point during the research, I may notify the researchers and withdraw from the study without penalty, and any data recorded will be discarded.
- I understand that this project has been reviewed and approved by the University of Oxford's Central University Research Ethics Committee.
- I understand that my personal data will remain strictly confidential and anonymous in accordance with the provisions of the Data Protection Act of 1998, kept securely in a password-controlled-server; and what will happen to the data at the end of the project.
- I understand that I will have the opportunity to review and comment on any analysis before publication.
- I understand how to raise a concern and make a complaint, and agree to participate in this study.

☐ I agree to voluntarily take part in this project.
☐ I confirm that I have read the participant information sheet.
☐ I am at least 18 years old.

Name of Participant: _____
Email: _____
Signature: _____ Date: _____

Name of Researcher_____ Signature and Date: _____

[Group 2, 3, 4:]

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

参加调研同意书

题目：跨语言的观念表述 （Cross-linguistic expressions of opinion）
研究者: Lucy Zhao, 教育系博士生（PRS DPhil Education）

同意声明:

- 我已经读过"参加调查信息说明书"，并且我所提出的问题都已得到了满意的回答。
- 我明白我可以在调研期间随时向调查者申请退出。 我的个人资料会被即时销毁。
- 我明白这项调研课题将由牛津大选伦理研究中心申阅。
- 我明白根据1998年制定的个人资料保护法，我的个人资料将会以不计名形式被严格保存在计算机系统中。调研项目结束后，我的个人资料将会得到妥善处理。
- 我明白在调研结果发表之前，我有权力查阅我的资料和提出意见.
- 我明白如何提出不同意见,我同意参加这项调研项目。

☐ 我同意自愿参加调查.
☐ 我确定我已经读过"参加调查信息说明书".
☐ 我是18岁以上.

参加调查者姓名: _____
电子邮箱: _____
签名: _____ 日期: _____

研究员姓名：_____签名和日期: _____

Appendix 3: Language questionnaires (All phases)

[Group 1:]

Monolingual Questionnaire for (name here): _____
Date of Birth_____
Your information will be completely anonymous and confidential under the Data Protection Act of 1998. The name and DOB above will only be used to match the questionnaire to the data set.

1. Are you a monolingual speaker of English? (IE: not proficient at a comfortable conversational level in any other language)
□ Yes. □ No (not qualified for participation in study).

2. Have you lived anywhere outside of America (excluding vacations)? If so, where, when, and for how long?_____

[Phase 1, Group 2:] [Phase 2, Groups 3 and 4:]

(actual version, followed by translated English version)

EFL 调查问卷
姓名:_____
出生日期：年_____月____日____

这份问卷是无计名试卷。在调查期间，你所有的个人资料都将受到 1988 资料保护法的保护。

1. 你能同时说流利的中文和英文吗? (例如：可以进行日常对话)
☐ 是，可以说流利的中文和英文 ☐ 只能说流利的英文 ☐ 只能说流利的中文

2. 你是不是在中国以外的国家和地区居住过(不包括去国外度假和旅游)? 如果有，请列出时间，地点，停留天数：_____

3. 除了中文和英文，你是不是还会其他语言？
☐ 不会 ☐ 会(请列出): _____

4. 你是什么开始学说英语的?
☐ 出生在双语家庭，从出生起就学说中英两种语言。
☐ 中文是母语, 英文是后学的，初学年龄：_____
☐ 其他: _____

5. 请从下列问题中选择最准切的答案.

	英语	中文	中英双语	(其他语言:请注明)
你的学校使用哪种语言教学?				
你的语言老师的国籍是哪里?				
你在家里时说哪种语言?				
你在外涉交时说哪种语言?				

EFL Questionnaire for (name here): _____
Date of Birth_____
Your information will be completely anonymous and confidential under the Data Protection Act of 1998. The name and DOB above will only be used to match the questionnaire to the data set.

1. Are you a proficient speaker of both Chinese and English? (IE can speak at a comfortable conversational level)
□ Yes, proficient in both □ Only proficient in English. □ Only proficient in Chinese.

2. Have you lived anywhere outside of China (excluding vacations)? If so, where, when, and for how long?_____

3. Are fluent in any languages other than Chinese or English?
□ No □ Yes (which?): _____

4. When did you first become exposed to English?
□ Both languages introduced simultaneously at birth.
□ Chinese at birth, English acquired later (When?:) _____
□ Other: _____

5. Please check the most appropriate box for each question below.

	ENG	CHI	Both/ equally	(Optional: please explain)
In which languages are you taught in at school?				
Which is your language teacher's native language(s)?				
In which language do you prefer to speak at home?				
In which language do you prefer to socialize?				

[Phase 2, Group 2:]
(actual version, followed by translated English version)

EFL 调查问卷
姓名: _____
出生日期：年_____月____日____

这份问卷是无计名试卷。在调查期间，你所有的个人资料都将受到 1988 资料保护法的保护。

1．你能同时说流利的中文和英文吗? (例如：可以进行日常对话)
☐ 是，可以说流利的中文和英文 ☐ 只能说流利的英文 ☐ 只能说流利的中文

2. 除了中文和英文，你是不是还会其他语言？
☐ 不会 ☐ 会(请列出): _____

3. 你是什么开始学说英语的?
☐ 出生在双语家庭，从出生起就学说中英两种语言。
☐ 中文是母语, 英文是后学的，初学年龄： _____
☐ 其他: _____

4. 你是什么年龄移居美国的?_____

5. 你是不是在中国以外的国家和地区居住过(不包括去国外度假和旅游)? 如果有，请列出时间，地点，停留天数：_____

6.你是不是在美国上过华人办的中文学校?
☐ 没有 ☐ 有 (每周上几小时课?) _____

7.请从下列问题中选择最准切的答案.

	英语	中文	中英双语	(其他语言:请注明)
你的学校使用哪种语言教学?				
你的语言老师们的国籍是哪里?				
你在家里时说哪种语言?				
你在外涉交时说哪种语言?				

EFL Questionnaire for (name here): _____
Date of Birth_____
Your information will be completely anonymous and confidential under the Data Protection Act of 1998. The name and DOB above will only be used to match the questionnaire to the data set.

1. Are you a proficient speaker of both Chinese and English? (IE can speak at a comfortable conversational level)
□ Yes, proficient in both □ Only proficient in English. □ Only proficient in Chinese.

2. Are fluent in any languages other than Chinese or English?
□ No □ Yes (which?): _____

3. When did you first become exposed to English?
□ Both languages introduced simultaneously at birth.
□ Chinese at birth, English acquired later (When?:) _____
□ Other: _____

4. At what age did you move to America?_____

5. Have you lived anywhere outside of China (excluding vacations)? If so, where, when, and for how long?_____

6. Do (did) you attend Chinese complementary school?
□ No □ Yes (how many hours per week?) _____

7. Please check the most appropriate box for each question below.

	ENG	CHI	Both equally	(Optional: please explain)
In which languages are you taught in school?				
Which is your language teacher's native language(s)?				
In which language do you prefer to speak at home?				
In which language do you prefer to socialize?				

Appendix 4: Perception task example trial page (Phase 1)

[Group 1:]

INSTRUCTIONS: You will be presented with a short story. Each page will contain two audio recordings in reference to the story and three questions asking your opinion about these recordings. Push the play button to hear the recordings. Choose one answer for each question.

The short story will be at the top of each page for your reference. You can play these recordings as many times as you like. There is no time limit, but try to answer the questions as quickly as possible. At the end of the survey, there will be two additional questions.

A teacher tells his students he is bringing a surprise to class tomorrow. He describes what it might be, giving hints but not telling them what it is. The following are recordings of a student guessing what it is.

>[AUDIO RECORDING: PROSODIC VARIATION 1]

1. **Student sounds...** (Choose one)

☐ **certain of her statement** [Statement seems direct and deliberate. Speaker believes her statement to be true.]

☐ **uncertain of her statement** [Statement seems direct but tentative. Speaker believes her statement may be true, but might not be.]

☐ **reluctant to present her statement** [Statement seems cautious. Speaker is apprehensive to say what they think.]

2. **Rate speaker confidence:** (Choose one)

☐	☐	☐	☐	☐	☐	☐
1	2	3	4	5	6	7
least			somewhat			most

[Group 2:]

你将听到一个短故事。每一页纸将包括两组针对这个故事的录音,和三个问题问及你对录音的看 法。按下播放键听录音。每个问题选择一个答案。每页纸的上方都有这个短故事做为你的参考。 你可以多次重放录音。虽然没有次数限制,但还是请争取尽快回答。在调查的最后, 将有两个副加问题。

老师告诉他的学生，明天他将带给他们一个惊喜。他描述了惊喜会是怎样的，但并没有明确说是什么。以下的录音是学生们对惊喜的猜测。

>[AUDIO RECORDING: PROSODIC VARIATION 1]

1. 学生的声音… (选择一个)

☐ **非常确定她的观点**[观点一看上去就很周到。讲话的人确信她的观点是对的。]

☐ **不确定她的观点**[观点一看上去就没把握。 讲话的人认为她的观点很有可能是错的。]

☐ **勉强表述她的观点**[观点看上去有些犹豫。讲话的人很担心她所说出的想法。]

2. **给**讲话人的自信程度打分**:** (选择一个分数)

☐ ☐ ☐ ☐ ☐ ☐ ☐
1 2 3 4 5 6 7
最低分 中等分 最高分

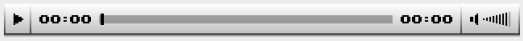
Appendix 5: Perception task condition distribution (Phase 1)

Key: Conditions
DIRECT = D.
Accent on PRN = P.
Accent on V = V.
No accent (front) = F.
No accent (mid) = M.
No accent (end) = E.

Prosodic variation condition scrambling for main task:

I think it's a cat.	P
I think it's a violin.	D
I think it's a cake.	M
I think it's a book.	E
I think it's a dog.	V
I think it's a guitar.	F
I think it's a pizza.	P
I think it's a lizard.	D
I think it's a trophy.	M
I think it's a hamster.	E
I think it's a flute.	V
I think it's a pie.	F
I think it's a goldfish.	P
I think it's a hat.	D
I think it's a butterfly.	M
I think it's a doll.	E
I think it's a bunny.	V
I think it's a painting.	F
I think it's a frog.	P
I think it's a drawing.	D
I think it's a snake.	M
I think it's a watch.	E
I think it's a bird.	V
I think it's a map.	F

A teacher tells his students he is bringing a surprise to class tomorrow. He describes what it might be, giving hints but not telling them what it is. The following are recordings of a student guessing what it is.



1. Student sounds...

- ◯ certain of her statement [Statement seems direct and deliberate. Speaker believes her statement to be true.]
- ◯ uncertain of her statement [Statement seems direct but tentative. Speaker believes her statement may be true, but might not be.]
- ◯ reluctant to present her statement [Statement seems cautious. Speaker is apprehensive to say what they think.]

2. Rate speaker confidence:

1	2	3	4	5	6	7
◯	◯	◯	◯	◯	◯	◯

Appendix 6: Phase 1 Pilot data

PARTICIPANTS

Originally, 67 individuals were recruited. However, participants with incomplete data (i.e. did not complete the task) were excluded. Accordingly, a final sample size of N = 38 was obtained (20 US, 18 CHI). Mean age was 36.76 (SD = 17.81). Given this large SD, age was analyzed as a covariate.

EXAMINATION OF TIMING

No participants were disqualified due to examination of timing. Qualtrics analytics showed that all participants took at least the minimum amount of time to listen to each audio recording (> 3 seconds) before choosing a response.

PERCEPTION TASK

Analysis included three mixed-design repeated measures ANOVAs. The first ANOVA was conducted with one within-subject factor (prosodic variation) tested at six levels[54] (D, PRN, V, F, M, E) and one between-subject factor (sociolinguistic background) tested at two levels[55] (US, CHI), and confidence rating (CR) as dependent variable being measured. Hypotheses are:

1) H1: Prosodic variation of [I + verb] statements affects CR of statements. Specifically, it is predicted that CR will decrease from Direct to PRN to V to no kontrast conditions.

H0: There is no difference in CR of [I + verb] statements between prosodic variations.

2) H1: Sociolinguistic background affects CR of [I + verb] statements.

[54] Refer to Section 5.2.2 for details.
[55] 2 groups; refer to Section 5.2.1 for details.

H0: There is no effect of sociolinguistic background on CR of [I + verb] statements.[56]

While sphericity cannot necessarily be assumed ($p = .001$), results calculated with sphericity vs. Greenhouse-Geiser measurements (=.678) do not differ regarding significance. Prosodic variation had a strong within-subject effect ($p < .01$, $F = 25.88$, $\eta_p^2 = .43$). Table A1 shows that mean CRs significantly decrease from Direct to PRN to V to no kontrast (FRONT/MID/END), thereby supporting Hypothesis 1.

Table A1: SDs and Means of prosodic variations.

Kontrast Placement	Sociolinguistic background	Mean	Standard Deviation
Direct (DIR)	US	6.51	.48
	CHI	5.97	1.08
	Total	6.26	.86
Pronoun (PRN)	US	5.99	.55
	CHI	5.06	1.38
	Total	5.55	1.12
Verb (V)	US	4.14	.76
	CHI	5.04	1.27
	Total	4.57	1.12
No Kontrast: Sentence initial (FRONT)	US	3.86	.70
	CHI	4.15	.94
	Total	4.00	.82
No Kontrast: Within-sentence (MID)	US	3.34	.97
	CHI	2.92	.76
	Total	3.14	.89
No Kontrast: Sentence final (END)	US	3.95	.77
	CHI	3.76	.87
	Total	3.86	.81

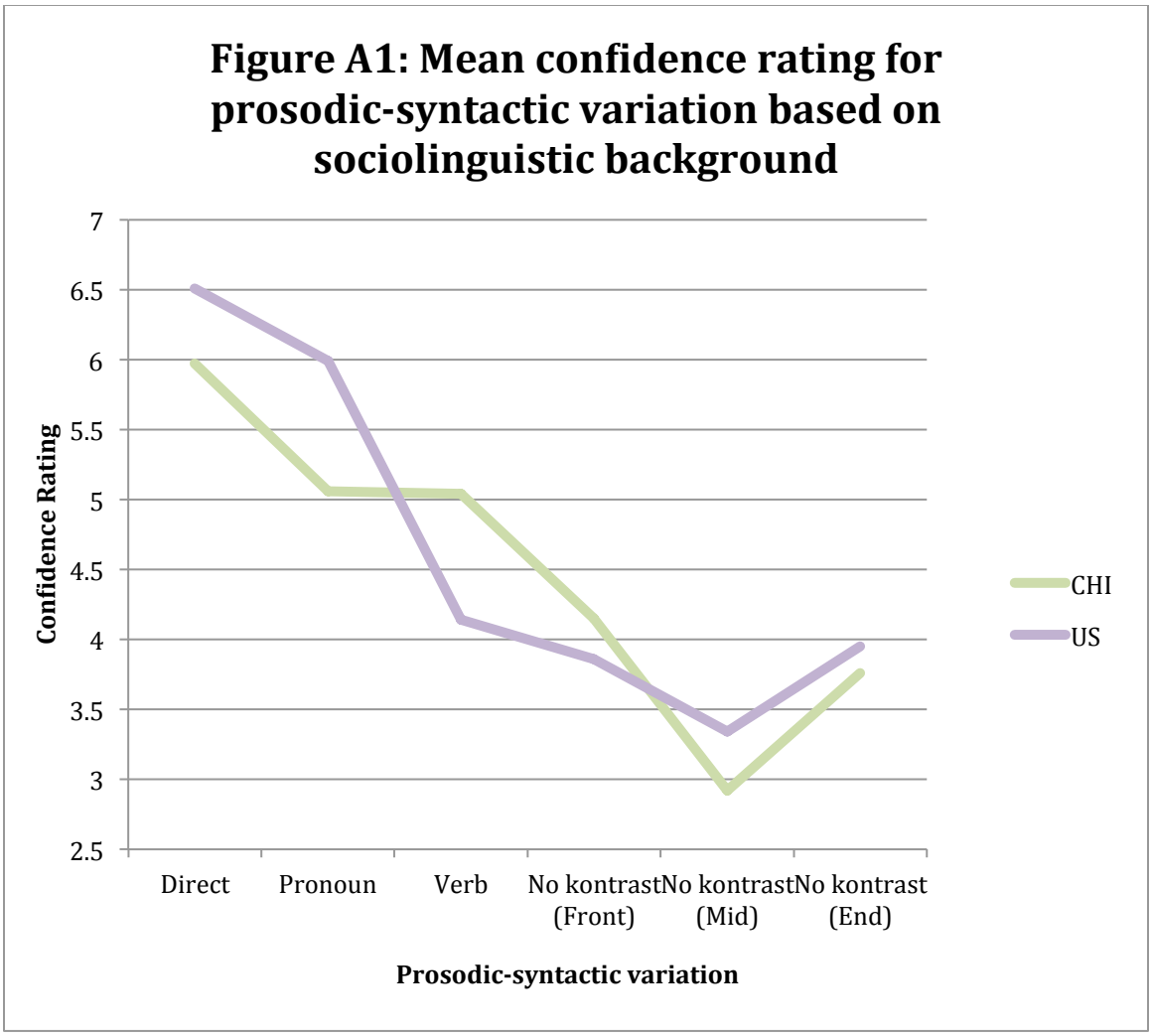
[56] Based on past research on universality of focus-marking (See section 3.3), this thesis predicts that results will support H0 instead of H1 for 2).

Post-hoc paired t-tests illustrate significant differences between all levels ($p < .001$) except FRONT/END ($p = .15$). Because Hypothesis 1 predicts that CR differs only for kontrast variation (vs. mere syntactic variation), given that FRONT and END are both no kontrast conditions, this result does not refute Hypothesis 1.

No between-subject effect of sociolinguistic background ($p = .22$) and slightly significant effect of the covariate, age ($p = .04$, $F = 13.08$, $\eta_p^2 = .12$), appeared, supporting both Hypothesis 2 and the predicted cross-linguistic congruence of pragmatic-prosodic mappings[57]. This renders influence from L1 pragmatic knowledge transfer (RQ 5) irrelevant, for both CHI/US speakers should possess congruent pragmatic knowledge regarding prosodic focus-marking's effects on perceived speaker confidence.

Given large SD of age between-groups, age was run as a covariate but did not interact significantly with prosodic variation ($p = .48$). However, a small interaction appeared between sociolinguistic background and prosodic variation ($p < .001$, $F = 8.88$, $\eta_p^2 = .20$). This is interesting given that age and sociolinguistic background produced opposite between-subject significance levels. The overall trend of CR (high-to-low) appears similar between US and CHI groups (Figure A1), with slightly higher (yet notably insignificant) means for US vs CHI. However, there is an observable inverse of means for V condition where CHI mean surpasses US mean (4.92 vs. 4.25; mean difference = .67).

[57] Note that in the second hypothesis, the effect of culture only relates to Proof of Concept (i.e. universality of prosodic focus-marking), and does not relate to trends regarding frequency of usage that may be more linguistically-specific to tonal vs. accent language characteristics.



Despite this inverse, both groups maintained the same CR trend (D > PRN > V > no kontrast). This, alongside V condition's small between-group difference (.67), suggests that results do not contradict hypotheses.

The second ANOVA was conducted with one within-subject factor (perceived pragmatic function) tested at three levels (Deliberate, Relectant, Uncertain) and one between-subject factor (Sociolinguistic background) tested at two levels (US, CHI). Hypotheses are:

1) H1: Perceived pragmatic function of a statement affects CR of [I + verb] statements.

Specifically, CR will decrease in order from Deliberate to Uncertain to Reluctant conditions.

H0: There is no difference in CR of [I + verb] statements between perceived pragmatic functions.

2) H1: Sociolinguistic background affects CR of [I + verb] statements.

H0: There is no effect of sociolinguistic background on CR of [I + verb] statements.[58]

Mauchly's test indicated that sphericity may be assumed ($p = 0.11$, $\chi^2(2) = 4.36$).

Perceived pragmatic function had a strong between-subject effect ($p < .001$, $F = 46.30$, $\eta_p^2 = .57$). Table A2 shows that CR means are highest for deliberate, then uncertain, and lowest for reluctant.

Table A2: Mean confidence ratings and standard deviations for perceived pragmatic function based on sociolinguistic background

Perceived pragmatic function	Sociolinguistic background	Mean	SD
Deliberate (D)	US	6.13	.36
	CHI	5.79	1.03
	Total	5.97	.76
Uncertain (U)	US	3.98	.66
	CHI	3.69	.79
	Total	3.85	.73
Reluctant (R)	US	3.05	.77
	CHI	2.86	.84
	Total	2.96	.80

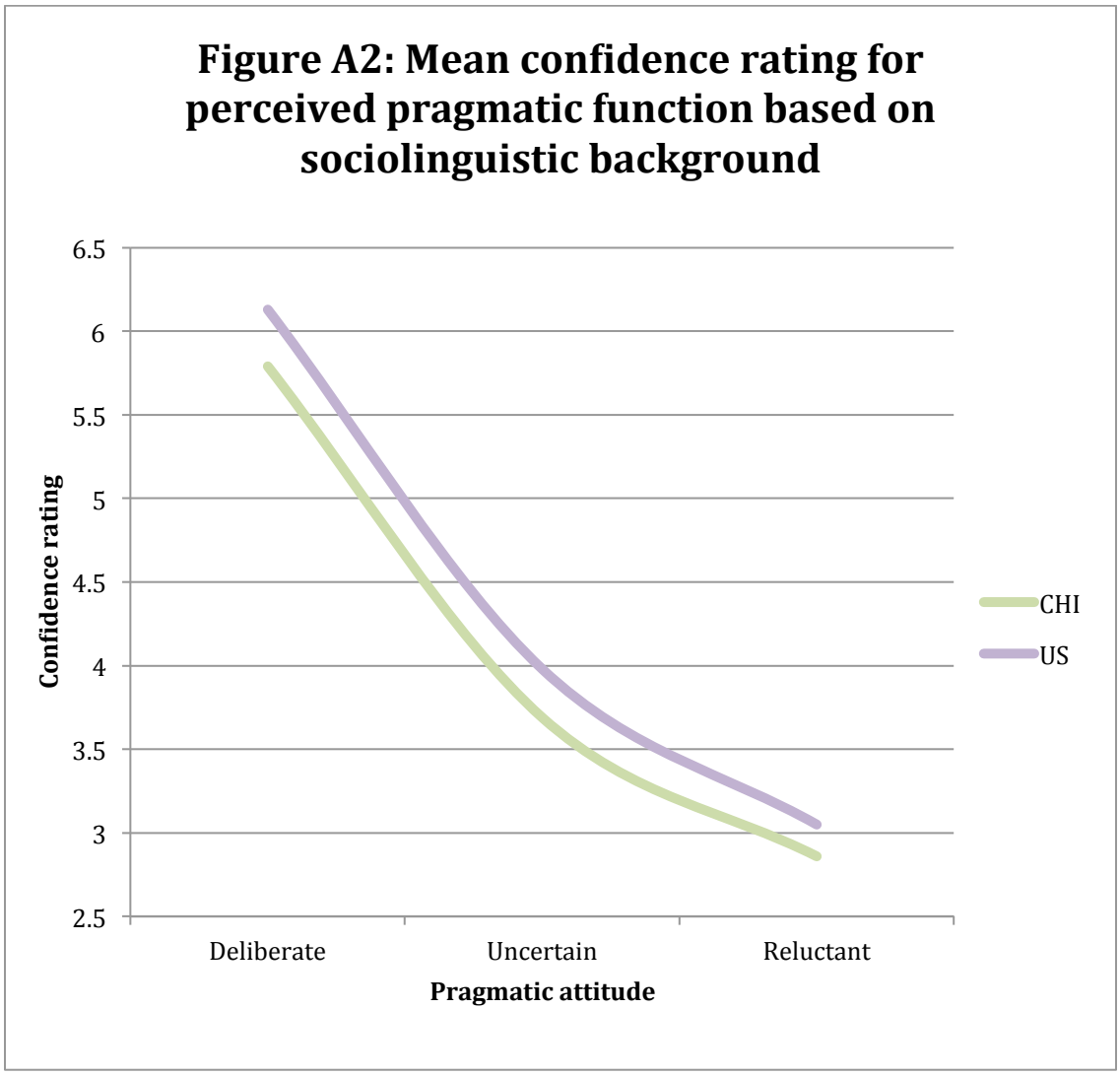
Post-hoc paired t-tests showed significant differences between all levels of perceived

[58] Based on past research on universality of focus-marking (See section 3.3), this thesis predicts that results will support H0 instead of H1 for 2).

pragmatic function ($p < .001$), thereby supporting Hypothesis 1.

For between-subject effects, sociolinguistic background was insignificant ($p = .06$) while covariate of age was barely significant (.04), mirroring ANOVA 1 and similarly supporting Hypothesis 2.

No interactional effects were found between pragmatic function and age ($p = .77$) or sociolinguistic background ($p = .84$). The parallel movement between-groups is visible in Figure A2. A visible difference of means appears, with a pattern of higher US and lower CHI CRs (4.45 vs. 4.06; mean difference = .39).



Although there was no strong effect of sociolinguistic background on CR, its .06 p value borders significance level; interestingly, this near significance mirrors expected differences between US and CHI sociolinguistic environment: the modest, face-saving nature of CHI may lead to increased caution toward giving out positive ratings in comparison to their expressive US counterparts (Heine et al., 1999).

A third analysis was implemented as a relationship check between CR of prosodic variation and perceived pragmatic function. Each pragmatic type was assigned a value based on mean CR calculated in Analysis 2 (see Table A3). Next, each participant's average CR per prosodic variation was tallied via average CR of pragmatic function chosen for that contour. This ANOVA was run with the same within-subject and between-subject factors as the first ANOVA, only with newly calculated CR values for prosodic variation based on pragmatic function CRs. Hypotheses are as follows:

1) H1: CR of [I + verb] statements follows the same prosodic variation trend as Analysis 1, when calculated via CR of perceived pragmatic function.

H0: There is no correlation between results of CR from Analysis 1 vs. Analysis 2.

2) H1: Sociolinguistic background affects CR of [I + verb] statements.

H0: There is no effect of sociolinguistic background on CR of [I + verb] statements.[59]

While sphericity cannot necessarily be assumed ($p < .01$), results calculated with sphericity vs. Greenhouse-Geiser measurements (=.79) do not differ in significance. For within-subject

[59] Based on past research on universality of focus-marking (See section 3.3), this thesis predicts that results will support H0 instead of H1 for 2).

effects, computed pragmatic variation appears significant ($p < .001$, $F = 27.87$, $\eta_p^2 = .44$), supporting Hypothesis 1. Means and SDs for calculated pragmatic variations are presented in Table A3.

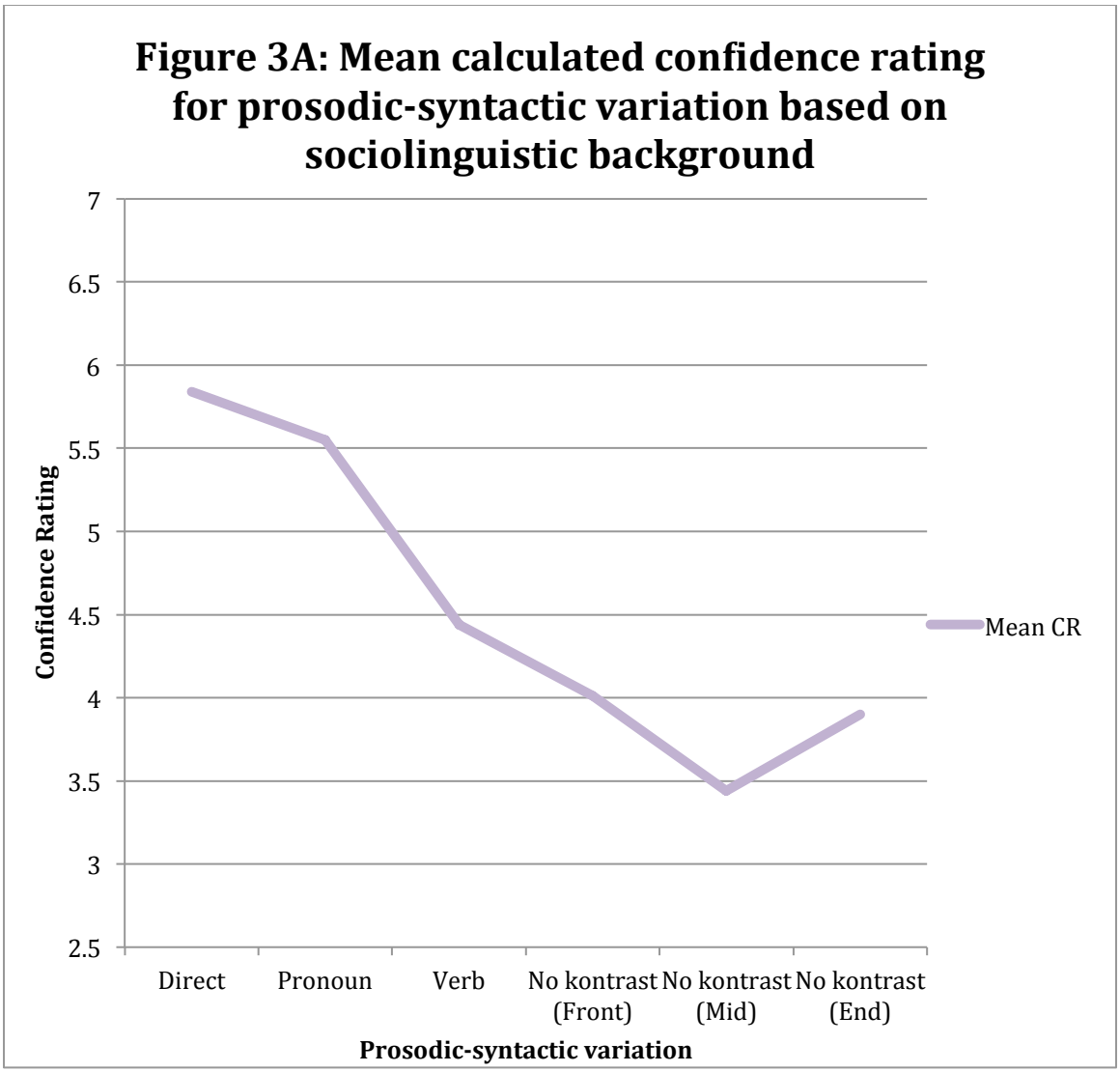
Table A3: Mean confidence ratings and standard deviations for prosodic variations computed via perceived pragmatic function

Kontrast Placement	Mean	Standard Deviation
Direct (DIR)	5.84	.29
Pronoun (PRN)	5.55	.60
Verb (V)	4.44	.82
No Kontrast: Sentence initial (FRONT)	4.01	.65
No Kontrast: Within-sentence (MID)	3.44	.25
No Kontrast: Sentence final (END)	3.90	.60

Post-hoc t-tests show significant within-subject variation ($p = .001$ between V/FRONT; $p < .001$ for rest) except for MID/END ($p = .27$); similarly to Analysis 1, the no kontrast condition of MID and END indicate that this does not contradict hypotheses.

No significant between-subject effect of sociolinguistic background ($p = .46$) or covariate of age ($p = .74$) appeared, supporting Hypothesis 2 across all analyses.

No interactional effect was observed between calculated prosodic variation and age ($p = .12$). There was a weak effect between sociolinguistic background and calculated prosodic variation ($p < .001$, $F = 12.66$, $\eta_p^2 = .27$, $OP = 1.00$) (Figure A3), paralleling Analysis 1.



Given FRONT/MID/END are similarly no kontrast conditions, the most apparent interactional effect on kontrast-based prosodic variation appears to be for V, which CHI participants rate as higher in confidence than US. Nevertheless, this difference appears insignificant via within-subject analysis and post-hoc t-tests, and calculated CR trends remain the same overall as predicted. Therefore, neither hypothesis is contradicted. Thus, calculated CR from Analysis 2 via factors of Analysis 1 provides beneficial correlative evidence for universal Proof of Concept of pragmatic-prosodic mappings for [I + verb].

CODING OF SPEAKER CONFIDENCE

Sociolinguistic background's affect on cognition was investigated through participants' definitions of "speaker confidence". Remarks were classified into five categories (Table A4).

Table A4: Meaning of "speaker confidence" for US vs. CHI participants

	Direct/firm tone of language	Belief of speaker toward statement	Belief of listener toward statement	Appeal/ Reputation of speaker	Comfortableness of speaker to share statement
US	11	9	0	1	4
CHI	5	3	6	5	0

"Directness of language" monopolized US remarks by comprising 44% of total remarks, vs. only 26.32% for CHI. Both US/CHI participants used words like "firm" and "strong," but while US participants deemed it "direct" and "deliberate," CHI participants expressed the same sentiment indirectly (ironically), expressing it as "us[ing] less 'I think' and other terms" and "say[ing] with less feeling and more absolutes." This frequent US mention, along with each culture's attitude toward directness/indirectness, coincides with each culture's philosophy toward value on self-expression vs. saving-face.

"Speaker belief toward statement" describes what one CHI participant called "empirical certainty of contents" in the statement, and is comprised of 36% US vs. only 15.79% CHI remarks. US remarks describe this as whether the speaker "*believes* [the statement] to be true" or "correct," with one participant stating, "Knowing you know the answer [or truth of statement] will give you more confidence." This US true/false dichotomy runs parallel to its absolute, brazen nature, yet its emphasis on certainty of one's own belief instead of "empirical" truth places individualist focus on knowing one's self above all, and with subjective belief one's opinion can never be faulted (refer to literature review). Fewer CHI remarks here may coincide with CHI's

notion of saving-face and reluctance to commit to a statement, for fear of limiting manoeuvring of future language and perhaps a malleable, changeable disposition of self (Norenzayan, Nisbett, & Smith, 1998).

"*Listener* belief toward statement" includes zero US vs. 6 (31.58%) CHI mentions, thereby being the largest CHI category. CHI participants described this as varying "degrees of trust" toward speaker, or how "convincing" and "reliable" a speaker seems leads one "to believe his opinion." One participant even noted that more speaker confidence means the "more [he] suspect[s]" whether the speaker is trustworthy. Perceived truth of a statement may depend on the listener's own beliefs, with one participant stating, "maybe I promote his views, maybe I doubt his views, depending on my own beliefs." CHI participants' definition of speaker confidence appears other-dependent, coinciding with a collectivist other-centric focus. A dearth of US mentions coincides with individualist self-centric focus, as viewed by the numerous US definitions of belief of one's *own* statement; this strongly coincides with Choi, Nisbett and Norenzayan's (1999) results on internal vs. external behavioural attribution (See section 2.1.3.1.).

"Speaker appeal/reputation" is again an other-constructed definition. Accordingly, data shows only 1 US vs. (4%) 5 CHI remarks (26.32%) in this category. While the one US remark describes this as how "knowledgeable" a speaker seems, CHI remarks viewed it more as how a speaker seems presentable to a *listener*. CHI remarks refer to possessing a "good ego state" which will "influence...and appeal to others;" the word "appeal" appeared in three remarks. Other CHI participants stressed its importance regarding "dignity" and "identity," reflecting the collectivist emphasis on appearance and judgment, and furthermore strengthened by no US mentions. Furthermore, CHI attitude toward identity as other-dependent corresponds with the aforementioned work of Choi, Nisbett and Norenzayan (1999) and Norenzayan, Nisbett, and Smith (1998), particularly the interdependent notion to determine one's identity through being situated in a broader social context.

"Comfortableness of speaker to share statement" is dominated by US remarks, with 4 (16%)

to 0 CHI remarks. US participants describe this as "comfortableness of sharing" and "willingness to reveal/share without fear of consequences ([of] what peers/teachers might think)". This last remark especially contrasts CHI remarks above regarding what others may think. US participants call this notion "self-assured," or the individualist's right of self-expression without fear of judgment expressed in the literature review.

The question, "Do you think the speaker has a standard or normal accent [and] why?" was included to confirm accent typicality of recordings. 100% of US participants answered "yes," with one even identifying it as "standard Midwestern... north of the Ohio river," or the "newscaster's accent." Six CHI participants raised concerns about naturalness of speech; however, the question may have been misrepresented, as 2/6 said unnaturalness was due to strange emphasis, as "emphasized statements may appear biased." Another raised concerns due to statements "being reread," which was due to methodology. Such remarks regarding "accent" appear to be more critique of prosody manipulation than of accent typicality. Overall, only 3/18 seemed to find fault in the speaker's accent; one even mentioned that given dialectal variation there may be no "standard" Chinese accent. Given the apparent lack of one "standard" Mandarin accent, an 88.33% standard accent approval appears valid for confirmation.

For the open-ended concerns/suggestions prompt, two US participants suggested that the 1-7 CR scale be clarified as to which value denotes more/less confidence. One individual stated that while he assumed 7 to be highest and 1 lowest, not all participants may assume the same. This will be clarified in the main phase' instructions.

Appendix 7: Sample selection of transcriptions (10%)

Numbered as x.x: Participant #.Prompt #. [I + verb] constructs are underlined.

1.12: Um: (1.1) just based upon that first ↑thing is uh I think that she's probably a person that likes to move ↑quick and when you move quick you make mistakes um: like the one that she made where she didn't put the right un date in and she probably isn't a great proofreader. I probably wouldn't hire her to be my publicist or: any type of person like that so sounds (1.0) pretty if I had to make ↑assumptions that's the assumption I would make.

2.10: I ↑think Charlotte should make the best effort to: contact as many people as she can and ask them to contact as many people they can to get that date changed so that: there's no confusion in: the: uh in the date and time of the party. That way: everybody can be happy. And I ↑think that: that's the best way to do it.

2.12: I ↑think Charlotte is very scatter-brained and: not very reliable as a person. Although it is natural to make mistakes um: she should really double check her work if something is as important as this. And I ↑believe that: she should really learn from this mistake. And I ↑feel that she wouldn't have if she- I ↑feel that she would have done this earlier.

2.15: I ↑think Gary might be a little bit irresponsible. But it is normal for people to: sleep in a little bit sometimes. And I ↑think that he might be a little bit immature. Um but that's something that can easily be corrected. And he should work to do- towards doing that.

3.8: ↑I don't have the attention span for anything longer than a half hour television ↑show. So: if it's ↑not modern family I don't really pay much attention. I just don't have the attention span for it. I could watch ↑sports because I come and go: and don't really care about the plot and don't get all that interested in it. But: in terms of a (1.2) actual scripted television I just don't care.

3.12: Charlotte should have proofread her invitation: so you ↑could say that she's kind of careless. But if it was ↑just a ↑date that she wrote on incorrectly I think that she could uh (1.3) that's an understandable mistake. But if she actually wrote the day: on wrong I think that's pretty careless.

3.15: Gary either doesn't know himself very well or is a- a little bit inconsiderate little bit arrogant. He oughta know if he can't get up in time to meet somebody at eight o'clock in the morning or: if: he ↑doesn't know that about himself that's uh that's something he should learn. Uh: I'm- I'm just not very sympathetic to Gary. I- ↑I have been the one kept waiting more often or not because I'm a morning person and very prompt.

4.1: Well probably the typical answer is stuff like the uh internet famous honey badger and stuff but uh (3.4) ma- maybe something super deep diving like a sperm whale comes up (1.2) up into to the surface of the ocean and then uh (1.1) close to a mile down to eat squid. That's a lot of uh pressure to get used to I imagine. Um (1.2) also a lot of dark must be spooky down there.

4.13: Uh call her and apologize I suppose. Uh (1.1) don't um someone who's making breakfast plans at eight um I'm imagining it's a weekend uh um I guess my right as a (.) listener here. Um (1.5) probably needs to be a little fuzzy on them. Um eight AM is a lot to expect out of somebody. Um oh um in that case a uh call with an apology and maybe a uh "I'm gonna dash out meet you at nine thirty" I ↑expect more than sufficient.

5.1: Oh um: that's a tough question. Uh: I ↑guess one of them would probably be the squirrels around here at the University of Michigan. They're very brave in going up to people. Um but uh: I ↑guess on a more global ↑level um: (1.0) I ↑feel like it might be some sort of animal that's not afraid to get close to people elephants or something and maybe it's sort of leading to their extinction ↑because they're not afraid of humans.

5.10: Well I ↑guess it sort of depends on what day it is right now. Um I ↑guess if (1.0) today's Wednesday and it's this coming Saturday or supposed to be um: Charlotte if ↑I were Charlotte and I were free Friday night I would switch the party to Friday night and sort of pretend that nothing happened. Uh: but if she's absolutely not free on Friday night then um I don't know call each guest individually. I hope there's not that many.

5.14: Uh if Gary has issues with morning um meetings or meeting people for breakfast he might suggest meeting someone at nine or at ten or later in the day especially if that's a habitual thing that he does. Um mm for all I know Gary might work ↑night shifts and ((then could)) not really get enough sleep. Um he could also set two alarms. He could set an alarm across the room.

6.7: Uh Marsailles is a really good place to travel in the summer. It's not quite as dirty as ↑Paris and although people say it's really dangerous it doesn't seem all that dangerous to me. Uh it might help that I'm really ↑big but uh it has really nice (2.2) coastal valleys. They call them fjords in France but they're not fjords. But you can go diving in the water and it's fuckin' fantastic.

7.7: Uh: a very good place to travel in the summer would be the south of France where the beaches are not as good as the beaches in ↑Los Angeles but the people are beautiful. The buildings are

beautiful. The architecture is a mixture of French and Italian and so is the food which makes it all the better.

7.16: Uh I ↑think that Jimmy (2.2) should (1.0) ↑think a little ↑more. Um (1.1) he shouldn't just (1.0) take what his father says uh: (1.0) to ↑complete heart because um even though it's ↑cloudy it doesn't mean it's going to rain. Um he can- but it ↑sounds like Jimmy's probably a little young and: he has to make these kinds of ↑assumptions to uh the ↑future make his own mistakes.

7.18: He asks questions. A lot of people don't ask questions. They just uh continue on (1.2) uncaring or they just buy into what people tell them so I think Jimmy's ↑probably going to grow up to be a nobel ↑laureate or ↑better (1.5) and I uh: as long as he keeps asking questions and keeps uh steps- steps lightly um doesn't always buy into the uh the things people tell him then I think Jimmy's gonna be okay.

8.11: Another way to solve the problem is to have- is to for her to move the party to Friday and: try to reschedule everything that- that she had planned and: and just go with it that way. Um: another way would be: (1.7) I don't know I guess it- it's either contact everybody and make sure they know it's Saturday or have the party on Friday.

8.12: Uh: Charlotte sounds I don't know it sounds like an- an- an in- an innocent mistake. Um (1.1) you know no matter how many times you double check everything mistakes can always creep in. Um: now if Charlotte does this uh if this is typical behavior for her then friends probably already know and they're probably calling her to find out "is the (.) is the party really on Saturday or is it Friday? Or is it really on Friday instead of Saturday?" Um but yeah it- it (1.7) again without having more information about her it's- I can't condemn her one way or the other. It sounds: pretty typical to me.

8.15: Mm: I- I- no ↑big impression. I don't know if he's young or old or: whether he's a chronic person that's ((there he is or)) chronic problem of- of not getting up in the morning. Um: from the one story it's difficult to tell if this is- if this is a habit of his or just uh just a: a one time thing. Um: I'll give him the benefit of the doubt for now.

9.5: Space travel is: uh: a pretty cool thing to think about. I don't ↑know how realistic it is but I ↑think that it would be pretty cool to take a space ship and: fly through space and see nebulas and: see other planets but I think (1.0) that landing on another planet would be pretty scary because you might not be able to breathe the ↑air and there could be aliens there would want to eat you.

9.7: In the summer it's good to go up north because if you go south it's too hot um and too sunny. So going up north is good because you get the perfect temperature and: good weather and: (1.0) if you go up north you also will get to a ↑lake and you can go swimming in it and that's (1.0) nice and cool: so you have a good (.2) mixture of hot air and cool water.

10.8: Um I ↑think that we're seeing a- a stretch in acceptable TV shows and- and I ↑think there's a large amount of the populace that buys ↑into them (1.0) because they've got into this ↑routing of watching- watching television as a staple for not having to think about whether or not their happy.

11.6: If it's about real ↑people and things that they confront and things that they ↑do and things that they enjoy and things that scare them and if it has some nice action and conflict and tension in it and: if: (1.3) it tells us something about what it's like to be a human being then ↑I think it's really interesting.

11.9: I ↑think it's the best thing to be. Unfortunately I still eat chicken and fish. (1.0) But if I were ↑really a good person I'd stop eating that as well. (2.0) It really bothers me what people do with animals. (1.7) Imp- but the ↑chicken I eat comes from Whole Foods where they get to run around.

12.6: Um that- I ↑think that it's also pretty different based on people's different um: I guess their different s- can't think of the ↑word what they lived through in their ↑experiences ha. Um different experiences kind of growing up and uh what kind of relates to you in general.

12.14: Gary can uh set a couple alarms. He can plan better. He can push back the breakfast. If he knows he's gonna have a late ↑night he can try and get up earlier have um different methods of waking him up. He could have uh he could have to ↑carpool so Jane would have to come pick him up.

12.15: I ↑think Gary's pretty tired. Um: I:↑think that maybe he's made a mistake. He's tired. He has things to do. He's kinda worn out even though I'm sure he loves breakfast or else he wouldn't have made a breakfast plan. Um he: is probably pretty sorry about it. So I ↑think he's just a good ↑guy who loves breakfast food probably dark-haired women.

12.16: I ↑think he's- Jimmy is just young and naïve and he is just trying to ↑connect to things that he has learned from his father who should have told him differently. Um: but I think for kind of like a mind that's: still trying to learn things it could make sense for him to think that every time it's cloudy it's gonna rain.

13.1: Bears are the bravest animal because they're big and brown and scary. And if you've ever seen the office Dwight likes bears. He likes beets and bears. I like bears a lot. Bears are ↑cool. I saw one up top in the Porcupine Mountains one time.

13.9: ↑I personally like meat. But I actually am pretty vegetarian because I usually don't cook meat. It's way too hard to cook meat slash too expensive. So I ↑think it's great one. It's really healthy. One of my friends he: became a vegan then became a vegetarian primarily to: (1.0) lose weight and like get in shape a little bit and it worked so well for him.

15.18: Uh as ↑far as I know ↑Jimmy seems like a really nice kid. Uh we don't get to know a whole lot about him. But he seems curious. And he asks his father questions about science and the world. So I have a pretty positive view of Jimmy even though I don't have every much information about him.

16.6: Um: I ↑think starting with the story line itself and uh carryin' it through the- the actual subject. And how you relate ↑with the subject um: makes a very very good uh- uh I just lost there. Uh I'm very good what is that baseline or- or subject so (1.4) I apologize on that but um (1.2) ↑I think that that if you- if they're able to carry the subject all the way through the entire novel it- it- it turns out really well.

16.14: I think ↑Jimmy needs to work on his conclusion a little bit more because: there will always be clouds in the sky and it doesn't necessarily mean that there's rain so I: so at that point I ↑think his parents need to kind of help him out a little bit more in- in straightening out that issue.

17.4: I ↑think it's very important to learn a second language because as you go through life you'll realize how much of that language is rooted in other languages and in infor- in- in detail that you'll run across. You'll recognize all sorts of words and ↑root words. And you can- for instance my job right now I'm able to read cases that are in Spanish.

17.5: That's an interesting question. Um you know I haven't paid much attention to the: the space program right now so I'm not really sure what's going on where we've gone and what we've been doing so uh I'd- I'd be interested to see whether you know what continue in furthering it.

17.17: One of the things that I find useful with my children that I do a lot is sometimes I don't know the details about why thing happen I use the internet and the web and lots of maybe YouTube videos to help explain things because I'm not a weatherman so I probably couldn't explain it well but there's lots of resources out there to help explain: uh what the rain the process of rain the different types of clouds.

18.4: Oh learning a second language is so: important. I studied Latin in: middle school and high school which um: I think helped a lot on my SATs but didn't really help me in other ways. Um: I learned ↑Swahili because I went on a study abroad to Kenya which was cool.

18.6: Um: okay so I've been reading a lot of books lately. [Clears throat] ↑I think what makes a story interesting- I like- first of all I like stories that you can't: you can't guess the ending: um right when you're reading the beginning so you don't really know what's going to happen. Um I also- for a story to be interesting for me: there has to be some sort of like connection. Like if I don't feel like I'm invested in the characters or in the plot then I usually just don't really give a shit about what's going on.

18.15: Uh: I don't know I understand where Gary's coming from because I've definitely stood people up before and it feels really shitty. But you can come back from it Gary it's okay. You can: be a better friend. Um: ↑initially I thought he was kind of a s- like a s- asshole standing this person up but it's not a judgment.

19.1: The bravest animal (1.5) oh my god that's a hard question. Are animals brave? Becau- I don't even know if they are ↑brave. Because (1.5) brave implies: that you're you know getting your courage up. And I don't- I don't know if the animals get their courage up. I ↑feel like they: are so: naturally: (1.0) intuitively responding and moving and doing ↑maybe brave things all the time but I don't ↑know if I'd call it brave.

19.4: My opinion is that it is good to learn a second language. And um you know I- I wanted ((chose)) I have a little five year old. I put him in like a Spanish immersion school like right when he was- cause I ↑feel like it's so great to just like get that mind you know expanded because a second language is just so hard to do to learn. ↑I think for me it's been really hard to learn like maybe doing it later. (1.1) But it's so: great.

19.17: 'Cause Jimmy's father should look into um (1.3) y'know how- he should explain that it's not just any-↑any cloud that um produces rain. (1.9) And he'll have to um (1.0) explain to ↑Jimmy how the raindrops get to a certain (2.1) fullness and that's the moisture in the cloud gets high enough for it to release its rain.

20.1: Uh the ↑bravest animal might be: the dolphin I ↑guess. I just- I'm thinking about this news article I read. Um and my it was a local paper where I'm from in South Carolina. And: uh: it was just talking about these this team of dolphins that saved a scuba diver. I just ↑thought it was really amazing. It's pretty↑risky situation where sharks around.

21.3: History (1.0) is very difficult. No actually I'm sorry let's go back. Geography I ↑think is most difficult subject um: (1.9) for at least for me it was. And it was because I was not focused on ↑geography so I didn't do well with that. Um I ↑did do well with math and art.

21.10: She needs to contact all: all of the um people who got ↑invitations and let them know that the date was wrong. (1.8) Unless she- ↑unless she can change the date [laughs]. If- if that would be: you know not a problem then that might be the best way to go about handling that.

23.1: Um I ↑mean it depends on how many people I guess. She could um she could send out a corrected invitation. That's probably not that hard or: if it was an e-mail you just send out another e-mail. Um (1.4) yeah I mean unless she can move the party to the date that's actually on the invitation I guess that's a possibility too.

24.11: Um: I ↑guess another way would be to move the party to Friday. Um (1.1) that might actually be easier cause well it depends on how much planning she put into this the Saturday date and how many people uh know the Saturday date but that would be ↑another way to solve the problem obviously just to just hold the party on Friday cause everyone thinks it's Friday already.

24.13: I ↑think Gary should (.6) apologize to Jane and: buy a new alarm clock. Or set his phone to go off so he's not late ↑again. (1.5) Cause it's ↑inconsiderate. (2.0) And um [laughs] uh: it's probably not the worst thing in the world because it was just a breakfast meet but um [clears throat] um ↑it's something he should probably try to avoid at least in the future.

24.18: Um my impression is that he's probably ↑young because his name is Jimmy and he's talking to his dad a lot about he doesn't know why it ↑rains. (1.2) Um and therefore (1.1) I ↑don't really have an

impression as to his reasoning abilities because uh I think ↑everyone when they're young has imperfect reasoning abilities (1.0) so (1.1) think Jimmy's a f- (1.5) fine young ↑man.

25.6: Just it has to be something that' personal: or something that people can relate to. Because if it's not then (1.8) I don't know. Or it can just be outrageously funny and totally (.) on the other end of the spectrum so it's like those two things for certain balanced.

25.12: Um: I ↑think Charlotte's probably really nice. She just wants to have a party. (1.0) She just screwed up on the date. You know that ↑happens pretty easily:. Um maybe she needs to ha- like have a ↑friend who can help her be her editor so whenever she does something like this she can be like "Help double ↑check this for me because I need to make sure this party's gonna be on the right ↑time and not the ↑wrong time. You know so and if she had ↑that she'll have a party planner person who's her ↑friend who could help her.

25.16: Uh: Jimmy's father just needs to give Jimmy a little more ↑information on uh: whether the clouds: are: (1.6) uh mean rain or not but um pretty much I think I think he'll ↑get it. You know he'll understand that he needs to just tell his son a little bit more information next time.

26:10: The ↑normal way will be of course try to ((list)) letting people know that "oh that's the wrong ↑time I and what's the time that day is the right one" (1.1) but it's always depends on how much time she left when like just say when she phone it. Maybe if there's too ↑short (1.0) or whatever some situation she cannot let all the people know about ↑it maybe she should.

26.15: ↑Still not enough information to judge one person by this (1.2) and of ↑course he didn't do well in this (1.2) in this (1.6) in this thing but and if he can ↑avoid such- try to avoids (1.2) this bad things happen ↑again (1.0) that would be good.

27.5: I ↑think (1.2) well that's you: like to me I think that's you need to wait very long time until it be like um you know very common ↑things. So right now I think maybe I- I ↑heard people are already sign up for you know some kind of space travel already but I don't know when that will happen so and: I feel it's not safe to go.

28.15: I ↑think he just need lot more uh (1.3) more like manage- time management and the- make sure like uh: don't make this kind similar mistake cause depends on what kind maybe this time is okay cause meet hi- meet him friends or something. In the future for like some very important event we don't want to miss it because that would be very expensive like very probably have big consequences so.

28.18: I ↑think Jimmy is a ((good)). She- he just try to use this something he learn- just he learned try to ↑apply this concept but I thought- I thought ((um they)) just did not use correctly but I ↑think if he can use this uh: if he can just like teach him more: then he should be like good learner I think.

29.12: I ↑think she's not (1.4) a person (1.6) uh (1.2) like (2.9) do things in a very (1.0) carefully in detail or something. So otherwise she won't put the ↑wrong date on the invitation because that's a ↑very important part of the invitation.

30.2: ↑I think the management is very important and (1.1) the very good communication and good uh: skills of uh knowledge. It's uh the three things most important. (1.6) [laughs] (1.4) ↑Management-

management like um: if you work with a group you- you can um: know uh your friends and they like to do or don't like to do something and you- the communication you- is a good like bridge you can. Um it's: more opportunity.

31.5: If ↑possible space travel uh (1.4) should be very fun generally and they should have uh (1.2) uh↑if possible I think I mean could be: a very exciting (1.0) and uh- uh (1.6) uh how to say excited-excited (1.0) travel. (4.6) And uh (1.3) maybe you can explore more nobody can explore conflict situation.

31.6: As a listener for the story you I mean (1.0) I think may ↑most interesting is like (1.1) uh: (1.3) matching the taste for the: listener's interest point. Or some- sometimes maybe (1.3) popular uh topics and then: (1.3) have some uh humanists in that maybe make it interesting.

32.1: Bravest animal I think is ↑dog. That's a um: I ↑think it's it ↑close to human. And they really having a bond relationship with human. And they save people's life on ↑helping people. And so it's a great animal. So yeah it's about dog.

33.3: Hm when ↑I'm in ↑school uh: the hardest subject ↑actually uh was ↑English. Yeah ↑other uh uh I'm doing pretty very well on ↑other: area math science um: a- all areas. I ↑think the problem with my English is when I went to college. Uh it's kind of hard to understand especially uh uh the ↑listening part uh because my: English teacher she's being a never speaking ↑proper English [laughs] so uh I ↑couldn't understand.

33.9: Yeah being vegetarian ↑I don't have a particular opinion about it. People have their choices. But for ↑myself I: I just would not be a vegetarian. And uh but I ↑generally thinking people who uh eat meat they seem ↑happier [laughs].

35.17: ↑I believe uh Jimmy's father probably uh want probably want to use uh ↑size (.4) to: correct Jimmy's [clears throat] Jimmy's uh ↑answer. And: so that require his father to uh: kno- to have some knowledge about why: uh why it's rain. So um yeah that- that's what ↑I can think.

36.11: Another way: I ↑think they have many way to solve the problem. They could use internet or make the phone calls and depend on how they connect to: those party members ((usually)) right and in my general we mostly use email more quickly. The new generation they use facebook:.

36.16: I ↑think Jimmy is a (.7) um: he didn't think too much. He's more: (.6) uh: (.9) more ↑pure and ↑naïve and uh (.5) um (1.1) I ↑think his answer's lovely probably as a kids. As a kid his answer be okay (.3) but as adult probably he need- he need do more (.3) scientific study. (.5) Um: (.) but in the ↑life those type of conversation they build bond between (.4) him and his daddy. I think this is a good conversation (.2) uh to build the ↑bon- bond together.

37.1: What is bravest? Uh bravest this kind of thing has something come ↑thinking come our: ↑intelligence and animal- animal does not have intelligence. Just they do not know any- every animal that's they can't say they're brave where- where brave and: I think simply they do not have thinking ((happily)) some time.

37.17: Mm: (3.3) explain not only have ↑cloud also have some ↑wind. Wind make clouds move and I think ↑that should lead ↑him and children should have enough (1.4) experience and whenever have cloud also have wind and then ↑come the rain. (3.3) That should be good ↑enough.

38.2: Uh: good leader uh first: (1.9) uh: must mm be a person that uh: smart ambitious and: also have power. (1.2) That's most I think most ↑important thing (4.2) [laughs] and ↑also uh have to be the ↑good uh have good uh: (1.7) how to say ↑personality.

38.6: You ↑have to be have some uh good ↑topic first and also: uh: ↑interesting uh: I ↑think for different people maybe different story to different people have uh if they ↑like it or they (1.0) maybe they feel interest. If they don't like it it's not interest [laughs] and but to ↑me I think the good story have to be: uh:.

38.10: He just need send in the invi- um call every person he invite so told them to come ↑Friday not ↑Saturday. He should do that. (1.1) ↑She should ((not)) she do that I think (4.2) so (1.2) mm: and say ↑sorry about that.

38.16: They didn't told me how old Jimmy is yeah and this first question because uh for the ↑children so they (1.1) just didn't uh ↑really understand a lot of things like this like how the ↑rain the ↑formed so (1.0) I ↑think Ji- Jimmy should ask more question about that (1.0) not just cloud uh ↑cloud.

40.4: Yeah I ↑think uh learning a second language is- is very important especially if you wanted uh: want to: like do anything that uh uh ou- out country ou- out side of your country. And ↑especially this time is a global economy. Everything- all the whole world is connected. So um: if your nati- even- even if your- your- your native language is the language spoke and very: I mean popular like English

I ↑think you still need to learn a second language like↑Chinese or: you know French or Germany. I ↑think it'd help you a lot.

40.5: Uh space travel I:↑think space uh will be: I mean is a- is a long shot. I ↑think um it will be not easy because ↑first I think ↑technology is not there yet and it will take- take uh take a long time. And: and it's also I ↑don't ↑think at this time it's economy I mean ↑possible. It will be eh very very expensive and I do ↑not ↑think it's uh it's sa- s- ↑safe yet.

41.18: Uh Jimmy is very ↑quick and he and also he is ↑easy to get a conclusion by himself also I might ↑guess that he is ↑young. He is not experienced. He may not know much may or maybe didn't take many years in school that why he could get this kind of ↑conclusion (1.3) and that's yeah Jimmy and yeah Jimmy's ↑quick and young. That's essentially my conclusion is.

42.18: I ↑think he's (1.2) ↑probably faster than that [laughs] but he um (2.6) hear the answer then he kind of get the conclusion. (1.4) Um but in ↑my opinion I ↑think his ↑father should give more details. [laughs] Just simply ask because of ((crow)) so (1.2) from his father's↑answer he get I ↑think he can get the right (1.3) conclusion [laughs].

43.9: Uh: my ↑opinion is ↑vegetarian- I ↑think will be ↑health right. But the ↑other side I'm not sure is enough for the your body nutrition for all bean diet. So I ↑think (1.0) I ↑think it's be okay for the vegetarian. (1.8) I ↑think I agree with that.

44.16: I think the conclusion is not correct. Because cloudy doesn't mean that it's raining. You know when it's rain: it- it's cloudy so the con- he messed up the concept of (2.6) what do you say in mathematical world (1.1) it's if and only if thing.

45.2: Confidence and respect. (.7) A ↑good leader doesn't necessarily mean they always make good ↑decisions. It's- they're a good leader because they inspire they have a lot of confidence which inspires respect so that people follow them. And obviously you know there is ↑some baseline level of good decision-making skills but in my opinion I think the vast majority of what makes a ↑leader is definitely confidence.

45.4: ↑I think↑everyone should learn a second language or more if possible because I'm of the belief that being more ↑cultured in general which is a consequence of learning a second language is ↑very helpful towards a person's (1.1) overall: intellectual well-being. (1.1) And uh: just that's- it (1.6) it- and (1.5) furthermore it just I think it- it allows better: globalization in the sense that people are more able to communicate with each other instead of being isolated in their own communities.

45.9: I am ↑not a vegetarian. But if I ↑were I: (1.4) my opinion would be that (1.8) I'm doing it for health benefits and ethical benefits as in (1.2) in ↑general eating vegetables and fruit is considered ↑healthy if you I mean if you're not exercising all the time. If you just get the necessary calories the nutrients that's enough.

46.6: Um: what makes a story interesting I think mm: (2.1) for ↑me (1.1) an ↑interesting dose- a really ↑interesting story needs to be: (1.4) the: (2.4) the ↑story- I don't know what makes a (1.4) I don't know how to describe that. The interesting story for ↑me is (1.2) you need to (1.7) I mean ↑attract person. Maybe different people have different you know it- different people may think this story is interesting. Another person may think this story is ↑not interesting. But for me a interesting story is a story with a happy end. I like that.

46.15: I ↑think (1.5) ↑he mm: I mean being ↑late happens to everybody. But I ↑think he needs really to be: (.3) um: to be more: uh more: ↑careful [laughs] ((towards)) not being late. And then [tisks] I don't know maybe he's not very: he's not ↑so (1.0) he- he maybe ↑this appointment is not ↑so important for him. Otherwise he would not be late.

47.9: I ↑think it's very healthy. (2.8) Keep going? Okay I ↑think it's very healthy:. And: it reduce CO2 omission globally. Also: yeah you just eat less meat and you can also (.9) like um use other protein source from beans and uh tofu without killing animals: so it's good.

47.18: I ↑think Jimmy is very young so he's probably elementary school student. Just tell him im- ((him impatient)) can explain everything don't blame him. And he will learn. So in the future he will become knowledgeable. (1.5) And: uh: (1.6) he is very good compared to the previous two guys or girls. (1.0) So generally speaking I don't wanna ↑blame him.

48.9: Hm: I: am not a vegetarian. But my sister is a vegetarian. I ↑think it is: I understand reasons why people choose to be a vegetarian. One if they think it's inhumane to animals. Or: for like environmental reasons. Like trying to eat less meat consumes lots of resources to keep or graze cattle or whatnot. (1.0) I don't know I- I get it but ↑I think uh: I don't eat meat very often but I still eat meat so: I think everything in moderation is good.

49.2: I ↑guess a good leader is really dominant expressive I don't know. Like if he seems really ↑honest (1.1) I guess. (1.2) Because those traits are important cause I guess they'll lets people put their trust in you so: that's how you're able to lead people.

49.5: Space traveling is: (1.2) never really ↑thought that much about space traveling but I ↑think it's something that people should definitely look into cause (1.7) the world itself is more than just ↑earth. I don't know it just ↑interesting to know what's out there than just (2.0) humans like on like planet.

50.12: Charlotte's girl looks like not very very careful about some details the small things but uh it- it's not a very big huge deal. Uh Charlotte is uh kind of like ↑maybe she's hurry doing some multi task at the same time. She just missed the date then really hurry to send something out uh but uh Charlotte maybe can ↑slow down a little bit but ↑that is I think just- ((just- just- just)).

51.3: Um personally I think it's ↑biology. (1.2) I don't know why: because I'm pretty good at physics and math. Uh I ↑find biology um a little bit hard to study because: it relates to- it's- it's always fancy human body but I don't want to understand it in a biological scientific way and it just put me off all the time for studying how things has been digested in your body and I hate it.

51.15: I think I ↑think he made an honest mistake as well. And I ↑think um: I ↑think I can't I- I ↑think he's um (2.1) I can't tell what kind of person he is or what his personality is unless I ↑know what he is going to do to handle the situation. So I ↑think it's perfectly okay for people may have sometimes sleep in. But as long as they handle it well they apologize and they make contact and reach out that's very fine.

51.18: Um: I ↑think he is um he is curious about things. That's why he um want to ask his dad when he something when he saw something. But also as I said earlier I think he has a tendency of overgeneralizing ↑things and not being crit- ↑critical about what he is hearing or what he is accepting. And again it's also difference on his age about like how old he is to able to- to make judgments and to make- able- able to make generalization statement.

52.2: Uh: first my mothers and my- my parents not leader and I think I never touch a leader so close but ↑I remember I watch a ↑book about how to be a leader in the summer ↑vacation. It's: it's I think it's ↑about your: it's about your uh maybe inside because something inside make- make your pu: make your surface different um.

52.5: Oh ↑I think space travel in the future oh I- I uh I remember when I was a child I dream of going outside I mean outside space. I dream of I can drive a: space- spacecraft to- uh ↑I think this- this important for our ↑human because our hu- ↑I think we born this we- we: we begin in this ↑plan but this↑plan- uh the ↑Earth ↑can limit us from this plan so I ↑think it's important for our human.

53.8: I think ↑sure. I think ↑one should learn a second language. Uh ↑personally I learn a third language and fourth language and a fifth language. I think languages uh language are ↑interesting and in another- another ↑language you are accessing ↑another culture and things like that.

53.12: [Laughs] Charlotte is actually uh she should another time pay more attention to things uh especially important things. Uh ↑I think this event wi- uh is actually very ↑important to her so she shouldn't make mistakes like that. Uh that is so oh my god.

54.4: Yea I think learning a: a second language is very important. Well for ↑me: even well for- for now even ↑now my English is not uh ↑good but I ↑think it's help me to uh to- to like open another ↑windows to ↑know the people to- to (1.2) I mean for- for- for the ↑people if you're- you're only to- to learn the things from the translation is- is quite different I mean to know the people to know the- the nother country people I mean by their by their ↑own language.

54.9: Well I think this is a personal choices and a habit. I mean this is uh I ↑think it's good. I mean in old days uh people already have a lot of I mean uh: eh more ↑food and um: I mean everything is ↑enough for their for- for: I mean to satisfy his healthy so I think ve- ve- vegetarian doesn't mea- uh: no need to- to have a ↑meat to well just forget it ↑okay?

54.11: Maybe have a two party yeah Friday and Saturday [laughs] or no. Yeah I ↑think: throwing the uh t- I ↑think just now I already a- answered the question ↑right I mean to just call the friend and- and ↑see the possibility to (1.0) or maybe ↑tell them "I'm sorry. It's w- it's incorrect should be Saturday." I mean just tell them the correct date.

55.3: Uh the hardest uh school subject for ↑me maybe: I think you know ↑my major is about ↑electrical engineering so maybe for me it's the ↑electromagnetic theory the Maxwe- Maxwell ↑Equation you know a lot of magnetic equations on the blackboard. You will see it. You will be ↑puzzled by its uh ↑derivation and the proce- procedure how to derive the equation. That's I think the most hard- hardest question.

55.14: ↑Maybe I think he can mm: just uh uh: make a ↑clock you know set a ↑clock a- at ↑before eight c'clo- before his scheduled time and get up ↑early and then he can manage all the procedures I think. (1.7) Get early is really hm ((usual)). Get up ↑early is a really good habit and can make you prevent a lot of troubles.

56.9: About religions uh: and I- I don't know about their belief. Uh: ↑I ↑we have the uh: specials restaurant for uh: that kind of people that don't eat meat uh: uh but recently the school didn't let us to enter that restaurant. I ↑think uh: we can't talk about s- uh: about ↑their belief.

56.10: That's complicated. Uh he should: (1.0) ↑I think she should: uh let his family to help him to call the uh: call the customers that he invited one by ones. It is: uh un-polite to uh his customers. Uh: (1.0) he have to uh spend the time to uh spend time doing it.

57.11: Um: she needs to change his change her school uh to F- to ↑Saturday and: and make some- make something- make some ↑other things to do on- on- on ↑Sunday and uh because they the ↑invitation because they- the ↑activities changed to- to Saturday and: uh if- if he's: uh if she's still use the: the- the old schedule um: she's not- she's not gonna join the activity on time.

58.16: I think Jimmy is a very sm- bright very- very ↑clever boy because his- his ↑father told him em it must be- it must have ↑clouds when it's rain so he can prove uh ever- every- every whenever- whenever it rains uh there must have- there must be ↑some clouds so ↑I think he's very clever and he's- he's ↑good at to con- to inclu- ↑clude some things.

59.3: I ↑think this may be the uh: (1.8) English [laughs]. Yeah ↑I think English is difficult for me because the- our environment for studying another language is very limited and we can only we ↑pay more attention on the uh ↑paper writing or reading and but not the: speaking or listening so so: so we have very (1.0) very strange accent.

59.10: I ↑think Charlotte: first has two choice. First he change his plan from Friday to Sat- from Saturday to Friday uh so he uh: does not need to- she does need to uh: correct or tell his friends the thing. And second maybe second he has right now to uh tell her friends that their mistake.

60.1: That's a tough question. Since I'm not a leader I will just guess. I ↑think a leader maybe have should have some- something called ↑leadership. Uh he need to tell uh very very ↑clear ↑direction to

his team- team- to his ↑crew and to guide them to the right ↑direction and to keep this term very um ↑firmly and: have a very (1.0) uh how to say very ↑strong relationship.

60.8: Ah current TV shows actually I- I don't watch quite a lot of TV shows. Recently my ↑wife watch some TV shows since (1.1) uh in ↑Chinese comedy "Baba qu na'er". It should be "Where I'm going ↑dad". I ↑think that's uh interesting. It's a- it's kind of family show. Uh: my wife like it.

60.16: As a ↑weatherman hehe this is a very I think: a very interesting ↑question. (1.1) I ↑think Jimmy is about maybe a very little young kid so his conclusion could very be ↑fine for him. Um to ↑us it's: ↑cloudy and rain has do now a very very very close ↑relationship. There are clouds maybe there ↑will be rain and (1.0) uh ↑some clouds cannot (1.0) just ↑cause rain.

61.5: Space travel in the future um: I can't really know. I- I don't really know cause uh: we are living in a- in an era that everything is developing really fast so maybe ↑next week we can have some new method and maybe after one ↑month we have other new technologies so uh you can't really tell but I ↑believe (1.6) maybe- maybe we- we have the chance to see: new aerospace- space ship that can bring everyone to the space.

61.9: Well first of ↑all I myself is not an- not a vegetarian (1.2) but it- it's truly okay you know for people ↑around me to be one and um: I don't really ↑see the necessity you know (1.0) not to eat animal uh I mean meat cause we are human being and uh: (1.0) I think it's our ↑nature you know (1.5) to feed our self you know with ↑any kind of food.

61.12: I ↑think she's (1.2) careless because you know ↑this to hold an ↑event you have to be: very careful because you are maybe contacting ↑many many people and if you make any mistake this can be inconv- inconvenient to you know a great number of people so she should be more ↑careful when planning all the event.

62.5: I never ↑thought that that. (2.2) Uh maybe (2.2) have you think about space ↑travel to other planet (1.4) Mars maybe? [laughs] Um mm of ↑course if I have the chance I will try yes and (1.8) but I also wan- don't wanna die [laughs] and don't want when I come back I'm still young and other people get old.

63.1: Uh maybe the ↑bravest animal is a (1.1) um: it's a ↑tiger or a ↑lion. I'm not so sure about it but ↑I think the most bravest animal should be the ↑wild ones because you are never trained in the ↑zoo and they maybe they would uh adjust to the wild ↑life very ↑naturally and: maybe they would prey for their: uh pray for the uh other animals for food.

63.2: I ↑think a good leader should be uh more or less ambitious. He would uh know uh what he wants to do and uh and what's more he w- he- he should know that what he ↑how can he do it so it's: uh usually a: a successful leader is a smart guy and uh: he can attra- uh he can attracts maybe people to focus on him.

63.12: Uh I think ↑Charlotte is more or less careless about the ↑arrangements. Maybe he is a very: uh is not a very ↑careful guy and uh um yeah it's fortun- ↑fortunately that uh she has found it uh before this party so there's a: chance an opportunity for her to correct it.

63.18: Uh what does the impression of Jimmy? I think Jimmy is um: (3.0) I don't know it depends on Jimmy's ↑age. If he's an adult ↑I think he's too uh: too quick to make a decision or make a connection between uh (1.0) two things or two conditions and uh there is something to correct about Jimmy's opinions or his beha- uh his habits of thinking about things and relate things.

64.2: I ↑think uh the ↑moral points make someone uh: (1.0) leader better I think. Uh: it has good uh: uh ↑technologies in the my ↑industry and can (1.0) teach everyone and- and can share the: ↑experience to someone else and uh (1.1) even uh and we ↑sometimes maybe we made ↑mistakes so for a good leader he can help to solve the ↑problem.

65.2: ↑I think what make a good leader is uh first of all he must have a cre- creative mind so uh I think uh the uh he can create some ↑interesting um some interesting ideas that made the team to become a strong team. Uh secondly I think maybe he must have a ↑power to make- to make the team to be a- a ↑real team. Uh they can com- compare it with each other.

65.4: Um I think it is cool to learning another language ↑ because I think you can broaden your ↑horizon and you can do more ↑knowledge about o- other country. Uh if you have some questions or some things to want to know you can come to another country to know ↑more knowledge and you can do ↑more things you want to do on another country that you cannot do on your own country.

65.12: Uh yes ↑I think she is little bit (1.4) mahong. (1.3) Huh how do you translate mahong to English? (1.4) Yeah um ↑I think she must be ↑carefully and uh in the future because uh this situation I don't- I ↑think she will not have this such a bad- bad- bad situation once more yeah.

65.13: Uh I think uh he will he- she might uh he can uh buy- bought um buy a ↑breakfast for that girl. ↑I think it is a good choice because she uh this way he can present he is a gentleman yeah. I think it's ↑good.

67.3: ↑I think: I think for ↑me it's English because every subject in my- in ↑now it's teached by English and I- I must to learn ↑many many English so I think ↑English is very difficult for me now. (2.8) Um (1.6) uh ↑calc- calculus you know I think calculus is difficult but teach calculus by English is ↑much difficult.

67.4: I think learning a second language (1.0) is ↑quite difficult- is quite difficult but I shall try my ↑best to learn- ↑learn it (1.1) like: ↑I think I like- I like German very much but I think it's ↑too difficult to me. (1.1) It's very very uh uh (1.7) I have- I have ↑tried to learn it before but I failed. Uh I can speak some- some ↑Japanese.

67.15: Ha I think he's so- he's so stupid and uh: he's someone- someone will wait for him and someone will some- someone will maybe ↑angry or- or maybe may have a bad impression on him. So ↑I think (MC) it's very very stupid.

68.6: Mm: I think um: a ↑story to be a- a- ↑good story maybe um should be um (4.0) ↑interesting and with ↑humor and um can um have maybe some a check on the peoples. Um (3.3) mm (2.6) so ma- maybe with a story uh about itself and uh about how to be success and um maybe people each ((like their)) own story.

69.8: Mm: I ↑like (1.1) ↑now the ↑most favorite TV show in uh in my opinion is (1.4) ↑reality- reality shows. I think people- people in reality- reality shows may: (1.5) ↑do acts in their ↑own opinion and

uh ↑show their real- ↑real personalities (1.3) for contacts and (1.5) communi- uh (2.8) com- common works.

71.5: Uh it will be very expensive at ↑first and only the ↑rich people can touch this field and I think um: years ↑later (1.0) this kind of activity may uh may become just a kind of entertainment for ↑normal people and everyone can touch it just like the: iPhone and at first they are very expensive and soon they just can a normal kind of ↑phone and everyone can use it.

71.18: He's real ↑cute and I really love uh I really love a ↑boy: with an- with an ↑imagination which adults cannot uh can ↑hardly have and some (1.3) ↑some of these part should be remain and that's okay. It will be a ↑huge profess- um um ↑property in his life.

72.2: So ↑if: you want to be a good leader I think first of all you need to know how to manage groups. You should know the others guys- the other guys in the groups their ↑capabilities and you can- you can manage the thems to how to say to: to perform their ↑best in this teams and this qui- I think it's a very ↑important properties for our leaders in the groups.

72.12: ↑I think it's okay. Some- ↑everyone will make mistakes so if you make ↑mistake you should correct it and then people will also like it- like you and for this cases I think if you found something is ↑not correct you should immediately correct it and then tell them people that it's ↑not correct and the- the right date is- is what and ↑people will also think you are ↑good ones.

73.8: I- I ↑never uh touched the subject like art. Uh (1.1) uh so I first uh I should uh investigate some mm the (1.2) uh (1.4) ↑knowledge about this subject and then I will mm ask some- somebody like my classmates or others to help me. Uh I ↑think it's a- a great- a- a great and a great job mm: and it- it's

beyo- ↑beyond my uh ability so I think it's a: it's ha- it's hard to me and ↑but- but I will- I will try and uh ↑challenge it and try my best to do it.

73.14: I- I will never be late in my interview with my friends so uh (1.0) if- if one of my friends he is late uh in our interview I will ((real an-)) I will be angry and um and: ask him or her to explain this thing or me so uh ↑but I ↑think uh it's not um: but we- we cannot um ↑think he or sh- uh him or her is a mm bad person or have something uh bad thoughts in uh in her or ↑him ((and uh)).

74.14: So Gary's problem but who knows uh this is you have a I ↑think it is a- a ↑accident you know will happen we cannot alter that everybody can do that they have an appointment at eight o' clock or ((fifteen)) eight ten or something ↑one hour or but I cant- I cannot ignore. It's I have to find something or they open can ↑open cannot open something like that so uh hard to see the next time never happen these things happen but I think but ↑Gary is a- is a real man with that real woman.

74.17: I ↑think he should tell- tell- tell Jimmy the- the whole story about the rain (3.0) why you get the rain is it clouds up we have the wet uh the air- the ↑air and we have cold air (2.0) mix it ↑together going to rain have then [laughs] also have the- the- the ↑wind why is the wind (1.5) something like that the whole story about that.

75.8: Uh ↑I think the uh TV shows is a little bit (1.3) ↑easy uh a little bit (2.3) mmm (1.7) uh the serious it too: (1.7) ↑weak and it can- cannot improve the (2.3) uh (9.9) can I say they can? Can I say it? I think the current TV shows a little ↑shallow um and uh (2.3) uh (1.5) it can uh- uh (1.0) it uh something like it'll include- include something that copy the uh ↑TV shows uh in the other countries (1.0) who don't have the (1.4) mm (1.8) pe- people.

76.18: 亮亮这个人吗就是比较片面的看问题。Uh 爸爸说↑一他就是一。他不知道一里面还有好几种的答法 uh 说法。他应该 uh 从这个里面重新去醒悟到一些的事情。他这个人可能就是很片面或者就是有点↑笨吗。

77.8: 当前的电视节目就是中医节目我觉得越来越喽越来越差。就没有以前那么好看。也没有以前有哪种节奏感后者一步一步↑代入感就比就说现已开始真样的一个新奇然后在一步一步地下去。现在就是很无聊的哪种。你都↑笑不出来的哪种终于节目我觉得↑并没有什么好看的。

77.13: 我↑觉得它可以定个闹钟。它可以定个闹钟。一个闹钟闹不醒的话可以定第二个。就是看他哪个 uh：熟睡程度怎怎样。她可以退－退前半个小时定个闹钟。如果他还没就可以闹钟会一直提醒你。

78.17: Uh 我觉得亮亮的爸爸如果想要亮亮↑彻底明白这个问题可能要给他给他讲的更详细一点儿。告诉他什么样的云彩出现的时候。Uh 天空中是什么↑原理所以才早上下雨了。这样呢亮亮真正明白这个原理时候他就不会简单的总结成有↑云就会下雨了。

79.7: ↑夏季最好的旅游地点我觉得↑是（1.3）uh：（1.2）具体地点不知道怎么说。也是比较哪种↑凉爽的吧。或者根据个人喜好－喜好－喜好－喜好自己想↑去的还是怎么样的。具体地点也要跟着个人的喜欢吧。

80.4: 如果能学习的话应该蛮不错的。但是现在：感↑觉学－学习第二种语言现在非常有必要出去玩或者怎么样。碰到－碰到别人感觉非常于必要要学。（1.4） Uh：如果要学的话应该要学英语（（不知别的））。

80.12: 应该犯错误应该人之长情吧。 每个人都会犯小错误。（1.4） 只要：大家相互理解应该还可以：原来也不是怎么样。不是怎么样大事情。Uh 所以应该没什么坏印象。

80.18: 感觉蛮↑天真的。比较可爱。Mm：见什么是什么。Uh：（1.3）但是也要去：学习才能（1.0）去说。（2.8）Uh：（1.9）感觉还是蛮可爱的。

81.8: 我对倒是－当前的电视节目我觉得影响还非常好。有体育的有文艺的((真实的有))军界的都有哈。但是就是手：mm 有些广告片哈 uh 我觉得是偏多那些。那广告效应太多。放个↑电视剧老是放广告。这样太耽误时间。而且看去来影响你的心情。

81.15: Um: 我对他也就是印象↑就是说呢可能都因为↑我想的好一些可能说他头一天太累了。睡过头了。特别是年轻人也能理解哈。如果说↑第二次再出现这个问题我↑觉得这个人不是 em 有特殊情况了。应该是这个人对自己（（？））一点了。不是很 em 搞-科研工作我觉得这个人就不太行。

82.10: Mm：（2.0）我↑觉得她应该再给他的所有朋友再发个请柬。然后附上 uh 一些：道歉的语气。然后还有要把时间改成星期六晚上。因为（1.7）因为就是在星期六晚上开。星期五是肯定准备不好的。

82.17: Uh：我↑觉得： 凉凉的爸爸可以在（1.0）↑下雨天的时候（1.0）带亮亮去看一次 uh－uh 去看一下外面的天就是有没有云彩。然后在－在↑晴天有云的时候再看一下外面天有没有下雨。然后 uh 用事实来证明给他的儿子。

83.1: 这个我还没－真没有好好想。（2.0）是最敢的动物应该是↑人吗。就是从主观的讲就是人闯到了文明。然后就是经历了这么多的东西。而且从 uh 我们的住房看就是出单这个事情觉得是。

83.6: 可能本身你经历的这个故事首先要比较有趣。然后你能够↑充分的故事。然后才－uh：一事自己前任经理的。二是你能够就迁徙的把这个故事满橙得熬打出来。我想他就是一个有趣的故事。

84.5: 可能有一个↑很有名的定律叫↑摩尔定律。摩尔定律↑说十八个月世界被改变一次。原来我们都↑不知道互联网世界是什么↑是九（1.1）↑九↑七年接触互联网↑中国↑最早的一批。那么当时↑我们提出就是↑卖东西。后来才知道是叫电子商务。嗯↑因为↑当时有一个唉很多东西不能完善。但是现在已经这个最平常的东西。所以↑从这个↑技术的哦发展方向来说太空旅行↑肯定也是↑未来的一个方向。

84.6: 啊首先我们要↑找到这个故事的↑核儿。比如说我们就是你是↑哪↑一方面的↑故事。你是属于那个是偏↑偏重于比如是 uh 这个可能要-要有[laughs] ↑分类[laughs]就故事的↑类别。比如说你是希望↑那个神话传说↑故事呢？还是说一个↑中国↑发展变化的故事呢？还是说↑一个- 一个-个比如说↑个人↑经历的故事？我觉得要首先要找到故事的↑核儿是什么。

84.15: Mm 我↑觉得↑印象（1.4）好像是↑一个比较比较↑没有↑心思的人。就是没有把什么事放在心上哈(1.0)第一。（DIR）第二呢就是(1.2)↑计划性不够哈。(1.3) 计划性↑不够哈[laughs]因

为你↑既然约好了嘛。还有这个是计划性（1.1）缺乏计划。第三呢我↑觉得可能是－是不是↑没有朋友啊什么的？如果跟朋友在一起可能↑不会－不会办这样子的事儿吧好像[laughs]。

85.13: ↑睡过头了的情况发生也正常。Uh 先道歉吧。然后以后要注意避免。自己可以设个闹铃之类的吧。然后：那当天就没办法了。已经发生了。（2.9）Uh：我↑觉得↑年轻人尤其这个时间观念还是很重要。还是要尽量避免。

86.3: 我↑认为学校最难的课程 uh：就是很多↑专业上的一些英语。因为很多词是比较生僻的。这对我们一些学英语来说人没有这样的一个环境。而且还要去每个↑词要去理解。这需要去花大量的时间。

86.10: 这主要看。如果说他自己星期五是有空的话可以－可以一由于这个。因为一个错误的话只要改改自己就好了。就是自己把活动放在星期五。如果觉得星期五那天她没空。我↑觉得也可以就是再寄第二份请柬。然后(再说明下问题。

86.11: 有什么办法？我↑觉得就是 mm 可以说寄第二份请柬。或者说是把真正的这个活动日期改掉。然后主要是要跟朋友说明情况。这个是最重要的。(2.1)然后嗯更需－更需要这个朋友能够理解她。

87.2: 哎呀这－这个问题对我来说很难。我从来没有想过这个问题。我↑觉得 uh（1.4）↑我可能－我是比较善于做一个 uh↑很好的执行者。我可能不太会－会领导别人会指挥别人。但是我能做一个 uh 能执－能按到领导的意图执行他安排的每一项↑工作。我↑觉得一个好的领导者来说他应该有一定得组织能力。

87.9: Uh 我对素食者没有设么看法。我↑觉得这是他的一种对生活的一种选择。Uh（1.0）我：可以↑认同他的看法。但是我自己我可能做不到。（2.3）mm↑其实素食来说 uh 我↑觉得挺好的。

87.10: Uh 飞飞应该－她应该马上通知 er 收到请柬的所有朋友。Uh（DIR）让他们改－uh 首先要表示↑道歉。（DIR）然后告诉他们正确的－的时间。欢迎大家能都来参加他的 party ［laughs］。↑我想他家可能都不会有什么问题的［laughs］。

88.1: 如果是分成种↑类这啊一种大类的话我感觉应该是↑狗吧。因为狗毕竟是人类的朋友。而且很多家里都会养猫狗这些动物陪伴自己。然后作为宠物来说实际它们自己本身也非常幸福。因为有人爱着它们。

89.5: 太空旅行我↑觉得这个是在未来的一个↑必然事件吧。因为人类现在探索太空主要就是（1.1）↑第一点知道生命的起源。第二点就是知道自己是从什么地方来的。因为（1.1）很多科学家都↑认为（1.0）地球上的生命都是从宇宙的生命分子里头传播过来的。

90.1: 最勇敢的动物是什么。是人的用情吧。用情－非得说是动物啊？这个－这个－就是人的一个心吧。Mm：怎么说呢？我↑觉得是那个心。这个↑真的没法说了。不知道说蛇么。（1.1）不行啊这点时间？（1.0）什么动物呢？老虎啊狮子啊我觉得↑都挺勇敢。

90.6: 有趣的故事啊我认为就是在你心前比较↑高兴的情况下。有时候特别↑搞笑的情况下可以写出一个有↑趣的故事。比－比如啊你前面↑发生的一些事啊。可以就是把它写在这个 ah 日记里啊。

90.13: 健健啊应该是 uh 晚上呢把表上一个↑闹钟。这样呢才不可以会有迟到。假如说他迟到的话也许是晚上↑感冒了呀拉↑肚子了 mm 这些问题不舒服导致他了迟到了。我↑认为是这样。

90.18: 对凉凉的印象啊↑觉得小孩儿还挺↑可爱的。爸爸说了天上有云他就觉得只要有云就下雨。觉得↑挺好玩儿的小孩儿我还比较喜欢他。挺好玩儿的挺活泼的。

91.5: ↑太空旅行 um↑就是↑现在有一些美国大片里面经常会有太空旅行。 我觉得↑蛮有意思的。但是现在对我们来说。可能还只是一个梦想吧。但是↑如果是有机会肯定还是会很多人想要去尝试。

91.7: Uh↑最好的啊↑是比如↑说有↑山或者有↑水的或者海↑边。这种可以带孩子一起去↑去玩的地方。然后也不要太热。Mm 比如↑说在中国的话就是嗯海南啊。或者是在北边一边的东北的 um 就觉得蛮↑凉快的地方。

91.15: Mm 啊首先他肯定是↑不守时。↑另外一个呢是比较啊比较↑懒惰的↑一个人。 啊从承诺别人八点。然后最后九点起床。嗯↑他应该是要变得更加↑勤快一点。更早起床。

92.18: 可能↑缺乏嗯↑少－↑缺少看书。那个对那个学习的知识↑掌握得不够充分。（1.0）建议他要多↑看书多学习。（3.3）↑在学习中提高自己的知识。

93.3: 啊最难的科程应该就是↑我不知道你们应该怎么说计量经济学。就里面有很多跟数学相关的建模啊什么的哎。这个当年我学的时候差点挂科了[laughs]嗯这样子。Uh 其实还有就是↑除此以外还有就是跟数学相关的可能都不太灵吧。因为 uh 数学的那个逻辑思维需要比较强。但我↑觉得我不是逻辑思维比较强的人。

93.8: 哈我已经很久不看电视了[laughs]。Uh 我↑觉得现在的电视节目可能就是↑娱乐性质的多一些吧。Uh 比较适合年龄就是年青年↑年青人看。另外就是多是一些就是表现那个 uh 就↑拍了很多比较重复的一些电视↑剧。就没有什么创新性。所以现在我觉得电视节目一般吧对我们。

94.3: 学校↑最难的我觉得↑对于我来说是数学。em(1.0) em 尤其是↑上大学时的高数。我觉得↑太头疼了。所以一般像我们女孩儿多比较都喜欢偏文类的。就是文科类的一点啊。数学能学的稍微简单一点。em（2.7）ah [laughs]ah[laughs]反正啊就是哦对！因为↑数学它比较讲究那个逻辑性嘛。逻辑性思维比较强一点。

95.1: 世界上最勇敢的动物啊 mm 这倒有点难说了。[laughs]只有我-我↑自己认为的哈。世界上最勇敢的动物我觉得还是:我觉得↑狼比较勇敢。[laughs]好可以了。

96.4: 我喜欢。我最想学的是韩语[laughs]。因为－因为以前↑看过学了一下。但是哎哟后来学着学着学了一下后来就唉太那个一个人后来就因为时间↑长事情也太多了再后来就把这个事情过去了。但如果有这个机会如果就是自己↑有什么学习韩语的我觉得我仍会试试的。

98.2: Uh：（1.2）据我现在觉得阿↑好的领导者应该（1.1）（uh）（1.2）身体力行。然后有什么事情都要和员工一起同甘共苦。（1.0）然后↑多关心一下员工。然后再（1.2）对这个员工的福利或者什么的方面我觉得他应该会想得会更多一点。

98.17: Mm 我↑觉得他爸爸应该就是（1.5）先要给他↑讲一讲。然后就是为什么会下雨。然后（1.0）就要说明它这个有多少个几率会下雨。（0.8）然后或者就是然后再下一次再遇上这一种情 uh 各种情况没有下雨或有下雨天上有云彩。Uh 然后给他一一地↑做分析。

99.12: ↑印像嘛这个（1.1）孩子比较粗心大意。怎么－怎么把时间居然写错了。（1.0）这样的话办事太有点有点不靠谱。（1.8）希望他↑下一次就是说做事情啊要把它计划好。都整理了好了然后再通知大家。

100.2: ↑首先就是说↑应该我觉得就是说。↑发生一件事情的时候就不要。比如说他犯了错误不要说就批评他。就是你要先想想这个事情是怎么发生的。然后要设身处地地站在他的角度如果说你做这件事情你会怎样怎样做。然后你再才去批评他。或者是怎么↑样。平常的时候我觉得我们领导就特别好。就感觉让人感觉很温暖。

100.7: 是哪里呢？唉呀当然了想不起来了。]我之前去过一次云南。我觉得云南还挺好的。虽然嗯我是↑在↑夏季之前去的去了西双版纳还有香格里↑那。觉得那里的↑空气特别好。然后去了-我比↑较喜欢谢娜和张继。然后他俩结婚的那个香格里那那个地方也去了。感觉特别棒。

```
22 TranscriptionComparator.sln
        @@ -0,0 +1,22 @@
+
+Microsoft Visual Studio Solution File, Format Version 12.00
+# Visual Studio 14
+VisualStudioVersion = 14.0.24720.0
+MinimumVisualStudioVersion = 10.0.40219.1
+Project("{FAE04EC0-301F-11D3-BF4B-00C04F79EFBC}") = "TranscriptionComparator",
"src\TranscriptionComparator.csproj", "{834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}"
+EndProject
+Global
+    GlobalSection(SolutionConfigurationPlatforms) = preSolution
+        Debug|Any CPU = Debug|Any CPU
+        Release|Any CPU = Release|Any CPU
+    EndGlobalSection
+    GlobalSection(ProjectConfigurationPlatforms) = postSolution
+        {834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}.Debug|Any CPU.ActiveCfg = Debug|Any CPU
+        {834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}.Debug|Any CPU.Build.0 = Debug|Any CPU
+        {834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}.Release|Any CPU.ActiveCfg = Release|Any CPU
+        {834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}.Release|Any CPU.Build.0 = Release|Any CPU
+    EndGlobalSection
+    GlobalSection(SolutionProperties) = preSolution
+        HideSolutionNode = FALSE
+    EndGlobalSection
+EndGlobal
6 src/App.config
        @@ -0,0 +1,6 @@
+<?xml version="1.0" encoding="utf-8" ?>
+<configuration>
+  <startup>
+    <supportedRuntime version="v4.0" sku=".NETFramework,Version=v4.6" />
```

```
+    </startup>
+</configuration>
30 src/ComparisonResult.cs
            @@ -0,0 +1,30 @@
+using System;
+using System.Collections.Generic;
+
+namespace TranscriptionComparator
+{
+    public class ComparisonResult
+    {
+        /// <summary>
+        /// Group the interviewee belongs to.
+        /// </summary>
+        public int GroupNumber { get; set; }
+
+        /// <summary>
+        /// Name of interviewee.
+        /// </summary>
+        public string SubjectName { get; set; }
+
+        /// <summary>
+        /// Name of the question / prompt given to the interviewee.
+        /// </summary>
+        public string PromptName { get; set; }
+
+        public string TextA { get; set; }
+        public string TextB { get; set; }
+        public int LevenshteinDistance { get; set; }
+        public List<Diff> Differences { get; set; }
+        public int TotalUnitsCount => Math.Max(TextA?.Length ?? 0, TextB?.Length ?? 0);
+        public string HtmlDelta { get; set; }
+    }
```

```
+}

2,302 src/DiffMatchPatch.cs
2,302 additions, 0 deletions not shown because the diff is too large. Please use a local Git client to view these changes.

41 src/HtmlRenderer.cs
            @@ -0,0 +1,41 @@

+using System.IO;

+

+namespace TranscriptionComparator

+{

+    public class HtmlRenderer : IComparisonResultRenderer

+    {

+        private readonly TextWriter _stream;

+

+        public HtmlRenderer(TextWriter stream)

+        {

+            _stream = stream;

+            _stream.WriteLine("<html><head><meta http-equiv=\"Content-Type\" content=\"text/html; charset=UTF-8\"
/></head><body><table><tr><th>Text A</th><th>Text B</th><th>Delta</th><th>Levenshtein</th></tr>");

+        }

+

+        public void Render(ComparisonResult result)

+        {

+            _stream.WriteLine("<tr>");

+            _stream.WriteLine("<td>");

+            _stream.WriteLine(result.TextA);

+            _stream.WriteLine("</td>");

+            _stream.WriteLine("<td>");

+            _stream.WriteLine(result.TextB);

+            _stream.WriteLine("</td>");

+            _stream.WriteLine("<td>");

+            _stream.WriteLine(result.HtmlDelta);

+            _stream.WriteLine("</td>");

+            _stream.WriteLine("<td>");
```

```
+            _stream.WriteLine(result.LevenshteinDistance);
+            _stream.WriteLine("</td>");
+            _stream.WriteLine("</tr>");
+            _stream.Flush();
+        }
+
+        public void Dispose()
+        {
+            _stream.WriteLine("</body></html>");
+            _stream.Flush();
+            _stream.Dispose();
+        }
+    }
+}
9 src/IComparisonResultRenderer.cs
            @@ -0,0 +1,9 @@
+using System;
+
+namespace TranscriptionComparator
+{
+    public interface IComparisonResultRenderer : IDisposable
+    {
+        void Render(ComparisonResult result);
+    }
+}
76 src/Iterators/TranscriptionComparator.cs
            @@ -0,0 +1,76 @@
+using System;
+using System.Collections;
+using System.Collections.Generic;
+using System.Linq;
+
+namespace TranscriptionComparator.Iterators
```

```
+{
+   public class TranscriptionComparator : IEnumerable<ComparisonResult>
+   {
+      private static readonly diff_match_patch DiffMatchPatch = new diff_match_patch();
+      private readonly TranscriptionDocument _transcriptionSectionsA;
+      private readonly TranscriptionDocument _transcriptionSectionsB;
+
+      public TranscriptionComparator(string transcriptionFilePathA, string transcriptionFilePathB)
+      {
+         _transcriptionSectionsA = new TranscriptionDocument(transcriptionFilePathA);
+         _transcriptionSectionsB = new TranscriptionDocument(transcriptionFilePathB);
+      }
+
+      public IEnumerator<ComparisonResult> GetEnumerator()
+      {
+         var transcriptionLines = _transcriptionSectionsA.Zip(_transcriptionSectionsB,
+            (a, b) => new Tuple<string, string>(a, b));
+
+         foreach (var line in transcriptionLines)
+         {
+            var differences = DiffMatchPatch.diff_main(line.Item1, line.Item2);
+
+            DiffMatchPatch.diff_cleanupEfficiency(differences);
+
+            DiffMatchPatch.diff_cleanupSemanticLossless(differences);
+
+            differences = CleanupCasingIssues(differences).ToList();
+
+            yield return new ComparisonResult
+            {
+               TextA = line.Item1,
+               TextB = line.Item2,
```

```
+                Differences = differences,
+                HtmlDelta = DiffMatchPatch.diff_prettyHtml(differences),
+                LevenshteinDistance = DiffMatchPatch.diff_levenshtein(differences),
+            };
+        }
+    }
+
+    private static IEnumerable<Diff> CleanupCasingIssues(List<Diff> differences)
+    {
+        for (var i = 0; differences.Count > i; i++)
+        {
+            var current = differences[i];
+
+            if (current.operation == Operation.DELETE && (i + 1) < differences.Count)
+            {
+                var next = differences[i + 1];
+                if (next.operation == Operation.INSERT)
+                {
+                    if (String.Compare(next.text.Trim(), current.text.Trim(), StringComparison.OrdinalIgnoreCase) == 0)
+                    {
+                        yield return new Diff(Operation.EQUAL, current.text);
+
+                        i++; // step over next
+                        continue;
+                    }
+                }
+            }
+
+            yield return current;
+        }
+    }
+
```

```
+       IEnumerator IEnumerable.GetEnumerator()
+       {
+           return GetEnumerator();
+       }
+   }
+}
36 src/Iterators/TranscriptionDocument.cs
            @@ -0,0 +1,36 @@
+using System;
+using System.Collections;
+using System.Collections.Generic;
+using System.Linq;
+using System.Text.RegularExpressions;
+using TranscriptionComparator.Strategies;
+
+namespace TranscriptionComparator.Iterators
+{
+   public class TranscriptionDocument : IEnumerable<string>
+   {
+       private readonly string _fileContents;
+       private readonly NormaliseText _normaliser;
+
+       public TranscriptionDocument(string fileContents)
+       {
+           _fileContents = fileContents;
+           _normaliser = new NormaliseText();
+       }
+
+       public IEnumerator<string> GetEnumerator()
+       {
+           return Regex.Replace(_fileContents, "\r\n\\s*\r\n", "\r\n\r\n")
+               .Replace("\r\n\r\n", "`")
+               .Split(`')
```

```
+            .Select(t => _normaliser.Perform(t))
+            .Where(t => !String.IsNullOrWhiteSpace(t))
+            .GetEnumerator();
+    }
+
+    IEnumerator IEnumerable.GetEnumerator()
+    {
+        return GetEnumerator();
+    }
+  }
+}
48 src/Program.cs
          @@ -0,0 +1,48 @@
+using System;
+using System.Diagnostics;
+using System.IO;
+
+namespace TranscriptionComparator
+{
+  class Program
+  {
+    static void Main()
+    {
+      Console.WriteLine("Please select comparison and press ENTER:");
+      Console.WriteLine("A: English");
+      Console.WriteLine("B: Chinese");
+
+      var selection = Console.ReadLine();
+
+      var args = new[]
+      {
+        Transcriptions.en_Interrating_Lucy,
```

```
+            Transcriptions.en_Interrating_William
+        };
+
+        if (String.Compare(selection, "B", StringComparison.OrdinalIgnoreCase) == 0)
+        {
+            args = new[]
+            {
+                Transcriptions.ch_Interrating_Lucy,
+                Transcriptions.ch_Interrating_Kang
+            };
+        }
+
+        var savePath = $"Comparison_{DateTime.Now.Ticks}.html";
+
+        var stopwatch = Stopwatch.StartNew();
+
+        using (var renderer = new HtmlRenderer(File.CreateText(savePath)))
+        {
+            foreach (var result in new Iterators.TranscriptionComparator(args[0], args[1]))
+            {
+                renderer.Render(result);
+            }
+        }
+
+        Console.WriteLine($"Completed in {stopwatch.Elapsed}");
+        Console.ReadLine();
+    }
+ }
+}
36 src/Properties/AssemblyInfo.cs
          @@ -0,0 +1,36 @@
+using System.Reflection;
+using System.Runtime.CompilerServices;
```

```
+using System.Runtime.InteropServices;
+
+// General Information about an assembly is controlled through the following
+// set of attributes. Change these attribute values to modify the information
+// associated with an assembly.
+[assembly: AssemblyTitle("TranscriptionValidator")]
+[assembly: AssemblyDescription("")]
+[assembly: AssemblyConfiguration("")]
+[assembly: AssemblyCompany("")]
+[assembly: AssemblyProduct("TranscriptionValidator")]
+[assembly: AssemblyCopyright("Copyright ©  2016")]
+[assembly: AssemblyTrademark("")]
+[assembly: AssemblyCulture("")]
+
+// Setting ComVisible to false makes the types in this assembly not visible
+// to COM components.  If you need to access a type in this assembly from
+// COM, set the ComVisible attribute to true on that type.
+[assembly: ComVisible(false)]
+
+// The following GUID is for the ID of the typelib if this project is exposed to COM
+[assembly: Guid("834a3ca3-e2d1-4d96-9c13-2f743dd9db31")]
+
+// Version information for an assembly consists of the following four values:
+//
+//      Major Version
+//      Minor Version
+//      Build Number
+//      Revision
+//
+// You can specify all the values or you can default the Build and Revision Numbers
+// by using the '*' as shown below:
+// [assembly: AssemblyVersion("1.0.*")]
```

```
+[assembly: AssemblyVersion("1.0.0.0")]
+[assembly: AssemblyFileVersion("1.0.0.0")]
132 src/Resources/ch-Interrating-Kang.txt
            @@ -0,0 +1,132 @@
+[Mandarin transcription here]
420 src/Resources/en-Interrating-Lucy.txt
            @@ -0,0 +1,420 @@
+[English transcription here]
22 src/Strategies/NormaliseText.cs
            @@ -0,0 +1,22 @@
+using System.Linq;
+
+namespace TranscriptionComparator.Strategies
+{
+    public class NormaliseText : ITextManipulationStrategy
+    {
+        private static readonly ITextManipulationStrategy[] Strategies =
+        {
+            new RemoveExtraWhitespace(),
+            new RemoveFillerWords(),
+            new RemovePunctuation(),
+            new RemoveUnintelligbleSpeech(),
+            new RemovePadding(),
+        };
+
+        public string Perform(string input)
+        {
+            return Strategies.Aggregate(input,
+                (current, strategy) => strategy.Perform(current));
+        }
+    }
+}
12 src/Strategies/RemoveExtraWhitespace.cs
            @@ -0,0 +1,12 @@
```

```
+using System.Text.RegularExpressions;
+
+namespace TranscriptionComparator.Strategies
+{
+    public class RemoveExtraWhitespace : ITextManipulationStrategy
+    {
+        public string Perform(string input)
+        {
+            return Regex.Replace(input, " +", " ");
+        }
+    }
+}
```

15 src/Strategies/RemoveFillerWords.cs

```
@@ -0,0 +1,15 @@
+using System;
+using System.Text.RegularExpressions;
+
+namespace TranscriptionComparator.Strategies
+{
+    public class RemoveFillerWords : ITextManipulationStrategy
+    {
+        private const string pattern = "\\b(uh|um|mm|em|er|ha)\\b";
+
+        public string Perform(string input)
+        {
+            return Regex.Replace(input, pattern, String.Empty, RegexOptions.IgnoreCase);
+        }
+    }
+}
```

10 src/Strategies/RemovePadding.cs

```
@@ -0,0 +1,10 @@
+namespace TranscriptionComparator.Strategies
+{
```

```
+    public class RemovePadding : ITextManipulationStrategy
+    {
+        public string Perform(string input)
+        {
+            return input.Trim();
+        }
+    }
+}
15 src/Strategies/RemovePunctuation.cs
              @@ -0,0 +1,15 @@
+using System;
+using System.Text.RegularExpressions;
+
+namespace TranscriptionComparator.Strategies
+{
+    public class RemovePunctuation : ITextManipulationStrategy
+    {
+        private const string pattern = @"[^\w\s\.\?]";
+
+        public string Perform(string input)
+        {
+            return Regex.Replace(input, pattern, String.Empty, RegexOptions.IgnoreCase);
+        }
+    }
+}
15 src/Strategies/RemoveUnintelligbleSpeech.cs
              @@ -0,0 +1,15 @@
+using System;
+using System.Text.RegularExpressions;
+
+namespace TranscriptionComparator.Strategies
+{
+    public class RemoveUnintelligbleSpeech : ITextManipulationStrategy
```

```
+    {
+        private const string pattern = "(\\[.*?\\])|(\".*?\")|('.*?')|(\\(.*?\\))";
+
+        public string Perform(string input)
+        {
+            return Regex.Replace(input, pattern, String.Empty, RegexOptions.IgnoreCase);
+        }
+    }
+}
7 src/Strategies/TextManipulationStrategy.cs
            @@ -0,0 +1,7 @@
+namespace TranscriptionComparator.Strategies
+{
+    public interface ITextManipulationStrategy
+    {
+        string Perform(string input);
+    }
+}
91 src/TranscriptionComparator.csproj
            @@ -0,0 +1,91 @@
+<?xml version="1.0" encoding="utf-8"?>
+<Project ToolsVersion="14.0" DefaultTargets="Build"
xmlns="http://schemas.microsoft.com/developer/msbuild/2003">
+  <Import Project="$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props"
Condition="Exists('$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props')" />
+  <PropertyGroup>
+    <Configuration Condition=" '$(Configuration)' == '' ">Debug</Configuration>
+    <Platform Condition=" '$(Platform)' == '' ">AnyCPU</Platform>
+    <ProjectGuid>{834A3CA3-E2D1-4D96-9C13-2F743DD9DB31}</ProjectGuid>
+    <OutputType>Exe</OutputType>
+    <AppDesignerFolder>Properties</AppDesignerFolder>
+    <RootNamespace>TranscriptionComparator</RootNamespace>
+    <AssemblyName>TranscriptionComparator</AssemblyName>
+    <TargetFrameworkVersion>v4.6</TargetFrameworkVersion>
```

```
+    <FileAlignment>512</FileAlignment>
+    <AutoGenerateBindingRedirects>true</AutoGenerateBindingRedirects>
+  </PropertyGroup>
+  <PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
+    <PlatformTarget>AnyCPU</PlatformTarget>
+    <DebugSymbols>true</DebugSymbols>
+    <DebugType>full</DebugType>
+    <Optimize>false</Optimize>
+    <OutputPath>bin\Debug\</OutputPath>
+    <DefineConstants>DEBUG;TRACE</DefineConstants>
+    <ErrorReport>prompt</ErrorReport>
+    <WarningLevel>4</WarningLevel>
+  </PropertyGroup>
+  <PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
+    <PlatformTarget>AnyCPU</PlatformTarget>
+    <DebugType>pdbonly</DebugType>
+    <Optimize>true</Optimize>
+    <OutputPath>bin\Release\</OutputPath>
+    <DefineConstants>TRACE</DefineConstants>
+    <ErrorReport>prompt</ErrorReport>
+    <WarningLevel>4</WarningLevel>
+  </PropertyGroup>
+  <ItemGroup>
+    <Reference Include="System" />
+    <Reference Include="System.Web" />
+  </ItemGroup>
+  <ItemGroup>
+    <Compile Include="ComparisonResult.cs" />
+    <Compile Include="HtmlRenderer.cs" />
+    <Compile Include="Strategies\RemovePadding.cs" />
+    <Compile Include="DiffMatchPatch.cs" />
+    <Compile Include="IComparisonResultRenderer.cs" />
```

```
+    <Compile Include="Strategies\NormaliseText.cs" />
+    <Compile Include="Program.cs" />
+    <Compile Include="Properties\AssemblyInfo.cs" />
+    <Compile Include="Strategies\RemoveExtraWhitespace.cs" />
+    <Compile Include="Strategies\RemoveFillerWords.cs" />
+    <Compile Include="Strategies\RemovePunctuation.cs" />
+    <Compile Include="Strategies\RemoveUnintelligbleSpeech.cs" />
+    <Compile Include="Strategies\TextManipulationStrategy.cs" />
+    <Compile Include="Iterators\TranscriptionComparator.cs" />
+    <Compile Include="Iterators\TranscriptionDocument.cs" />
+    <Compile Include="Transcriptions.Designer.cs">
+      <AutoGen>True</AutoGen>
+      <DesignTime>True</DesignTime>
+      <DependentUpon>Transcriptions.resx</DependentUpon>
+    </Compile>
+  </ItemGroup>
+  <ItemGroup>
+    <None Include="App.config" />
+  </ItemGroup>
+  <ItemGroup>
+    <EmbeddedResource Include="Transcriptions.resx">
+      <Generator>PublicResXFileCodeGenerator</Generator>
+      <LastGenOutput>Transcriptions.Designer.cs</LastGenOutput>
+      <SubType>Designer</SubType>
+    </EmbeddedResource>
+  </ItemGroup>
+  <ItemGroup>
+    <None Include="Resources\ch-Interrating-Kang.txt" />
+  </ItemGroup>
+  <ItemGroup>
+    <None Include="Resources\ch-Interrating-Lucy.txt" />
+  </ItemGroup>
```

```
+ <ItemGroup>
+   <None Include="Resources\en-Interrating-Lucy.txt" />
+ </ItemGroup>
+ <ItemGroup>
+   <None Include="Resources\en-Interrating-William.txt" />
+ </ItemGroup>
+ <Import Project="$(MSBuildToolsPath)\Microsoft.CSharp.targets" />
+ <!-- To modify your build process, add your task inside one of the targets below and uncomment it.
+      Other similar extension points exist, see Microsoft.Common.targets.
+ <Target Name="BeforeBuild">
+ </Target>
+ <Target Name="AfterBuild">
+ </Target>
+ -->
+</Project>
```

127 src/Transcriptions.Designer.cs
Diff suppressed. Click to show.

133 src/Transcriptions.resx
@@ -0,0 +1,133 @@

```
+<?xml version="1.0" encoding="utf-8"?>
+<root>
+ <!--
+   Microsoft ResX Schema
+
+   Version 2.0
+
+   The primary goals of this format is to allow a simple XML format
+   that is mostly human readable. The generation and parsing of the
+   various data types are done through the TypeConverter classes
+   associated with the data types.
+
+   Example:
+
```

+ ... ado.net/XML headers & schema ...

+ <resheader name="resmimetype">text/microsoft-resx</resheader>

+ <resheader name="version">2.0</resheader>

+ <resheader name="reader">System.Resources.ResXResourceReader, System.Windows.Forms, ...</resheader>

+ <resheader name="writer">System.Resources.ResXResourceWriter, System.Windows.Forms, ...</resheader>

+ <data name="Name1"><value>this is my long string</value><comment>this is a comment</comment></data>

+ <data name="Color1" type="System.Drawing.Color, System.Drawing">Blue</data>

+ <data name="Bitmap1" mimetype="application/x-microsoft.net.object.binary.base64">

+ <value>[base64 mime encoded serialized .NET Framework object]</value>

+ </data>

+ <data name="Icon1" type="System.Drawing.Icon, System.Drawing" mimetype="application/x-microsoft.net.object.bytearray.base64">

+ <value>[base64 mime encoded string representing a byte array form of the .NET Framework object]</value>

+ <comment>This is a comment</comment>

+ </data>

+

+ There are any number of "resheader" rows that contain simple

+ name/value pairs.

+

+ Each data row contains a name, and value. The row also contains a

+ type or mimetype. Type corresponds to a .NET class that support

+ text/value conversion through the TypeConverter architecture.

+ Classes that don't support this are serialized and stored with the

+ mimetype set.

+

+ The mimetype is used for serialized objects, and tells the

+ ResXResourceReader how to depersist the object. This is currently not

+ extensible. For a given mimetype the value must be set accordingly:

+

+ Note - application/x-microsoft.net.object.binary.base64 is the format

+ that the ResXResourceWriter will generate, however the reader can

+ read any of the formats listed below.

```
+
+   mimetype: application/x-microsoft.net.object.binary.base64
+   value   : The object must be serialized with
+           : System.Runtime.Serialization.Formatters.Binary.BinaryFormatter
+           : and then encoded with base64 encoding.
+
+   mimetype: application/x-microsoft.net.object.soap.base64
+   value   : The object must be serialized with
+           : System.Runtime.Serialization.Formatters.Soap.SoapFormatter
+           : and then encoded with base64 encoding.
+
+   mimetype: application/x-microsoft.net.object.bytearray.base64
+   value   : The object must be serialized into a byte array
+           : using a System.ComponentModel.TypeConverter
+           : and then encoded with base64 encoding.
+   -->
+ <xsd:schema id="root" xmlns="" xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:msdata="urn:schemas-
microsoft-com:xml-msdata">
+   <xsd:import namespace="http://www.w3.org/XML/1998/namespace" />
+   <xsd:element name="root" msdata:IsDataSet="true">
+     <xsd:complexType>
+       <xsd:choice maxOccurs="unbounded">
+         <xsd:element name="metadata">
+           <xsd:complexType>
+             <xsd:sequence>
+               <xsd:element name="value" type="xsd:string" minOccurs="0" />
+             </xsd:sequence>
+             <xsd:attribute name="name" use="required" type="xsd:string" />
+             <xsd:attribute name="type" type="xsd:string" />
+             <xsd:attribute name="mimetype" type="xsd:string" />
+             <xsd:attribute ref="xml:space" />
+           </xsd:complexType>
+         </xsd:element>
```

```
+        <xsd:element name="assembly">
+          <xsd:complexType>
+            <xsd:attribute name="alias" type="xsd:string" />
+            <xsd:attribute name="name" type="xsd:string" />
+          </xsd:complexType>
+        </xsd:element>
+        <xsd:element name="data">
+          <xsd:complexType>
+            <xsd:sequence>
+              <xsd:element name="value" type="xsd:string" minOccurs="0" msdata:Ordinal="1" />
+              <xsd:element name="comment" type="xsd:string" minOccurs="0" msdata:Ordinal="2" />
+            </xsd:sequence>
+            <xsd:attribute name="name" type="xsd:string" use="required" msdata:Ordinal="1" />
+            <xsd:attribute name="type" type="xsd:string" msdata:Ordinal="3" />
+            <xsd:attribute name="mimetype" type="xsd:string" msdata:Ordinal="4" />
+            <xsd:attribute ref="xml:space" />
+          </xsd:complexType>
+        </xsd:element>
+        <xsd:element name="resheader">
+          <xsd:complexType>
+            <xsd:sequence>
+              <xsd:element name="value" type="xsd:string" minOccurs="0" msdata:Ordinal="1" />
+            </xsd:sequence>
+            <xsd:attribute name="name" type="xsd:string" use="required" />
+          </xsd:complexType>
+        </xsd:element>
+      </xsd:choice>
+    </xsd:complexType>
+  </xsd:element>
+ </xsd:schema>
+ <resheader name="resmimetype">
+   <value>text/microsoft-resx</value>
```

+ </resheader>

+ <resheader name="version">

+ <value>2.0</value>

+ </resheader>

+ <resheader name="reader">

+ <value>System.Resources.ResXResourceReader, System.Windows.Forms, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089</value>

+ </resheader>

+ <resheader name="writer">

+ <value>System.Resources.ResXResourceWriter, System.Windows.Forms, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089</value>

+ </resheader>

+ <assembly alias="System.Windows.Forms" name="System.Windows.Forms, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089" />

+ <data name="ch_Interrating_Kang" type="System.Resources.ResXFileRef, System.Windows.Forms">

+ <value>Resources\ch-Interrating-Kang.txt;System.String, mscorlib, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089;utf-8</value>

+ </data>

+ <data name="ch_Interrating_Lucy" type="System.Resources.ResXFileRef, System.Windows.Forms">

+ <value>Resources\ch-Interrating-Lucy.txt;System.String, mscorlib, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089;utf-8</value>

+ </data>

+ <data name="en_Interrating_Lucy" type="System.Resources.ResXFileRef, System.Windows.Forms">

+ <value>Resources\en-Interrating-Lucy.txt;System.String, mscorlib, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089;utf-8</value>

+ </data>

+ <data name="en_Interrating_William" type="System.Resources.ResXFileRef, System.Windows.Forms">

+ <value>Resources\en-Interrating-William.txt;System.String, mscorlib, Version=4.0.0.0, Culture=neutral, PublicKeyToken=b77a5c561934e089;utf-8</value>

+ </data>

+</root>
Lock conversation

Appendix 9: Information sheet (Phase 2)

[Group 1:]

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk



Participant Information Sheet

Title of Project: Cross-linguistic expressions of opinion
Researcher: Lucy Zhao, PRS DPhil Education

Hello, my name is Lucy Zhao, and I am a DPhil student at Oxford University under the supervision of Dr. Victoria Murphy. I would like to invite you to participate in this research project on the influence of culture and language on communication. You are being asked to participate due to your unique language skill set. Participation is completely voluntary; choosing not to take part will not be detrimental to you in any way. The survey involves voicing your thoughts on intriguing and sometimes quirky opinion questions.

Purpose of study: This research will explore how language knowledge and language environment affect self-expression and communication, particularly in expression of opinions. Participants in this study involve students from America and China.

Benefits of study: While participants do not directly benefit, this study does indirectly benefit participants through potential findings and implications for language learning and communication. This research aims shed light on how language acquisition and language-learning environment affects communication style and cross-cultural interaction, which has implications for improving language-teaching methods. Ultimately, I aim to improve cross-cultural communication between native and foreign language speakers.

Risks of study: There are no foreseeable or expected risks involved in participating in this project.

Description of procedure: If you agree to participate, I will ask you to fill out a brief (<1 minute) online questionnaire about your language background. You may then be asked to participate in the next phase, which consists of a short (~5 minute) oral survey containing general and diverse opinion questions and follow up language questionnaire. The researcher will conduct the survey either over Skype or in-person at your chosen date and time of convenience. This survey will be audio recorded for purpose of analysis; your data will remain strictly confidential and secured in a password-protected-server at all times. Depending on your responses, you may be asked to take two short (5

minute) online language quizzes to measure vocabulary depth and breadth. Additionally, I will provide you with any significant findings resulting from this research if desired.

Contact information: If you have any questions or concerns, or desire more information, please email the primary researcher, Lucy Zhao, at: lucy.zhao@education.ox.ac.uk; or the supervisor, Dr. Victoria Murphy, at: victoria.murphy@education.ox.ac.uk. I thank you for your time and look forward to your response!

[Group 2, 3, 4:] (actual version, followed by translation)

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk

参加调查信息说明书

题目：跨语言的观念表述(Cross-linguistic expressions of opinion)
研究者: Lucy Zhao, 教育系博士生 （PRS DPhil Education

我叫 Lucy Zhao, 我是牛津大学Victoria Murphy 导师的博士研究生。我选择请你们参加这项语言文化的研究课题，是因为你们具有特殊的语言能力。参加或是退出调查，是完全自愿的。调查题目也包括你对有些问题的看法。

研究目的: 这项研究是为了发现语言知识和语言环境是如何影响我们的语言表达,特别是如何影响我们对事物的看法。被调查者可以是美国学生或是中国学生。

研究获得的利益: 尽管参加调查的学生将不会直接受益，但是，他们可以间接获益于语言的学习和交流。这项研究着眼于找出语言环境是如何影响人们的表达方式和文化交流，从而改进语言教育方式。最终的目的是要改善在以讲英语为母语和讲外语的人们之间的文化交流。

研究的风险: 对参加调查者没有任何风险。

调查研究的程序: 如果你愿意参加调查，请你上网回答有关你的语言背景问题（约1分钟）。然后，你将参加约5分钟的口头调查, 题目包括对不同层次的语言问答。根据你的回答，你可能还需要再做两个5分钟的网上语言和词汇测试。这项调查将会在你选择的时间地点通过Skype或面对面的方式进行。 出于研究分析的目的，调查将被录音; 在整个研究期间，你的个人资料会被严格保存在计算机系统中。我会届时提供研究的进展情况。

联系方式: 如果你有任何问题, 请联系项目研究人员Lucy Zhao，电子邮箱: lucy.zhao@education.ox.ac.uk; 或者联系Victoria Murphy导师，电子邮箱: victoria.murphy@education.ox.ac.uk.
感谢你的时间和支持!

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk



Participant Information Sheet

Title of Project: Cross-linguistic expressions of opinion
Researcher: Lucy Zhao, PRS DPhil Education

Hello, my name is Lucy Zhao, and I am a DPhil student at Oxford University under the supervision of Dr. Victoria Murphy. I would like to invite you to participate in this research project on the influence of culture and language on communication. You are being asked to participate due to your unique language skill set. Participation is completely voluntary; choosing not to take part will not be detrimental to you in any way. The survey involves voicing your thoughts on intriguing and sometimes quirky opinion questions.

Purpose of study: This research will explore how language knowledge and language environment affect the cognition behind spoken discourse, particularly in expression of opinions. Participants in this study involve students from America and China.

Benefits of study: While participants do not directly benefit, this study does indirectly benefit participants through potential findings and implications for language learning and communication. This research aims shed light on how language acquisition and language-learning environment affects communication style and cross-cultural interaction, which has implications for improving language-teaching methods. Ultimately, I aim to improve cross-cultural communication between native and foreign language speakers.

Risks of study: There are no foreseeable or expected risks involved in participating in this project.

Description of procedure: If you agree to participate, I will ask you to fill out a brief (<1 minute) online questionnaire about your language background. You may then be asked to participate in the next phase, which consists of a short (~5 minute) oral survey containing general and diverse opinion questions and follow up language questionnaire. The researcher will conduct the survey either over Skype or in-person at your chosen date and time of convenience. This survey will be audio recorded for purpose of analysis; your data will remain strictly confidential and secured in a password-protected-server at all times. Depending on your responses, you may be asked to take two short (5 minute) online language quizzes to measure vocabulary depth and breadth. Additionally, I will provide you with any significant findings resulting from this research if desired.

Contact information: If you have any questions or concerns, or desire more information, please email the primary researcher, Lucy Zhao, at: lucy.zhao@education.ox.ac.uk; or the supervisor, Dr. Victoria Murphy, at: victoria.murphy@education.ox.ac.uk. I thank you for your time and look forward to your response!

Appendix 10: Production task (Phase 2)

[Group 1, 2, 3:]

Instructions: Participants are to answer the question given in the prompt until instructed to stop. Any questions which the participant might ask the researcher about the prompt must be answered with "it's up to you" or "however you want to interpret it" so as not to lead or guide the participant's response.

Set 1: Direct opinion questions

1) What is your opinion on being a vegetarian?

2) Which is the bravest animal and why?

3) What is a good place to travel in the summer and why?

4) What is the hardest school subject and why?

5) What is your attitude toward current TV shows and why?

6) What are your ideas on space travel in the future?

7) What makes someone a good leader and why?

8) What makes a story interesting?

9) What is your opinion on learning a second language?

Set 2: Problem-resolution opinion questions

1) Jimmy asks his dad why it's raining, and his dad says it's because there are clouds in the sky. Jimmy concludes that every time it's cloudy, it must rain.

-What is your attitude toward Jimmy's conclusion?

-How can Jimmy's father correct Jimmy's belief?

-What is your impression of Jimmy?

2) Gary tells Jane that he is going to meet her for breakfast at 8, but then accidentally sleeps in until 9.

-What should Gary do?

-How can Gary avoid this problem in the future?

-What is your impression of Gary?

3) Charlotte is planning a big party for Saturday night. She sends out the invitations, but realizes afterward that the date on the invitation say Friday instead of Saturday.

-What should Charlotte do?

-What is an alternative way to solve this problem?

-What is your impression of Charlotte?

[Group 4:]

Instructions: 你在回答问题时不要停，我说停，你才停。如果你有问题，我只能对你说，“这要取决于你自己”或者说“无论你怎样理解都可以”。 我不影响你的观点。

Set 1: Self-opinion questions

1) 你对素食主义者有什么看法?

2) 世界上最勇猛的动物是什么? 为什么?

3) 夏季最好的旅游地点是哪里? 为什么?

4) 学校最难的课程是什么? 为什么?

5) 你对当前的电视节目有什么看法? 为什么?

6) 你对未来的太空旅行有什么看法? 为什么?

7) 怎样才能成为一名 好的领导者? 为什么?

8) 怎样才能写出一个有趣的故事?

9) 你对学习第二种语言有什么看法?

Set 2: Opinion of others questions

1) 亮亮问爸爸, 天为什么会下雨？爸爸回答这是因天上有云彩。于是，亮亮的结论是，每当天上有云，天就会下雨。

-你对亮亮的结论是怎样看的?

-亮亮的爸爸应该怎样纠正亮亮的看法?

-你对亮亮有什么印象?

2)健健告诉 红红他明早八点钟和她吃早饭。 可是他早上睡过头了，九点才醒 -健健应该怎么办?在

-健健以后应该怎样做才能不犯同样的错?

-你对健健有什么印象?

3) 霏霏计划星期六晚上办一个大宴会. 她把请柬寄了出去。但是她发现她把日期写错了，写成了星期五，而不是星期六。

-霏霏应该怎么办?

-有什么办法可以补救错误?

-你对霏霏有什么印象?

Appendix 11: Semantic annotation instructions and example

"PREP:

1) Download the excel sheet "interratercodingsheet.xlsx" and the folders titled "annotation sound files" and "pitch contour images".

2) Click on the folder "annotation sound files". Notice how there are Groups 1, 2, and 3 subfolders. Make sure your computer is set to view files by name from A-to-Z. We will start with group 1, then 2, then 3. If you click on the folder "pitch contour images" you will see the exact same set up.

3) Open up the excel sheet. Notice how the file names go alphabetically from A-to-Z, resetting twice as the list goes down to correspond switching from Group 1, to 2, to 3 subfolders.

a. The first column ('file') is the .wav file name. Second column ('at time:') is at which time in the .wav file the annotation unit occurs. The third column ('coding') is for you to write down your semantic annotation of the unit. The fourth column ('transcription') is a basic transcription of the unit for annotation for guidance.

4) Import a group's subfolder of "annotation sound files" into praat (start with 1, then 2, then 3, to follow the order in the excel sheet). These should come up alphabetically to match the order in the excel sheet. Note that some files contain multiple units within the same file occurring at different times.

a. Click 'view and edit' and play the file starting at the time listed in the "at time" column corresponding to the file name on the excel sheet.

b. If the unit is a sentence (refer to transcription column), then there is no corresponding image under "pitch contour images." If the unit is only a construct (i.e 'I think', 'I feel like' etc.) then go into the group subfolder under "pitch contour images" and open it up. You will refer to the pitch contour seen in this image, as well as your own interpretation of pitch from listening to the .wav

file (including surrounding context before and after the construct), in order to make your annotation decision.

CODING:

5) (note: not all choices will necessarily be represented in subsample.)

a. for sentences: Coding decisions are based on grammatical structure and meaning. There are 5 choices. The first four are labeled as "Direct," meaning:

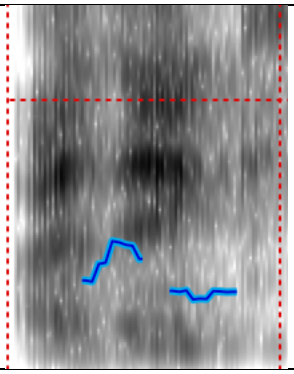
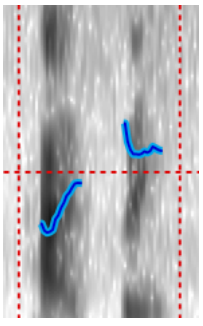
i. "Sentence cannot contain any weak modality (such as "sometimes" or "maybe") or qualifiers ("to me" "for some people") or negations (N is not) which would contradict the directness of the statement."

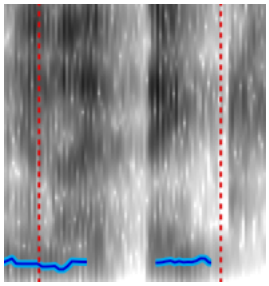
Annotation	Definition	Example
DIR (N is)	Direct sentence with structure [Noun][is…X] which expresses speaker's opinion. N is = It is, that is, I am, he/she is, you/they are…etc.	YES: That's really great. YES: The best color is green. NO: Perhaps that's great. (weak modality) NO: Green is a color. (fact, not opinion)
DIR (partial)	Direct sentence that appears as a partial utterance. In reality, it may correspond as a clipped variation of DIR (N is), such that the first part of the sentence is omitted, i.e. "(The best movie is) X". Think of it as a shortened answer to a posed question, i.e. "What is the best movie?" "X."	YES: Definitely lord of the rings. YES: By practicing every day. NO: Lord of the rings I guess.
DIR (command)	Direct sentence that appears as a verbal command.	YES: Go to the store. YES: Make sure he doesn't come. NO: Maybe go to the store.
DIR (fact)	A DIR (N is) structured statement which is not an opinion, i.e. a direct statement of fact.	YES: I am eight years old. YES: You are last place in the race. NO: I am old enough to go out alone. (opinion) NO: You are bad at sports (opinion)

Not DIR	Not a direct sentence, i.e. not any of the above four choices. Could be due to weak modality or differing sentence structure.	YES: Rabbits could be a good pet. (not N is structure) YES: Sometimes he is nice. (weak modality) YES: This is not great. (negation)

b. For constructs: Refer to the pitch contour image alongside playing the clip; the part for analysis is within the dotted red rectangle. These should have the same name as the .wav clip. If there are multiple units within the same file, then whichever unit comes second has a _2 after the file name, i.e. "kesummer_2.png".

c. All constructs follow the basic structure [I + Verb of cognitive belief]. There are 5 possible choices (note that DMF, DMM, and DME are the exact same pitch contour and differ only in syntactic position):

Annotation	Definition	Example image	Syntactic position
MC	Main clause construct. Stress (raised pitch) on the "I", therefore pitch contour falls from H-to-L. If V="**know**" then automatically is MC, regardless of stress.		Front of sentence
EM	Epistemic marker construct. Stress (raised pitch) on the "V", therefore pitch contour rises from L-to-H. If negation of V (i.e. "I don't know", "I don't think" etc) then automatically is EM due to weak modality.		Front of sentence
DMF	Discourse marker construct. No stress at all,		Front of sentence. Can be preceded by modifiers

	therefore pitch contour does not rise or fall (relatively stable).	and  (remember: it's about pitch contour, not absolute pitch)	etc. that are not part of the main clause, i.e. "In my opinion [I think] she is cool" is fine.
DMM			Middle of sentence. Appears to almost break up the sentence, i.e. "She is I think cool."
DME			End of sentence, i.e. "She is cool I guess".

Analyze pitch contour (with corresponding sound playing) first. If it is a DM, then listen to the context before and after the unit to determine syntactic position and identify which DM.

Simply write down which annotation you choose in column 3 under 'coding.' Remember: not all annotations are necessarily represented in this subsample, so choose based on the criteria above.

Sentence = DIR (N is), DIR (partial), DIR (command), DIR (fact), not DIR.

Construct = MC, EM, DMF, DMM, DME.

☺"

Example of English coding sheet:

File	at time:	Coding	Transcription
AA	28.5		I think
AB	10		I guess
AC	15.5		Funny story is interesting
AD	0		I think
AE	4.6		I think
AF	7.2		I guess

Example of Chinese coding sheet:

File	at time:	Coding	Transcription
AA	0		我觉得
AB	11.4		我觉得
AC	12.3		我觉得
AD	6.1		它们都是群居的
AE	16.2		感觉
AF	4.9		感觉

Appendix 12: Phase 2 Pilot data

PARTICIPANTS

Sample size was N = 16 (4 per group). Participants were recruited via personal contacts and snowball sampling.

Mean age per group is provided in Table A5. A one-way ANOVA showed no statistically significant main effect of group for age of participants ($F(1,3) = .771$, $p < .532$) , with post-hoc paired t-testing showing no significant difference in mean age between any groups ($p = 1.00$). Thus, there appears to be no confounding variable from age, and for the pilot phase ANOVA will be run without a covariate.

Table A5: Means and SDs for age of participants per group

Group	Mean age	SD
1	22.50	1.73
2	28.50	10.47
3	27.25	6.56
4	23.50	1.29
Total (average)	25.25	6.13

LENGTH OF AUDIO RECORDING

Part 1: Self-opinion

Means and SDs for utterances per prompt are presented in Table A6.

Table A6: Means and SDs for utterances per prompt for Self-opinion set

Group	Mean utterances/prompt	SD
1	4.88	.31
2	5.09	.40
3	4.47	1.34
4	5.67	1.05
Total (average)	5.03	.91

A one-way ANOVA showed no significant effect of group on mean utterances per prompt ($F(1,3) = 1.26$, $p = .33$). Additionally, post-hoc paired t-testing showed no significant difference between any groups ($p = .49$ between Native CHI and EFL CHI); $p = 1.00$ between all other comparisons). Thus, data length between participants should not act as a confounding variable.

Part 2: Opinion of others

Means and SDs for utterances per prompt are presented in Table A7.

Table A7: Means and SDs for utterances per prompt for Opinion of others set

Group	Mean utterances/prompt	Standard deviation
1	4.58	.30
2	4.08	.28
3	4.32	.82
4	4.58	.91
Total (average)	4.39	.62

A one-way ANOVA shows no significant effect of group on mean utterances per prompt ($F(1,3) = .57$, $p = .65$). Additionally, post-hoc paired t-testing showed no significant difference between any groups ($p = 1.00$ between all comparisons). Thus, the length of data between participants should not act as a confounding variable.

FREQUENCY OF FUNCTIONAL [I + VERB] USAGE

Three analyses were carried out on the data from the production task as presented in section 4.4.3. Results are presented for each section of the production task below.

Percentage of pragmatic function of [I + verb] to total utterances produced:

To investigate the frequency of pragmatic function of [I + verb] usage (RQ 3), a repeated measures ANOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at two levels (US, CHI), with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

Part 1: Self-opinion

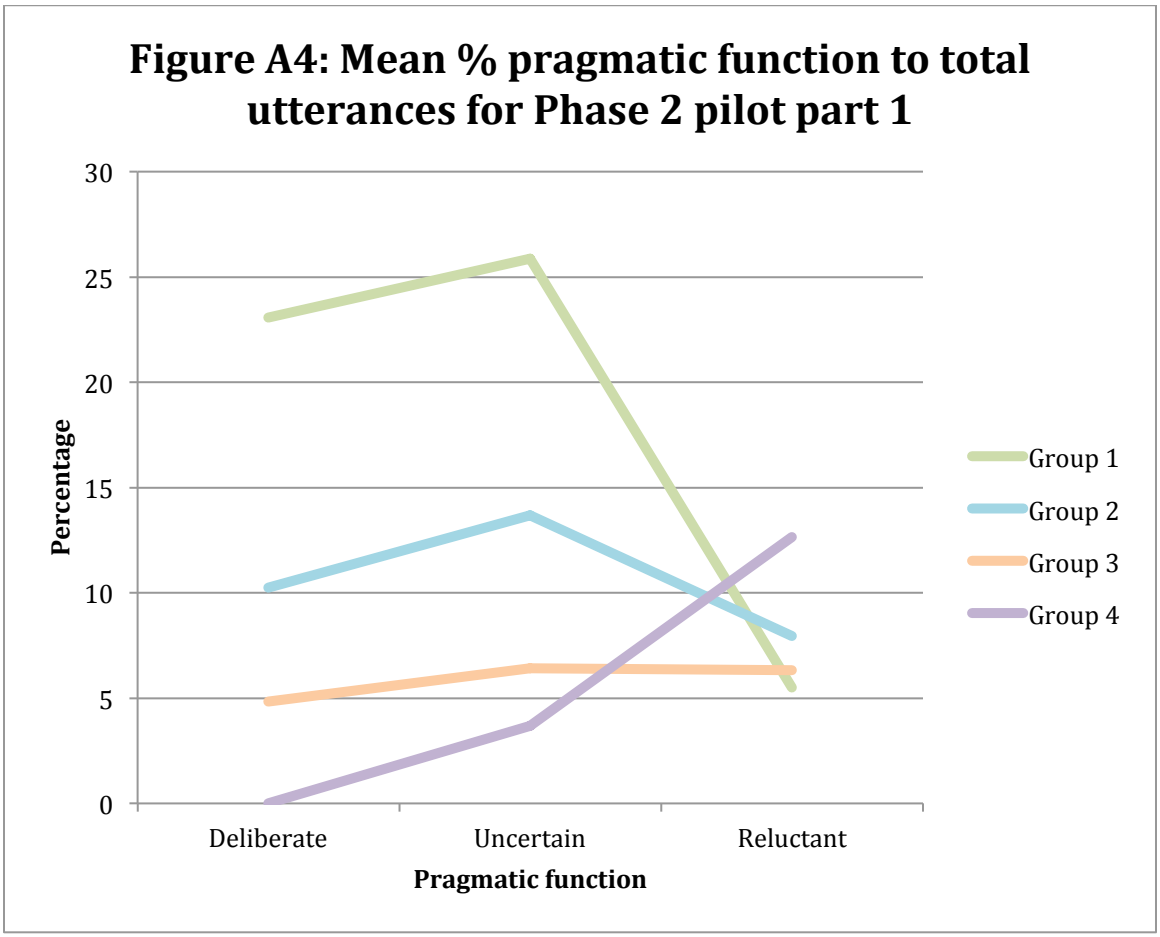
Means for Percentage of pragmatic function to total utterances are presented in Table A8.

Table A8: Means and SDs for % pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	1	23.08	9.83
	2	10.26	5.32
	3	4.83	5.94
	4	0	0
	Total	9.54	10.55
Uncertain (U)	1	25.87	10.08
	2	13.69	12.28
	3	6.43	8.32
	4	3.69	4.14
	Total	12.42	12.10
Reluctant (R)	1	5.50	2.901
	2	7.95	2.12
	3	6.33	6.00
	4	12.65	9.55
	Total	8.11	6.02

Notice the trend within each condition: % Deliberate decreases as it travels from group 1 to 2 to 3 to 4, with more US sociolinguistic background mirroring higher % Deliberate and more CHI sociolinguistic background mirroring lower % Deliberate. The data for % Uncertain mirrors this trend as well. For % reluctant, the trend is the opposite: there is an increase in % from group 1 toward group 4, with more US sociolinguistic background mirroring lower % Reluctant and more CHI sociolinguistic background mirroring higher % Reluctant. Additionally, group 2 (EFL speakers in US) mirrors group 1 (US) more than group 3 (EFL speakers in China) does, whilst group 3 mirrors group 4 (CHI). Overall, it appears that the more aggressive and confident the pragmatic function, the more it increases with US sociolinguistic background. Conversely, the

less confident the pragmatic function, the more it increases with CHI sociolinguistic background. These trends may be observed in Figure A4.



Mauchly's test indicated sphericity ($p = .74$) and Levene's test indicated homogeneity of variance within all levels ($p > .05$), signifying that ANOVA testing may proceed. Results showed a large between-group effect of sociolinguistic background ($F(1,3) = 10.63$, $p = .001$, $\eta_p^2 = .73$). Additionally, while there was no significant main effect of pragmatic function ($F(1,2) = 1.28$, $p = .30$, $\eta_p^2 = .10$), there did appear a large interactional effect between pragmatic function and group ($F(1,6) = 3.41$, $p = .01$, $\eta_p^2 = .46$).

This interaction was further investigated through post-paired t-testing on all three pragmatic function conditions between-groups and is presented in Table A9, with significant interactional differences in bold.

Table A9: Post-hoc paired t-testing for % pragmatic function to total utterances

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
1 to 2	Deliberate	12.83	.09	4.48
	Uncertain	12.18	.52	6.51
	Reluctant	-2.45	1.00	4.12
1 to 3	**Deliberate**	**18.26**	**.01**	**4.48**
	Uncertain	19.44	.07	6.51
	Reluctant	-.83	1.0	4.12
1 to 4	**Deliberate**	**23.08**	**.001**	**4.48**
	Uncertain	**22.18**	**.001**	**6.51**
	Reluctant	-7.15	.68	4.12
2 to 3	Deliberate	5.43	1.00	4.48
	Uncertain	7.26	1.00	6.51
	Reluctant	1.62	1.00	4.12
2 to 4	Deliberate	10.25	.25	4.48
	Uncertain	10.00	.90	6.51
	Reluctant	-4.70	1.00	4.12
3 to 4	Deliberate	4.83	1.00	4.48
	Uncertain	2.74	1.00	6.51
	Reluctant	-6.32	.94	4.12

Whilst visual exploration of Figure A4 may suggest differences between many interactional levels, it appears that significant differences appear between Native US and EFL CHI groups at Deliberate condition and between Native US and Native CHI groups at Deliberate and Uncertain condition.

To investigate the between-group effect of sociolinguistic background, post-hoc paired t-testing was conducted. Mean differences appear between Native US and EFL CHI groups ($p <$

.01) and Native US and Native CHI groups ($p < .01$), with no statistically significant mean difference appearing between any other pairs ($p = .08$ between Native US and EFL US; $p = .52$ between EFL US and EFL CHI; $p = .40$ between EFL US and Native CHI; $p = 1.00$ between EFL CHI and Native CHI). While these latter pairings do not show significant difference, it is interesting to note that the p value approaches significance value with Native CHI comparisons the more US linguistic and sociolinguistic background the comparison group possesses; that is, the p value between Native CHI and EFL US groups is far closer to suggesting a significant mean difference than between Native CHI and EFL CHI groups.

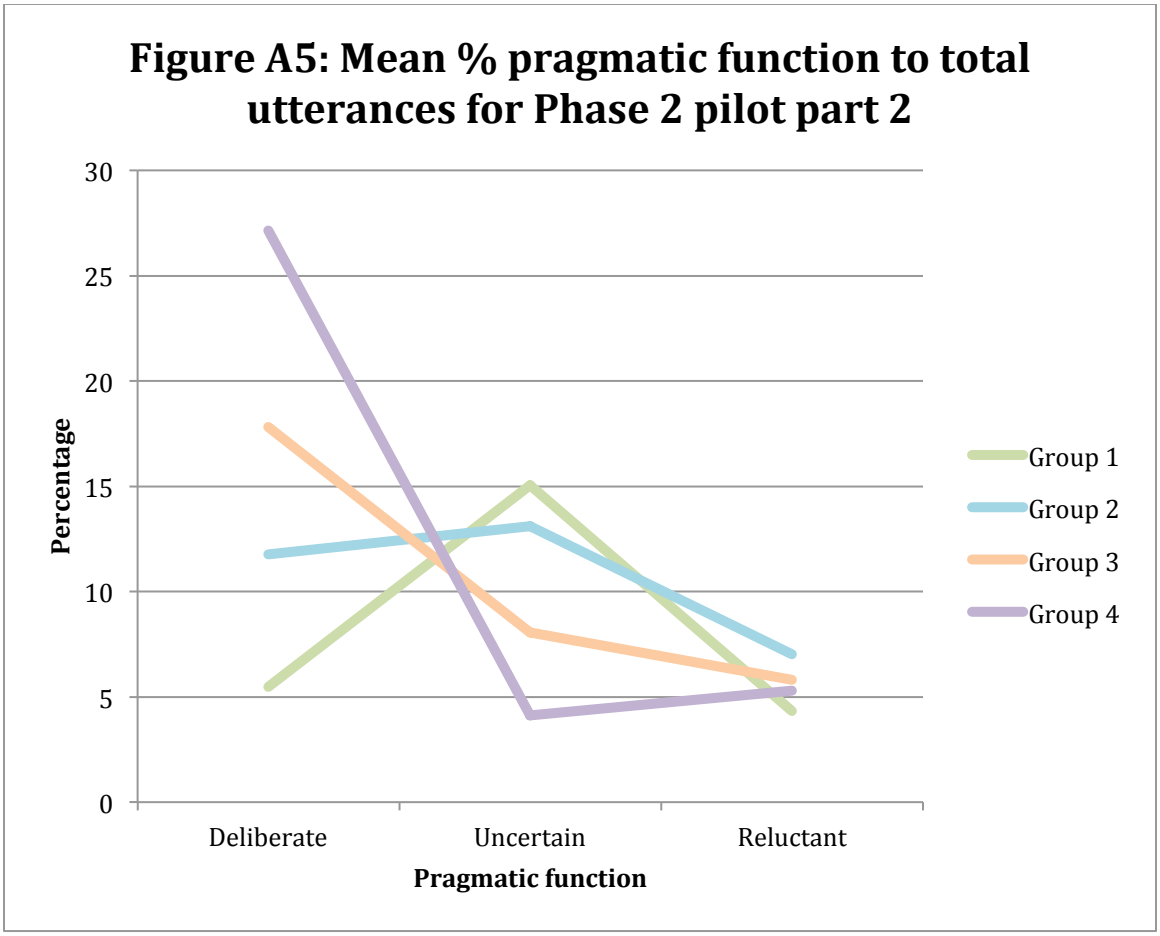
Part 2: Opinion of others

Means for Percentage of pragmatic function to total utterances are presented in Table A10.

Table A10: Means and SDs for % pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	1	5.48	3.28
	2	11.76	9.18
	3	17.80	8.40
	4	27.15	6.47
	Total	15.55	10.47
Uncertain (U)	1	15.06	8.91
	2	13.09	12.66
	3	8.05	13.12
	4	4.12	4.75
	Total	10.08	10.32
Reluctant (R)	1	4.33	8.66
	2	7.02	3.46
	3	5.81	4.65
	4	5.30	6.31
	Total	5.62	5.54

Notice that the inverse in trend of Percentage of pragmatic function to total utterances for Deliberate and Reluctant conditions between part 1 and part 2 of the production task: For part 2, % Deliberate increases as it moves from Native US to EFL US to EFL CHI to Native CHI; for Reluctant condition, % is lowest for group 1 and highest for group 2, but with little mean difference between-groups. Trend for % Uncertain still mirrors that of part 1, with % decreasing from Native US to EFL US to EFL CHI to Native CHI; however, mean differences between groups as not as large as with part 1. These trends may be observed in Figure A5.



While sphericity cannot necessarily be assumed ($X^2(1,2) = 6.16$, $p = .05$), results calculated with sphercity vs. Huynh-Feldt measurements (.96) do not differ. Additionally, Levene's test

indicated homogeneity of variance at all levels ($p > .05$). Thus, ANOVA testing will be conducted.

Results revealed a significant within-subject effect of pragmatic function ($F(1,2) = 5.22$, $p = .01$, $\eta_p^2 = .30$). No between-subject effect of sociolinguistic background was found $F(1,3) = .67$, $p = .59$, $\eta_p^2 = .14$). However, there was a large interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 2.72$, $p = .04$, $\eta_p^2 = .41$).

To investigate the within-group effect of pragmatic function, post-hoc paired t-testing was conducted. A mean difference of 5.46 appears between Deliberate and Reluctant conditions (std. error = 2.35, $p < .01$) but not between Deliberate and Uncertain ($p = .58$) or Uncertain and Reluctant conditions ($p = .38$). Additionally, paired-samples t-tests confirm this result, with a significant difference ($t(1,15) = 3.39$, $p < .01$) appearing between Deliberate and Reluctant conditions but no significant difference ($p > .05$) between any other pairings.

To investigate this interactional effect, post-hoc paired t-testing was conducted between-groups on all three pragmatic function conditions. Results are presented in Table A11, with significant interactional differences in bold.

Table A11: Post-hoc paired t-testing for % pragmatic function to total utterances

Group comparison	Pragmatic function	Mean % difference	Significance	Std. Error
1 to 2	Deliberate	-6.28	1.00	5.09
	Uncertain	1.97	1.00	7.37
	Reluctant	-2.69	1.00	4.31
1 to 3	Deliberate	12.32	.19	5.09
	Uncertain	-7.01	1.00	7.37
	Reluctant	-1.48	1.00	4.31
1 to 4	**Deliberate**	**-21.67**	**.01**	**5.09**
	Uncertain	10.94	.98	7.37
	Reluctant	-.97	1.00	4.31

2 to 3	Deliberate	-6.04	1.00	5.09
	Uncertain	5.04	1.00	7.37
	Reluctant	1.21	1.00	4.31
2 to 4	Deliberate	-15.39	.06	5.09
	Uncertain	8.97	1.00	7.37
	Reluctant	1.72	1.00	4.31
3 to 4	Deliberate	-9.35	.55	5.09
	Uncertain	3.93	1.00	7.37
	Reluctant	.51	1.00	4.31

It is interesting to note that within the Deliberate condition, there is no significant mean difference between groups 1 and 2, but there is a near significant mean difference between groups 2 and 4. This may suggest a shift in cultural mindset in group 2 (EFL speakers in US), resulting in these individuals changing their pragmatic speaking style to mirror that of US native speakers vs. CHI speakers in their old cultural environment.

Percentage of pragmatic function of [I + verb] construct to total constructs used

To further explore the frequency of pragmatic function of [I + verb] usage (RQ 3), a second repeated measures ANOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at two levels (US, CHI), with a DV of Percentage of pragmatic function of [I + verb] to total [I + verb] constructs produced.

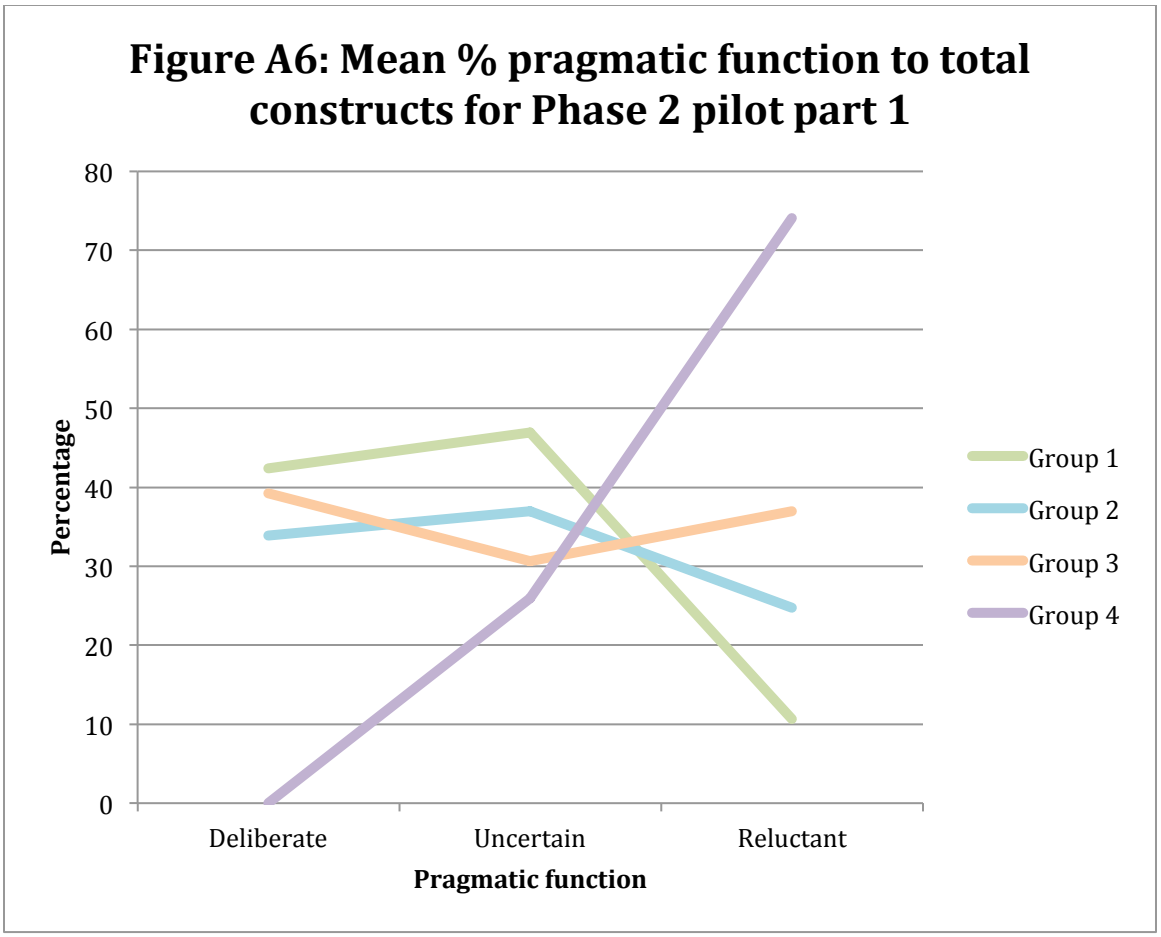
Part 1: Self-opinion

Means for Percentage of pragmatic function to total utterances are presented in Table A12.

Table A12: Means and SDs for % pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	1	42.40	11.86
	2	33.90	20.35
	3	39.21	45.36
	4	0	0
	Total	28.88	28.79
Uncertain (U)	1	46.94	8.33
	2	36.94	20.63
	3	30.63	33.88
	4	25.94	28.01
	Total	35.11	23.48
Reluctant (R)	1	10.66	5.21
	2	24.75	5.49
	3	36.99	31.02
	4	74.06	28.01
	Total	36.62	30.85

Notice a decreasing trend in both Deliberate and Uncertain conditions as it travels from Native US to EFL US to EFL CHI to Native CHI; this trend matches the trend for these conditions in the first analysis. Similarly, the increasing trend in Reluctant condition moving from Native US to EFL US to EFL CHI to Native CHI also mirrors this condition's results in the first analysis. These trends may be observed in Figure A6. Once again, Native US and Native CHI groups appear to have inversely correlated trends regarding confidence of pragmatic function. As predicted, EFL US's performance appears more closely correlated with Native US's performance vs. Native CHI, whilst EFL CHI's performance tends to appear more similar to that of Native CHI.



While Levene's test indicated homogeneity of variance within Uncertain condition ($p = .19$), homogeneity of variance may not necessarily be assumed at Deliberate ($p < .001$) and Reluctant conditions ($p = .01$). Nonetheless, Mauchly's test indicated sphericity ($p = .92$), and ANOVA testing was conducted.

No significant within-subject effect of pragmatic function ($F(1,2) = .32$, $p = .73$, $\eta_p^2 = .25$) or between-subject effect of sociolinguistic background ($F(1,3) = 1.30$, $p = .32$, $\eta_p^2 = .03$) was found. However, as predicted, there appeared a significant interactional effect between sociolinguistic background and pragmatic function ($F(1,6) = 2.84$, $p = .03$, $\eta_p^2 = .42$).

Post-hoc paired t-testing was conducted to investigate this interaction. While no significant mean differences were found between-groups in Deliberate ($p = .22$ for Native Us and Native

CHI groups, p = .51 for EFL US and Native CHI groups, p = .31 for EFL CHI and Native CHI groups, p = 1.00 for rest) and Uncertain (p = 1.00) conditions, significant mean differences did appear within the Reluctant condition. Table A13 presents these results, with significant mean differences appearing in bold.

Table A13: Post-hoc paired t-testing for % Reluctant to total constructs produced

Group comparison	Mean % difference	Significance	Std. Error
1 to 2	-14.09	1.00	15.02
1 to 3	-26.32	.63	
1 to 4	**-63.40**	**<.01**	
2 to 3	-12.23	1.00	
2 to 4	**-49.31**	**.04**	
3 to 4	-37.08	.18	

These findings suggest that for Self-opinion prompts, Native CHI speakers preferred using less confident [I + verb] constructs, whilst Native US speakers preferred using more confident [I + verb] constructs. EFL CHI group's performance appeared to mirror that of Native CHI group, given that they share in CHI sociolinguistic background; similarly, Native US group's performance mirrored that of Native US group, with both groups sharing in US sociolinguistic background.

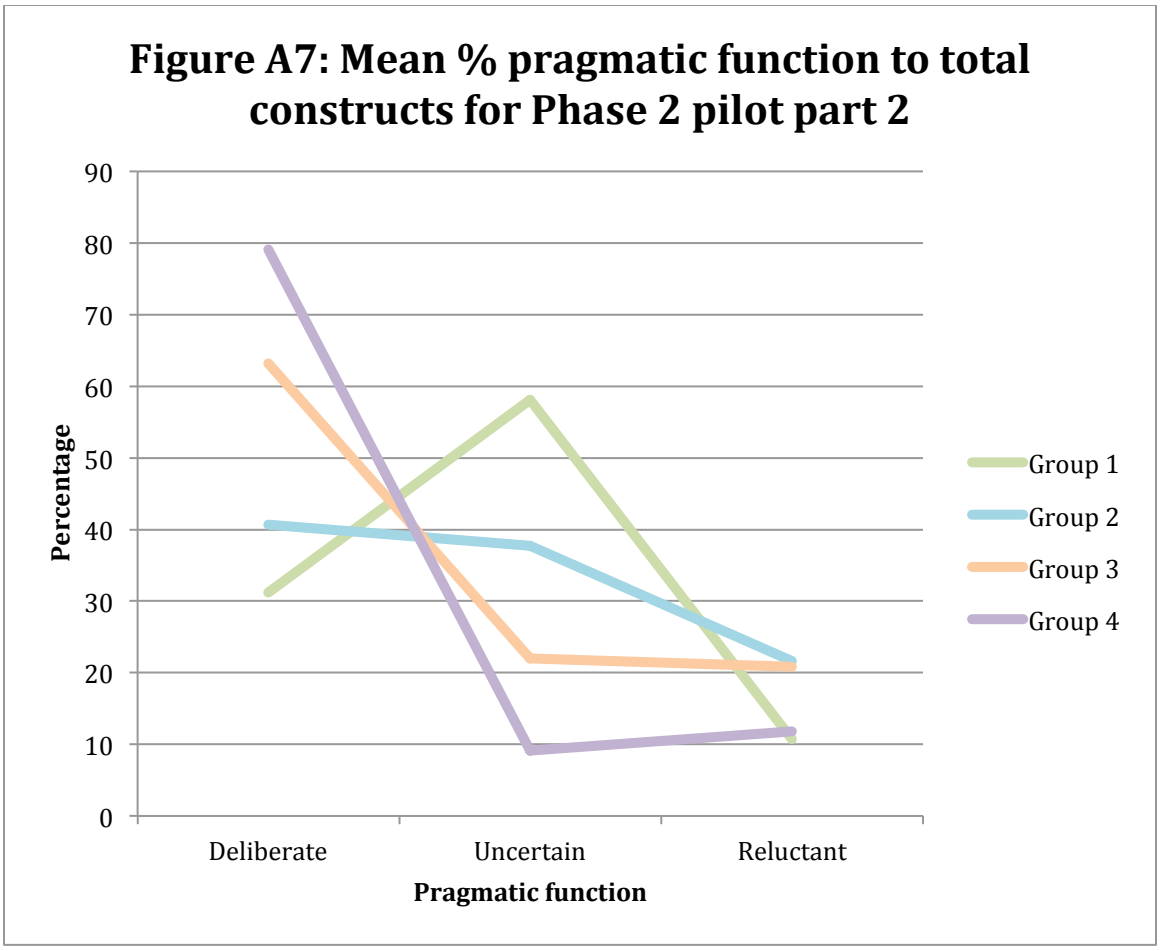
Part 2: Opinion of others

Means for Percentage of pragmatic function to total constructs are presented in Table A14.

Table A14: Means and SDs for % pragmatic function to total utterances

Pragmatic function	Group	Mean	SD
Deliberate (D)	1	31.19	28.01
	2	40.63	24.09
	3	63.15	30.69
	4	79.10	22.61
	Total	53.52	30.66
Uncertain (U)	1	58.10	22.11
	2	37.75	25.13
	3	21.99	35.11
	4	9.10	9.42
	Total	31.73	29.07
Reluctant (R)	1	10.72	21.43
	2	21.61	2.83
	3	20.81	16.97
	4	11.81	13.68
	Total	16.24	14.67

For Opinion of others prompts, trends for all three conditions match the trends for these conditions in analysis 1 and may be observed in below in Figure A7.



Mauchly's test indicated sphericity (p = .12) and Levene's test indicated homogeneity of variance in all levels (p = .96 for Deliberate, p = .33 for Uncertain) except Reluctant (p = .02). Given these results, ANOVA testing was conducted.

Results revealed a large within-subject effect of pragmatic function ($F(1,2) = 7.30$, $p < .01$, $\eta_p^2 = .38$). However, there was no significant between-group effect of sociolinguistic background ($F(1,3) = 1.00$, $p = .43$, $\eta_p^2 = .20$) or interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 2.46$, $p = .05$, $\eta_p^2 = .38$).

Post-hoc paired t-testing showed a significant difference between Uncertain and Reluctant levels (p = .003) but not between Deliberate and Uncertain (p = .22) or Reluctant (p = .30) pairings. Additionally, paired sample t-tests were conducted to investigate the within-subject

effect of pragmatic function. Results confirmed a significant difference between Uncertain and Reluctant levels ($t(1,15) = 4.02$, $p = .01$).

Calculated CR of speaker confidence:

CR values were calculated for each participant based on CR values of prosodic-syntactic variations from Proof of Concept results. Each prosodic-syntactic variation was multiplied by the percentage of total constructs comprised of that particular prosodic variation for each participant, resulting in an assigned average CR for speaker confidence for each participant.

To investigate the effect of sociolinguistic background on CR values of speaker confidence, a one-way ANOVA was run with one between-group factor (sociolinguistic background) measured at four levels (Native US, EFL US, EFL CHI, Native CHI), with average CR as the dependent variable.

Part 1: Self-opinion

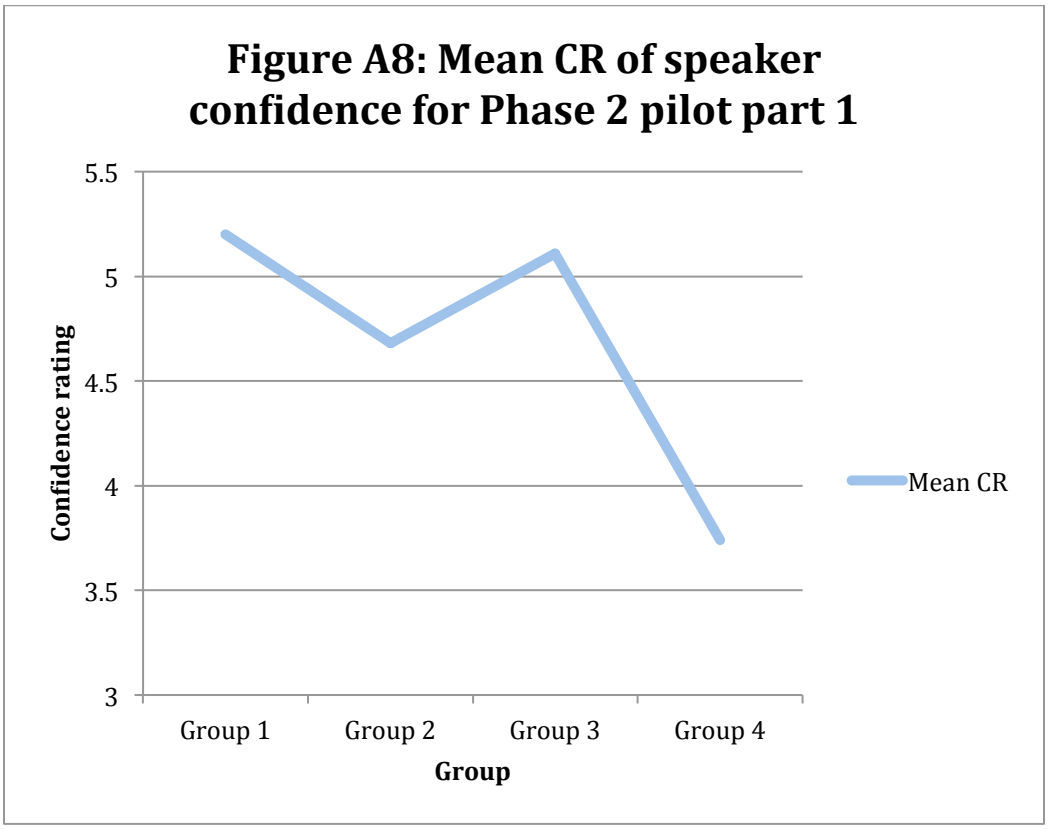
Means for CR of each group are presented below in Table A15.

Table A15: Means and SDs for CR of speaker confidence

Group	Mean	SD
Native US	5.20	.18
EFL US	4.68	.60
EFL CHI	5.11	1.20
Native CHI	3.74	.36
Total	4.68	.87

As predicted, CR is highest for Native US and lowest for Native CHI. However, predicted trend shifts between EFL US and EFL CHI, with EFL CHI exhibiting a higher CR that is closer

to that of Native US, whilst EFL US's CR value is lower than that of EFL CHI's CR value. This trend may be observed in Figure A8.



Levene's test suggested homogeneity of variance ($p = .09$), indicating that ANOVA may be conducted. Results showed a significant between-group effect of sociolinguistic background ($F(1,3) = 3.74$, $p < .05$). To investigate this effect, post-hoc paired t-testing was conducted. While results showed no significant between-group differences ($p = .08$ between Native US and Native CHI groups; $p = .10$ between EFL CHI and Native CHI groups, $p = .50$ between EFL US and Native CHI groups, and $p = 1.00$ for rest), it can be noted that the mean difference of 1.45 between groups 1 and 4 was closer to significance level than any others for CR of Self-opinion prompts.

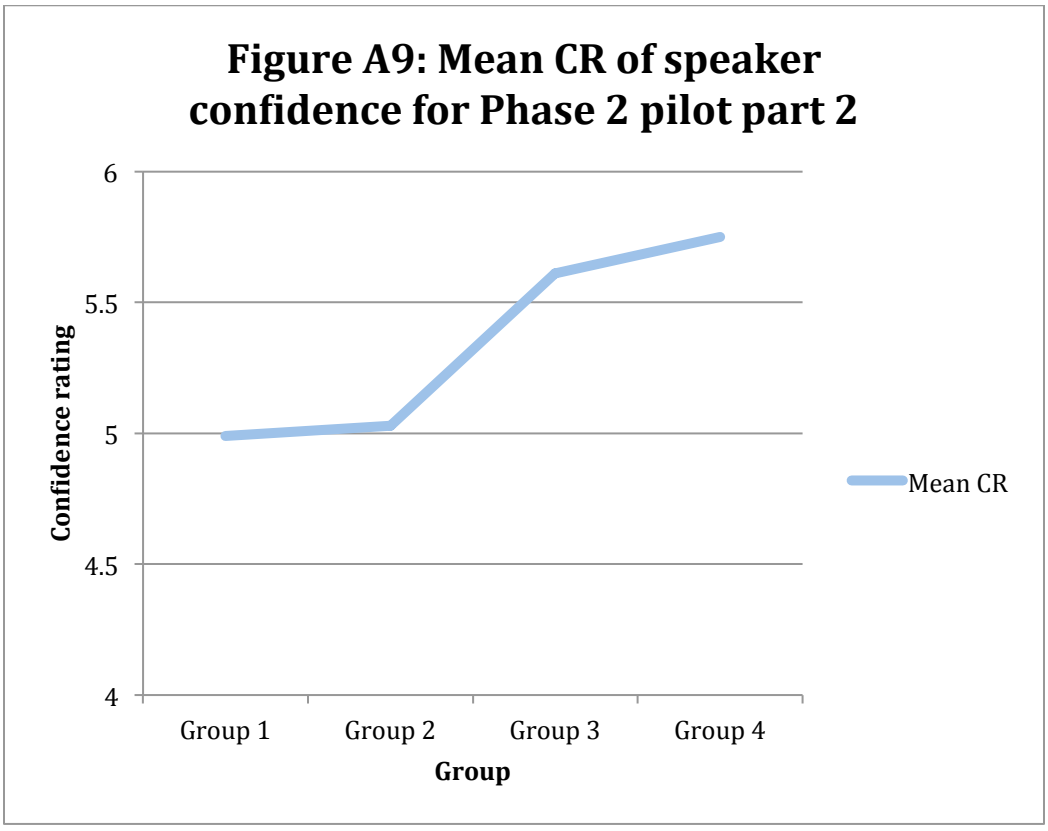
Part 2: Opinion of others

Means for CR of each group are presented below in Table A16.

Table A16: Means and SDs for CR of speaker confidence

Group	Mean	SD
1	4.99	.55
2	5.03	.36
3	5.61	.68
4	5.75	.47
Total	5.34	.59

As predicted, there appears to be an increasing CR trend from group 1 to 2 to 3 to 4. Compared to part 1, mean CRs are closer in value. This trend is visible below in Figure A9.



Levene's test suggested homogeneity of variance ($p = .35$), indicating that ANOVA may be conducted. Results however showed no between-subject effect of sociolinguistic background ($F(1,3) = 2.19$, $p = .14$). It must once again be noted that due to the small sample size of this pilot, results must be interpreted with extreme caution. Nevertheless, pilot results do show that the production task was successful in eliciting production of various [I + verb] constructs from participants, thereby showing promise for data collection in the main phase.

Appendix 13: ANCOVA results of Confidence Rating of [I + verb] prosodic variations with potential outliers removed

ANCOVA results below were run with participant 4 (US group) and participant 30 (US group) removed, resulting in n = 28 for US group 1 and n = 29 for CHI group 2. It may be noted that results do not differ in significance with results that do not exclude these two cases. Table A17 presents the mean CRs of prosodic variations by group.

Table A17: Means, SDs, and confidence intervals of prosodic variations.

Kontrast Placement	Sociolinguistic Background	Mean	SD
Direct (DIR)	US	6.16	.72
	CHI	6.19	.87
	Total	6.18	.79
Pronoun (PRN)	US	6.11	.60
	CHI	5.23	.89
	Total	5.67	.90
Verb (V)	US	4.45	.75
	CHI	4.90	.83
	Total	4.68	.82
No Kontrast: Sentence initial (FRONT)	US	3.68	1.12
	CHI	4.4	1.38
	Total	3.96	1.30
No Kontrast: Within-sentence (MID)	US	3.14	1.03
	CHI	3.38	1.10
	Total	3.26	1.07
No Kontrast: Sentence final (END)	US	3.45	.84
	CHI	3.50	.96
	Total	3.47	.90

While sphericity cannot necessarily be assumed ($X^2(1,14) = 26.95$, $p = .02$), results calculated with Mauchly's sphericity vs. Greenhouse-Geiser (.82) or Huynh-Feldt measurements

(.93) do not differ regarding significance. Since the Greenhouse-Geiser measurement did not fall below .75, the Hunyh-Feldt estimate was chosen as a correction for sphericity given that Greenhouse-Geiser measurements may be too conservative (Hyunh & Feldt, 1976, Field, 2013:1778). Levene's test indicated homogeneity of variances between all levels of prosodic variation ($p > .05$).

Results showed a significant within-subject effect of prosodic variation ($F(1,4.12) = 27.15$, $\eta_p^2 = .34$, $p < .001$). Conversely, no between-subject effect of sociolinguistic background ($F = .49$, $\eta_p^2 = .01$, $p = .49$) or covariate of age ($F = .17$, $\eta_p^2 = .01$, $p < .01$) appeared. Additionally, age did not interact significantly with prosodic variation ($F(1,4.12) = 3.62$, $p = .10$). There was, however, a moderate interaction between sociolinguistic background and prosodic variation ($F(1,4.12) = 6.20$, $\eta_p^2 = .10$, $p < .001$).

To further investigate the interaction between prosodic variation and sociolinguistic background, post-hoc paired t-tests with were conducted on means adjusted for covariate of age on Pronoun, Front, and Mid conditions between-groups where any gap in confidence rating may have occurred, given that these were the conditions that showed any mean differences between groups. Pronoun level results showed a significant mean difference ($p < .01$) of .80 on a scale of 7. An independent sample t-test further suggests a significant between-group difference at Pronoun condition ($t(1,55) = 4.84$, $p < .001$). Front condition results showed no significant mean difference between groups ($p = .12$). Additionally, no significant difference between-groups occurred at Mid level ($p = .06$).

Appendix 14: Results from repeated measures ANOVA with M = 100

Pooled F-test results show that perceived pragmatic function had a strong between-subject effect ($p < .001$, $F(2,111.08)= 327.51$). Table A18 shows that CR means are highest for deliberate, then uncertain, and lowest for reluctant[60].

Table A18: Pooled mean confidence ratings and standard deviations for perceived pragmatic function based on sociolinguistic background

Perceived pragmatic function	Sociolinguistic background	Mean	Std. Error
Deliberate (D)	US	6.02	.11
	CHI	6.08	.13
	Total	6.08	.08
Uncertain (U)	US	4.26	.13
	CHI	4.28	.18
	Total	4.27	.11
Reluctant (R)	US	2.61	.14
	CHI	3.18	.18
	Total	2.89	.12

Post-hoc paired t-tests (Table A19) show significant differences between all levels of perceived pragmatic function ($p < .001$), thereby rejecting the null hypothesis and supporting Hypothesis 1.

[60] SD and confidence intervals cannot be calculated due to pooled results. Standard error is given instead.

Table A19: Post-hoc paired T-tests for pooled MI results

Pair	Mean difference	Significance	95% confidence intervals	
			Lower bound	Upper bound
D – U	1.81	.000	1.58	2.04
U – R	1.38	.000	1.14	1.61
R – D	-3.19	.000	-3.47	-2.91

For between-subject effects, sociolinguistic background was insignificant ($F(1,51.87) = .97$, $p = .33$). Covariate of age was also found to be insignificant ($F(1,108.71) = .75$, $p = .39$), mirroring ANOVA 1 and similarly supporting Hypothesis 2. A small, slightly-significant interactional effect was found between pragmatic function and sociolinguistic background ($F(2,111.31) = 3.32$, $p = .04$).

Appendix 15: Non-parametric Kruskal-Wallis test for Phase 1 calculated CR

Kruskal-Wallis results (See Table A20) did not differ from ANCOVA in terms of significance. Significant between-group differences were found within PRN, V, and F conditions.

Table A20: Kruskal-Wallis test results for Phase 1 calculated CR

	DIR	PRN	V	F	M	E
Chi-Square	.08	14.86	8.80	5.20	.07	.16
Degrees of freedom	1	1	1	1	1	1
Significance	.777	.000	.003	.023	.792	.688

Appendix 16: Effect of L2 proficiency on frequency of [I + verb]

An in-depth sub-analysis was carried out on the two EFL groups to identify whether frequency of functional [I + verb] usage varied as a function of L2 proficiency, as identified by the Lex30 score (RQ 6). For the Analysis 1 above, the same analysis was run but with EFL US and EFL CHI participants only (N = 50), subsequently followed by a similar analysis with the addition of the between-subject factor of Lex30 score, as well as the potentially confounding between-subject factors of Age of L2 acquisition and Native language of L2 instructor(s). The first N = 50 analysis serves to act as a comparison model for the second analysis, which acts to answer RQ 6 as well as to control for potential confounds.

Results are presented below for both prompt sets of the open production task. For each prompt set, differences between the original N = 100 ANCOVA results and subsection N = 50 results are first presented, followed by a comparison with N = 50 ANCOVA with inclusion of covariates.

Part 1: Self-opinion

ORIGINAL MODEL: To act as a comparison model for the second set (which serves to answer RQ 6 and control for confounds), a repeated measures ANCOVA was run with one at N = 50 within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at two levels (EFL US, EFL CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 9.39, p = .01), results calculated with Mauchly's sphericity vs. Huynh-Feldt measurements (.83) do not differ regarding significance.

Whilst Levene's test did not confirm homogeneity of variance at at Uncertain condition ($p < .01$), it did indicate homogeneity of variance at Deliberate ($p = .59$) and Reluctant ($p = .10$) conditions.

ANCOVA results revealed no between-group effect of sociolinguistic background ($F(1,1) = 1.22$, $p = .28$, $\eta_p^2 = .03$, OP = .19). Additionally, there was a significant within-group effect of pragmatic function ($F(1,2) = 6.63$, $p < .01$, $\eta_p^2 = .12$, OP = .86). Moreover, results also showed a large interactional effect between pragmatic function and group ($F(1,2) = 13.43$, $p < .001$, $\eta_p^2 = .22$, OP = .99). There did not appear to be a statistically significant effect of age ($F(1,1) = .39$, $p = .54$, $\eta_p^2 = .03$, OP = .09), but results showed an interactional effect between pragmatic function and age ($F(1,2) = 5.03$, $p = .01$, $\eta_p^2 = .10$, OP = .75).

Results between this repeated measures ANCOVA subsection run with N = 50 and the original repeated measures ANCOVA run with the whole sample (N = 100) are similar in terms significant interactional effect between pragmatic function and group, differing only in that the N = 50 ANCOVA also found an interactional effect between pragmatic function and age. However, whilst for the N =50 ANCOVA the observed power for the first interactional effect of pragmatic function/group, at .99, is well above the .80 threshold for reliability (Cohen, 1988), the second interactional effect of pragmatic function/age is only .75, meaning that significance value may not be trustworthy. Thus, an observation of effect size may prove more telling, and at $\eta_p^2 = .10$ it appears that age may account for 10% of variance, although this is less than half of the variance that the interactional effect between pragmatic function/group may explain.

In terms of the interactional effect, post-hoc paired t-tests showed significant mean differences between groups at Deliberate ($p < .001$) and Reluctant ($p < .001$) conditions, but not at Uncertain condition ($p = .12$). Compared to the original N = 100 ANCOVA, the significance for Deliberate and Reluctant conditions remain the same, with a slight change in Uncertain condition which appeared as borderline significant ($p = .05$) in the original N = 100 ANCOVA.

ADDITION OF COVARIATES: To investigate whether frequency of pragmatic function of [I + verb] usage varies as a function of L2 proficiency (RQ 6) as well as to account for possible confounding variables, a comparable repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and two between-group factors of sociolinguistic background measured at two levels (EFL US, EFL CHI) and Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with age,Age of L2 acquisition, and Lex30 score as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 16.66, p < .001), results calculated with Mauchly's sphericity vs. Huynh-Feldt measurements (.88) do not differ regarding significance. Whilst Levene's test was borderline significant for homogeneity of variance at Uncertain condition (p = .05), it did successfully indicate homogeneity of variance at Deliberate (p = .68) and Reluctant (p = .38) conditions.

Table A21 below presents ANCOVA results in terms of significance values (F included if significant), effect size, and observed power. Significant p values with sufficient observed power (> .80) are presented in bold; for all other cases, observed power may be too low to depend on significance values, and thus effect size should be taken into consideration instead.

Table A21: ANCOVA results of significance values, effect size and observed power

Within-subject effect	Significance	η_p^2	Observed power
Pragmatic function	.38	.02	.20
Pragmatic function * Age	.13	.05	.40
Pragmatic function * Age L2 acquisition	.60	.01	.12
Pragmatic function * lex30	.31	.03	.23
Pragmatic function * Sociolinguistic background	**.001 (F(1,2) = 9.05)**	**.17**	**.96**

Pragfn * Native language of L2 instructor	.68	.01	.10
Pragmatic function * sociolinguistic background * Native language of L2 instructor	.83	<.01	.07

Between-subject effect	Significance	η_p^2	Observed power
Age	.62	.01	.08
Age L2 acquisition	.54	.01	.09
Lex30 score	.49	.01	.10
Sociolinguistic background	.14	.05	.31
Native language of L2 instructor	.82	.001	.06
Sociolingustic background * Native language of L2 instructor	.24	.03	.22

The large interactional effect between pragmatic function and sociolinguistic background (group) mirror that of previous analyses. Note that here, the interactional effect between age and pragmatic function is no longer significant, with its effect size being halved from 10% to 5% of total variance explained; this mirrors the original analysis of N = 100 ANCOVA.

In terms of the significant interactional effect between pragmatic function and group, post-hoc paired t-tests showed significant mean differences between groups at Deliberate ($p < .001$) and Reluctant ($p < .004$) conditions, but not at Uncertain condition ($p = .12$). These results mirror the significance of results for N =50 ANCOVA without addition of covariates.

In terms of effect size, the interactional effect between pragmatic function and sociolinguistic background, as predicted, accounted for the largest percentage of variation at 17%. The variables that accounted for the second most amount of variance were sociolinguistic background and age at 5% each, followed by the interactional effect between pragmatic function and Lex30 score, as

well as interaction between sociolinguistic background and Native language of L2 instructor, at three percent.

Below in Table A22 is a comparison of effect sizes between N = 50 original ANCOVA and N = 50 with addition of covariates ANCOVA.

Table A22: Comparison of effect sizes between original N = 50 ANCOVA and N = 50 with addition of covariates ANCOVA

Within-subject effect	Original η_p^2	Additional covariates added η_p^2
Pragmatic function	.12	.02
Pragmatic function * Age	.10	.05
Pragmatic function * Sociolinguistic background	.22	.17

Between-subject effect	Original η_p^2	Additional covariates added η_p^2
Age	.03	.01
Sociolinguistic background	.03	.05

When compared to one another, it appears that the variable explaining the largest amount of total variance, pragmatic function * sociolinguistic background, was not affected by the addition of covariates, still maintaining over three times the effect size of any other variable in the additional covariates added ANCOVA. The effect of pragmatic function * age is halved, which now mirrors the original N = 100 ANCOVA results but still maintains a 5% explanation of variance nonetheless. The effect of sociolinguistic background was raised by 66.67%, but still only accounts for 5% of variance. Pragmatic function's direct effect was substantially lowered from 12% to 2% with the addition of covariates added to analysis.

In summary, it appears that for N = 50 ANCOVA of the direct-opinion prompt set, frequency of [I + verb] usage (as determined by percentage of pragmatic [I + verb] construct out of total

utterances) varies as a factor of the interaction between sociolinguistic background and pragmatic function (in accordance with previous results), with not much variation being a function of L2 proficiency or other confounding variables analyzed.

Part 2: Opinion of others

ORIGINAL MODEL: To act as a comparison model for the second set (which serves to answer RQ 6 and control for confounds), a repeated measures ANCOVA was run with one at N = 50 within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and one between-group factor (sociolinguistic background) measured at two levels (EFL US, EFL CHI), with age as a covariate, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed (X^2 (1, 2) = 12.90, $p < .01$), results calculated with Mauchly's sphericity vs. Huynh-Feldt measurements (.86) do not differ regarding significance. Levene's test indicated homogeneity of variance at all levels (p = .13 Deliberate; p = .86 Uncertain, p = .24 Reluctant).

ANCOVA results revealed a between-group effect of sociolinguistic background ($F(1,1) = 15.90$, $p < .001$, $\eta_p^2 = .25$, OP = .97). Additionally, there was a significant within-group effect of pragmatic function ($F(1,2) = 13.74$, $p < .001$, $\eta_p^2 = .23$, OP =1.00). Moreover, results also showed a large interactional effect between pragmatic function and group ($F(1,2) = 13.43$, $p = .01$, $\eta_p^2 = .12$, OP = .84). There did not appear to be a statistically significant effect of age ($F(1,1) = .22$, $p = .64$, $\eta_p^2 = .08$, OP = .09), but results showed an interactional effect between pragmatic function and age ($F(1,2) = 4.50$, $p = .02$, $\eta_p^2 = .16$, OP = .74).

Results between this repeated measures ANCOVA subsection run with N = 50 and the original repeated measures ANCOVA run with N = 100 are similar in terms significant interactional effect between pragmatic function and sociolinguistic background, differing only in

that the N = 50 ANCOVA also found a between-subject effect of age and interactional effect between pragmatic function and age. However, whilst for the N =50 ANCOVA the observed power for the first interactional effect of pragmatic function/group, at .97, is well above the .80 threshold for reliability (Cohen, 1988), the second interactional effect of pragmatic function/age was only .74 and the between-subject effect of age was only .09, meaning that these significance values may not be trustworthy. Thus, an observation of effect size may prove more telling, and at $\eta_p^2 = .16$ it appears that the interaction between pragmatic function and age may account for 16% of variance and age for 8% of variance, thereby strengthening the decision made to include age as covariate in analyses.

In terms of the interactional effect, post-hoc paired t-tests showed significant mean differences between groups at Deliberate ($p < .001$) but not Reluctant ($p = .19$) and Uncertain conditions ($p = .98$). As predicted, compared to the original N = 100 ANCOVA, these results directly mirrored significance findings at all levels.

ADDITION OF COVARIATES: To investigate whether frequency of pragmatic function of [I + verb] usage varies as a function of L2 proficiency (RQ 6) as well as to account for possible confounding variables, a comparable repeated measures ANCOVA was run with one within-group factor (pragmatic function) measured at three levels (Deliberate, Reluctant, Uncertain) and two between-group factors of sociolinguistic background measured at two levels (EFL US, EFL CHI) and Native language of L2 instructor measured at two levels (US/Non-CHI, Only CHI), with age,Age of L2 acquisition, and Lex30 score as covariates, and with a DV of Percentage of pragmatic function of [I + verb] to total utterances produced.

While sphericity cannot necessarily be assumed ($X^2 (1, 2) = 10.61$, $p = .01$), results calculated with Mauchly's sphericity vs. Huynh-Feldt measurements (.96) do not differ regarding significance. Whilst Levene's test did not necessarily indicate homogeneity of variance at

Deliberate condition ($p < .01$, it did successfully indicate homogeneity of variance at Uncertain ($p = .40$) and Reluctant ($p = .26$) conditions.

Table A23 below presents ANCOVA results in terms of significance values (F included if significant), effect size, and observed power. Significant p values with sufficient observed power (> .80) are presented in bold; for all other cases, observed power may be too low to depend on significance values, and thus effect size should be taken into consideration instead.

Table A23: ANCOVA results of significance values, effect size and observed power

Within-subject effect	Significance	η_p^2	Observed power
Pragmatic function	.30	.06	.26
Pragmatic function * Age	.19	.08	.34
Pragmatic function * Age L2 acquisition	.94	< .01	.06
Pragmatic function * lex30	.91	.01	.06
Pragmatic function * Sociolinguistic background	**.01 (F(1,2) = 5.50)**	**.21**	**.82**
Pragmatic function * Native language of L2 instructor	.66	.02	.11
Pragmatic function * sociolinguistic background * Native language of L2 instructor	.59	.02	.13

Between-subject effect	Significance	η_p^2	Observed power
Age	.33	.02	.16
Age L2 acquisition	.98	<.001	.05
Lex30 score	.61	.01	.08
Sociolinguistic background	**<.001 (F(1,1) = 17.09)**	**.28**	**.98**
Native language of L2 instructor	.43	.01	.12
Sociolinguistic background * Native language of L2 instructor	.26	.03	.20

The large interactional effect between pragmatic function and sociolinguistic background (group), as well as the large between-subject effect of sociolinguistic background, mirror that of previous analyses. Note that here, the interactional effect between age and pragmatic function is no longer significant, with its effect size being halved from 16% to 8% of total variance explained; this mirrors the original analysis of N = 100 ANCOVA, much as it was halved in the same analysis for prompt set 1 above.

In terms of the interactional effect, post-hoc paired t-tests showed significant mean differences between groups at Deliberate ($p < .001$) but not Reluctant ($p = .18$) and Uncertain conditions ($p = 1.00$). Much as it did in the same analysis for prompt set 1 above, these results mirror the significance of results for N =50 ANCOVA without addition of covariates.

In terms of effect size, the between-subject effect between pragmatic function and sociolinguistic background accounted for the largest percentage of variation at 28%. The variables that accounted for the second most amount of variance was the interaction between pragmatic function and sociolinguistic background at 21%, followed by the interactional effect between pragmatic function and age at 8%, and then pragmatic function at 6. Age of L2 acquisition as a variable accounted for almost no variance.

Below in Table A24 is a comparison of effect sizes between N = 50 original ANCOVA and N = 50 with addition of covariates ANCOVA.

Table A24: Comparison of effect sizes between original N = 50 ANCOVA and N = 50 with addition of covariates ANCOVA

Within-subject effect	Original η_p^2	Additional covariates added η_p^2
Pragmatic function	.12	.06
Pragmatic function * Age	.16	.08

Pragmatic function * Sociolinguistic background	.22	.21

Between-subject effect	Original η_p^2	Additional covariates added η_p^2
Age	.08	.02
Sociolinguistic background	.25	.28

When compared to one another, it appears that the main variance-inducing variable of pragmatic function * sociolinguistic background is actually increased by 3% to explain a total 28% of variance, followed by pragmatic function's interaction of sociolinguistic background explaining 21% of variance. The effect size was halved for the interaction between pragmatic function and age and decreased by four fold for effect of age, which leads results to more closely mirror the original N = 100 ANCOVA results where both these variables account for only 1% of variation, thus suggesting a better model with covariates added. Pragmatic function's direct effect was halved with the addition of covariates added to analysis.

In summary, it appears that for N = 50 ANCOVA of the direct-opinion prompt set, frequency of [I + verb] usage (as determined by percentage of pragmatic [I + verb] construct out of total utterances) varies as a factor of the interaction between sociolinguistic background and pragmatic function (in accordance with previous results), with not much variation being a function of L2 proficiency or other confounding variables analyzed. Additionally, because the addition of covariates led results to more closely mirror results of the N = 100 ANCOVA, it appears that addition of these three covariates led to a better model fit of analysis. At this level, L2 proficiency does not appear to play a large role in accounting for total variance.