

From pitch to purpose: The prosodic-pragmatic mapping of [I + verb] belief constructions in English and Mandarin

Introduction

This paper explores how prosodic variation via focus-marking acts as a pragmatic tool to create conversational implicatures. A cross-linguistic study in Mandarin and English was conducted to investigate this prosodic-pragmatic link through a particular pragmatically ambiguous construct: [I + verb] preceding an opinion. Given that opinions inherently express speaker belief, the addition of [I + verb] belief constructs (i.e. “I think,” “I feel,” “I believe”) appears semantically redundant. Pragmatically, however, different prosodic variations of [I + verb] may perform either boosting (Main Clause) or hedging (Epistemic marker, Discourse marker) functions. Following an overview of the prosodic-pragmatic interface and alternative semantics theory, the cross-linguistic use of pitch as a cue for prominence¹ will be discussed, leading to a developed conceptual framework for the pragmatic-prosodic mapping of [I + verb].²

¹ Whilst there are various types of prominence, in the context of this paper we refer to pragmatic prominence via prosodic focus-marking (specifically, prosodic variation of focus placement).

²Definition of abbreviations:

CR = Confidence rating

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. The prosody-pragmatics interface

As pragmatic communication is considered a cognitive process, various theories of how this process functions have been posited (See Grice, 1975/1989; Sperber & Wilson, 1995). 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).

A speaker draws upon their 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 is explored in the next section.

Given 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 of the speech signal are grouped (syntactically parsed based on pragmatic interpretation) and given

prominence (Pierrehumbert, 1980). Wilson and 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.

Focus-marking by placing pragmatic prominence within a prosodic grouping of syllables allows speakers to pragmatically guide the inferred meaning of their utterance. Viewing pragmatic discourse as a cognitive tool, prosody is used to create conversational implicatures and direct speaker interpretation (Calhoun, 2007). 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 the perspective of information structure (See Steedman, 2000; Jackendoff, 1977/2002), 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).

1.1 Alternative semantics theory

Within information structure, focus-marking may be explicated via alternative semantics (see Rooth, 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, 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). For 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 veridical verb of "know" instead of a weaker veridical 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 focus-marking, 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 focus-marking as a pragmatic tool has been presented in past research, there appears to be little empirical evidence specifically regarding cross-linguistic similarity vs. variation of this pragmatic-prosodic link. Thus,

the current study was implemented to establish a 'proof of concept', i.e., an empirical link.

1.2 Cross-linguistic realizations of focus-marking

Research shows that prosodically, contrastive focus (i.e. "contextually contrastive", Wichmann et al., 2009:10) appears to be realized universally as a pragmatic tool for presenting new information. Indeed, there is congruent evidence that use of pitch (F0) to signify contrastive focus (hereafter referred to as "focus") is used by both American English speakers (e.g. Lehiste, 1970; Pierrehumbert & Hirschberg, 1990) and Mandarin speakers (e.g. Zou, Bao, & Luo, 2010)³. Whilst other prosodic cues such as duration and intensity may also be used to mark contrastive focus (e.g. Katz & Selkirk, 2011), F0 manipulation was chosen for investigation given a higher accuracy as a cue for focus-marking which is present in both Mandarin and English, with Mandarin speakers being better able to interpret English focus-marking using F0 (Wang, 2008), and a higher reliance on F0 in English to signify focus (Lehiste, 1970; Dehé & Wichmann, 2010). However, languages vary in their manner of manipulating F0 to indicate lexical identities (Zimmermann & Onea, 2011).

In an intonation, or stress, language like English, focus is marked by pitch accents. In English, this typically involves increased F0, with H*⁴, i.e. a local F0

³ Of course, non-prosodic cues for focus-marking exist in various languages, but for the purpose of this paper only prosodic cues will be explored.

⁴ According to Pierrehumbert & Hirschberg (1990:289), an H* pitch accent marks an item as "salient" and "new" in the discourse". A contrast may also be marked by L+H* (Pierrehumbert & Hirschberg, 1990: 296), i.e. a pitch accent also involving a high peak associated with the stressed syllable. For the purpose of this paper, we use H* here to refer to focus but do not mean to exclude L+H*.

maximum associated with the stressed syllable, being a typical focal accent (Pierrehumbert & Hirschberg, 1990). Contrastive focus within a syntactic grouping is marked with a pitch accent, which functions as the grammatical equivalent of focus in information structure (Zimmermann & Onea, 2011:1657). However, while a pitch accent involving a high peak is typically used for focus-marking in English (Pierrehumbert, 1980/1990; Zimmermann & Onea, 2011), utilization of F0 for focus-marking changes in a tonal language like Mandarin.

Mandarin employs four distinct tones for lexical distinction. Each lexical item is a single syllable with a tonal marking, which establishes a syllable's local pitch contour. Consequently, 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.

Mandarin focus may be therefore more accurately realized through variation of F0 range expansion, or raising/lowering of minF0/maxF0 of a syllable. Although different tones inherently possess varied F0 ranges, the largest F0 range within a prosodic grouping appears to indicate prosodic prominence (Jin, 1996; Chen & Gussenhoven, 2008). That is, focus is marked by maximum F0 range relative to other lexical items (Schack, 2000). Furthermore, post-focal F0 range compression in

Mandarin creates the same sort of downward F0 trend movement as post-focal deaccenting in English, and F0 deaccenting may even be more accurately characterized as F0 compression (Flemming, 2008). Thus, based on the similar behaviours of F0 accenting and F0 range expansion, prosodic focus-marking in Mandarin and English appear fundamentally comparable.

In sum, F0 manipulation may be used cross-linguistically in both Mandarin and English as a tool for focus-marking to direct pragmatic interpretation. English does so through a local F0 maximum, and Mandarin through F0 range expansion. Given a congruous cross-linguistic global intonation trend, F0 peaks in English and F0 range expansion in Mandarin may be treated similarly as prosodic cues for pragmatic inference. Thus, both these manipulations of F0 may be viewed as expressions of F0 range to signify focus-marking. This framework for prosodic variation will be applied and discussed further in the methodology (section 4). In what follows, we refer to F0 accent placement in English and F0 range expansion in Mandarin together as local pitch events in regard to prosodic focus-marking.

2. The prosodic-pragmatic mapping of [I + verb]

This study explores the above topics through a specific syntactically ambiguous construct: The [I + verb] belief construction (see Dehé & Wichmann, 2010 for overview). Recall that pragmatic knowledge is pragmatically derived from sociocultural environment (i.e. linguistic and cultural background); thus, it is possible that this could influence pragmatic prosodic mapping of [I + verb]. While such an 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 local pitch events (i.e. use of prosodic variation for focus-marking) 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.

Local pitch variation may lead to three prosodic forms of this construct with respect to prominence, each of which fulfills a different pragmatic function. 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 jude,' 'wo ganjue,' 'wo xiangxin') appears semantically unnecessary, especially

since the profile determiner of the statement is the indicative opinion clause (Langacker, 1987). 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" ('juede') may be the most flexibly used, other cognitive verbs, such as "guess," "believe," and "feel," and their roughly Mandarin counterparts "xiang" ('guess'), "xiangxin" ('believe') and "renwei" or "ganjue" ('feel') are also susceptible to polysemous interpretations (Guanwu, 2008; Huang, 2011). Based on these cross-linguistic similarities, 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 correspond to their translated Mandarin counterparts.

Syntactically, a matrix clause acts as the main clause of the utterance, with the embedded (complement) clause designated as a subordinate clause (Dehé & 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 parenthetical clause (Dehé & Wichmann, 2010).

This notion derives from Langacker's (1991:436) theory that only the main clause may act as the profile determiner of the utterance; that is, the main clause carries the most semantic value in the utterance. 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/her* 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], which may be classified into three pragmatic forms: main clause (MC), epistemic marker (EM), and discourse marker (DM) (Quirk et al.,

1985:1112). Corresponding work in Mandarin has also focused on the qualitative difference between *wo juede*'s ('I think') EM usage (Endo, 2013) and DM usage (Huang, 2011; Liu, 2013). Dehé 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. MCs appear the most assertive, committing the speaker directly to his or her statement (Dehé & 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 in both Mandarin and English as either a hedger of passive politeness or a formulaic filler for online-planning of speech (Guanwu, 2008:1702; Baumgarten & House, 2010; Aijmer, 1997).

Note that Dehé and Wichmann's (2010) research on sentence-initial [I + verb] starts out distinguishing only between MC vs. CC (comment clause) forms of [I + verb], where CC comprises both EMs functioning to express speaker attitude and DMs used for interpersonal/interactional purposes (i.e. encompassing two different [I

+ verb] pragmatic functions). However in their prosodic analysis and subsequent discussion, they identify three categories: MCs (prominence associated with pronoun), CCs (equivalent only to present EMs, prominence on verb) and DMs (unstressed; hedging, discoursal function). DMs used non-pragmatically as gap fillers were not included in their work (note that interpersonal/interactional DM use vs. gap filler DM are virtually indistinguishable in terms of prosody and syntactic position, and thus would be difficult to accurately identify all non-pragmatic gap filler DMs from pragmatic DM usage). In order to refine a more defined prosodic-pragmatic framework and to render more precisely the terminology, the umbrella term CC and pragmatic-only DM term were foregone in lieu of the separate, prosodically-distinguishable pragmatic classifications of EM (focus-marking on verb) and all-encompassing DM (no focus-marking). This distinctive cue of prosody to signify the pragmatic function of [I + verb] forms is discussed in the following section.

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

This section explores how different [I + verb] variations are inferred through prosodic cues. Since pragmatic tools create context-dependent implicatures, they are employed based on one-time situational use, making context a relevant cue for form in both Mandarin (Guanwu, 2008) and English (Sperber & Wilson, 1997). Regarding linguistic cues, Dehé and Wichmann's (2010) prosodic analysis of "I think" and "I believe" in British English found that prosody was the most reliable cue for determining pragmatic function, with Wang (2008) similarly determining F0 to be the most accurate tool for determining focus for Mandarin speakers. Table 3.2

accordingly revisits alternative semantic forms for prosodic variations in focus-marking of the same statement (see section 2.2).

Table 3.2: Focus placement⁵ to alternative semantics implicature for [I + verb] variations

Focus location	Example (italics = focus-marked)	Alternative form put forth by [I + verb] belief construction
Pronoun	<i>I</i> think cats are nice.	X think(s) cats are nice.
Verb	I <i>think</i> 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 no focus 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 value; because DMs by definition retain no semantic information, they receive no focus-marking, and thus no prosodic prominence via a local pitch event.

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

⁵ Here and in the tables below, focus placement/focus-marking involves a local pitch event, i.e. a pitch accent in English and pitch expansion in Mandarin, as explained above.

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

Focus location	[[focus semantic value]]F = {set of semantic objects}	Pragmatic function	[I + verb] Form
Pronoun	[[I]]F = {me vs. anyone else}	Focus-marking on speaker = increased commitment to statement	MC
Verb	[[think]]F = {think vs. know}	Focus-marking on uncertainty verb ⁶ = decreased commitment to statement	EM
None	N/A	No focus-marking = interpersonal marker of reluctance to present one's statement	DM

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

⁶ Note that in the case of “know”, being the one certainty verb of belief, focus-marking 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 veridical 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.

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

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 changes in word order during spontaneous speech, previous research suggests that prosodic cues seem to take precedence in the event of conflicting prosodic-syntactic cues (Dehé & Wichmann, 2010).

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" (Huang, 2011)	X	X	
p medial/final (Huang, 2011)			X
2+ times use per utterance (Guanwu, 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	DM

In summary, prosodic [I + verb] variations should express different pragmatic attitudes as per alternative semantics theory, with more deliberate (MC) usage expressing most confidence, uncertain (EM) usage expressing less confidence, and more reluctant (DM) usage expressing least confidence. Moreover, given the universality of focus-marking as a cue for prominence, this trend of confidence rating should manifest cross-linguistically. This scale of speaker stance and confidence was tested in the study described below.

4. Methodology

The potential cross-linguistic phenomenon of pragmatic-prosodic mapping of confidence rating for [I + verb] was investigated in this research for English and Mandarin. Utilizing the continuum of speaker stance scale introduced above, participants were asked to rate recorded audio utterances containing different prosodic-syntactic [I + verb] variations on confidence (rating from 1-7) and perceived pragmatic function (whether the utterance sounded deliberate, uncertain, or reluctant).

Given incongruous classifications and methodological shortcomings for classifying [I + verb] variations in both Mandarin and English, as well as no research having been previously conducted regarding the existence of a confidence-mapping trend of various pragmatic-prosodic [I + verb] forms, this Proof of Concept study was designed to identify the cognitive function and perceived confidence rating of each [I + verb] form through participants' interpretations of [I + verb] prosodic variation, thereby injecting initial insight into previous claims of [I + verb]'s pragmatic functions.

From past research in focus-marking and alternative semantics (see sections 1.1 – 1.2), the following hypotheses were formed:

H1) Prosodic variation and pragmatic attitude of [I + verb] statements are correlated in terms of confidence rating (CR). Specifically, CR should decrease from direct statement (DIR; no [I + verb]) to PRN to V to no focus conditions for

prosodic variation, and from Deliberate to Uncertain to Reluctant conditions for pragmatic attitude.

H2) There is no effect of sociolinguistic background⁷ on trend of CR of [I + verb] statements (That is, the same trend of CR occurs regardless of linguistic/cultural background).

The following interpretive, mixed-methods task was designed to investigate these hypotheses cross-linguistically and cross-culturally between Mandarin and English. In the subsequent sections, information regarding participants, materials, procedure, and analysis are presented.

4.1 Participants

Participants were fluent native speakers of American English and Mandarin who had resided most of their lives (<2 years in any other country) in the United States for American English speakers (US group) or China for Mandarin speakers (CHI group). Participants were recruited via snowball sampling with personal contacts in both countries, as well as through recruitment at universities via email adverts. 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 to thus allow for a representation of sociolinguistic background that more wholly encompasses the population

⁷ i.e. linguistic and cultural background. Whilst in this case sociolinguistic background may equate to L1 background, this is not the case for individuals who have either a) acquired more than one language, or b) lived in other sociolinguistic environments (e.g. immigrated).

(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 .80 observed power threshold for reliability (Cohen, 1988). 60 participants completed the task in total. However, one participant within the CHI group was excluded as answers were submitted before the minimum time required to listen to the full 2-3 second audio clip, thereby indicating improper performance of the task. Thus, final sample size for analysis was $n = 30$ in the US group and $n = 29$ in the CHI group.

Participants were aged 18 to 71 (34 males, 26 females). Mean age was 35.27 ($SE = 2.54$) for the US group and 44.21 ($SE = 2.67$) for the CHI group. 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 ($F(1,57) = .03$, $SE = 3.68$, $p = .02$), age was included as a covariate in analyses.

4.2 Materials

Materials used in this study are listed below.

4.2.1 Language questionnaire

An online-language questionnaire screened for participant-group inclusion based on language and sociolinguistic background (Appendix A). This questionnaire was adapted from Zhao's (2013) study on bilingual French-English vs. monolingual English children and appeared as the first page before the perception task survey. Participants were only permitted to proceed to the

English or Mandarin perception task if their answers indicated that they were a monolingual speaker of either English or Mandarin, respectively.

4.2.2 Perception task

An original online survey was created via Qualtrics (Qualtrics Lab, Inc, 2015) and validated through pilot testing. 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 the effect of prosodic variation on semantic interpretation.

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

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

Prosodic-syntactic form (based on focus placement)	Label for study's use	Associated pragmatic form	Example of construct (<i>focus in bold-italics</i>)
Direct statement (no [I + verb])	DIR	Direct	"It's a #."
Focus-marking on pronoun	PRN	Main clause	" <i>I</i> think it's a #."
Focus-marking on verb	V	Epistemic	"I <i>think</i> it's a #."

		marker	
No focus-marking ([I + verb] sentence-initial)	FRONT	Discourse marker	"I think it's a #."
No focus-marking ([I + verb] sentence-medial)	MID		"It's, I think, a #."
No focus-marking ([I + verb] sentence-final)	END		"It's a # I think."

Utterances containing different prosodic-syntactic [I + verb] variations of target nouns (see example utterances in Table 4.3) were audio-recorded from a native speaker with a "typical" accent in each language. The standard features of the speaker's accent were confirmed by an open-ended question posed at the end of the pilot phase of the study in both the Mandarin/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 and American English dialects existing in the US, responses from the pilot confirm overall standardness of accent.

"Think" ('juede' in Mandarin), the most commonly used belief verb, was used across trials. All target nouns (see Appendix D) were imageable, 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 was based on the structure of Meyer and Mleineik’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 (1 = lowest and 7 = highest confidence rating). This question is derived from Kurumada, Brown and Tanenhaus' (2012) confidence rating task, which asked participants to rate self-confidence in regard to their answer choice.

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 (Wilson & Wharton, 2006:1559), 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 (Boersma & Weenik, 2016), to ensure that all recordings consistently followed the specific pitch-contours of each [I + verb] variation (see Appendix B for examples).

Questions and format in each language, as well as a screenshot from an example prompt, may be found in Appendix C. This task included 24 target nouns with each of the six prosodic variations appearing four times. The condition distribution for prosodic variation is provided in Appendix D.

4.3 Procedure

Participants were recruited via verbal/email enquiry through academic contacts at American and Chinese universities⁸, and further on through snowball sampling. Qualtrics IP address tracking of participant submissions ensured that all participants completing this online task resided in the appropriate country.

Before data collection, participants received written explanation of the study and signed an electronic consent form. Next, participants filled out an online language questionnaire (see section 4.2.1). If responses indicated appropriate linguistic and cultural background for group inclusion, participants were immediately directed to the online perception task in their native English/Mandarin.

Participants were provided with written instructions and then presented with 24 consecutive trials administered on separate pages (Appendix C).

⁸ Universities fitting these criteria included Washington University and University of Michigan in America, and universities in international cities like Beijing, Hangzhou and Nanjing in China.

5. Results

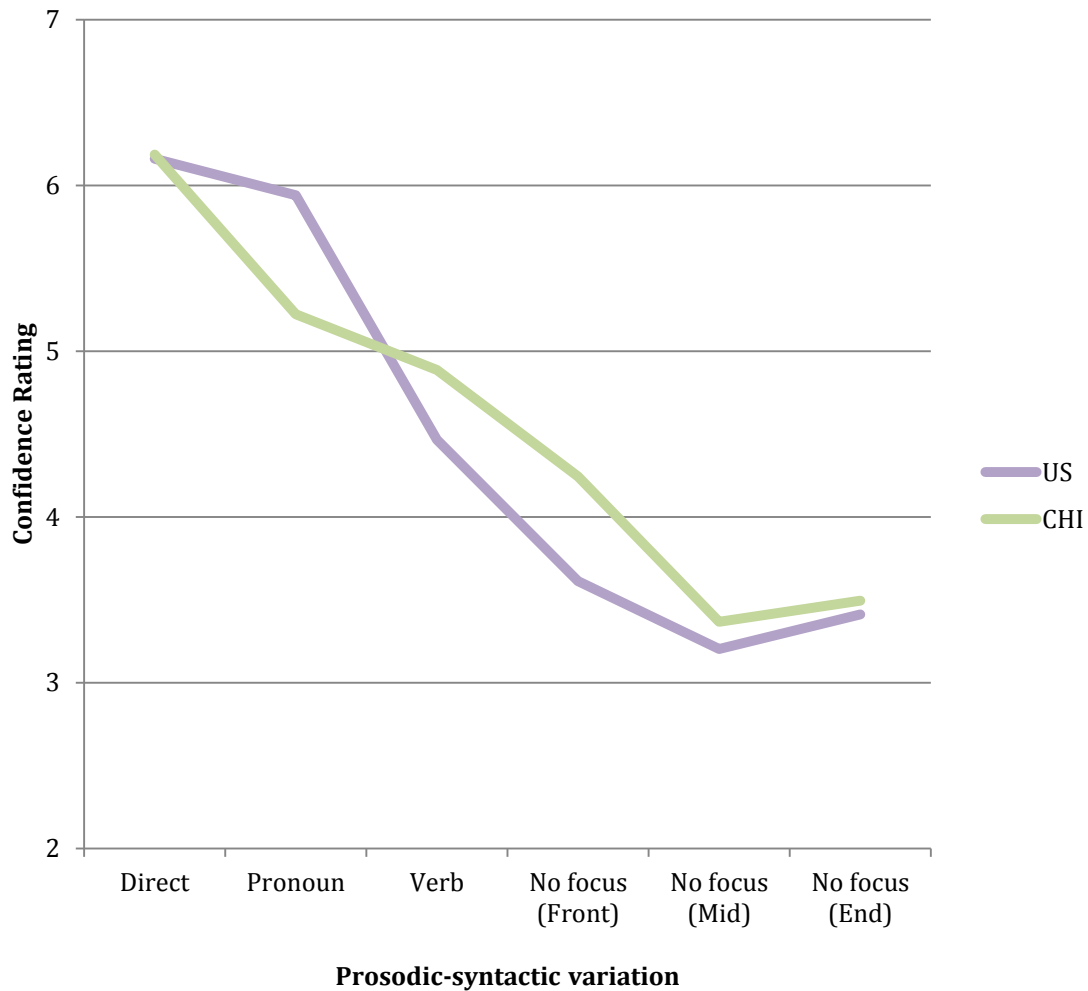
Findings are first summarized in an overview of results section, and then followed by in-depth, detailed statistics. Raw data is accessible through Mendeley Data ([Authors], 2017).

5.1 Overview of results

Quantitative analysis for the perception task consisted of three analyses each containing a 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 is also presented in Appendix E. Age was analyzed as a covariate due to a significant mean difference in age between groups (see section .1). All confidence rating (CR) values were compared within-group and between-group. Additionally, a one-way ANCOVA examination of response timing ensured that there was no mean difference in length of responses.

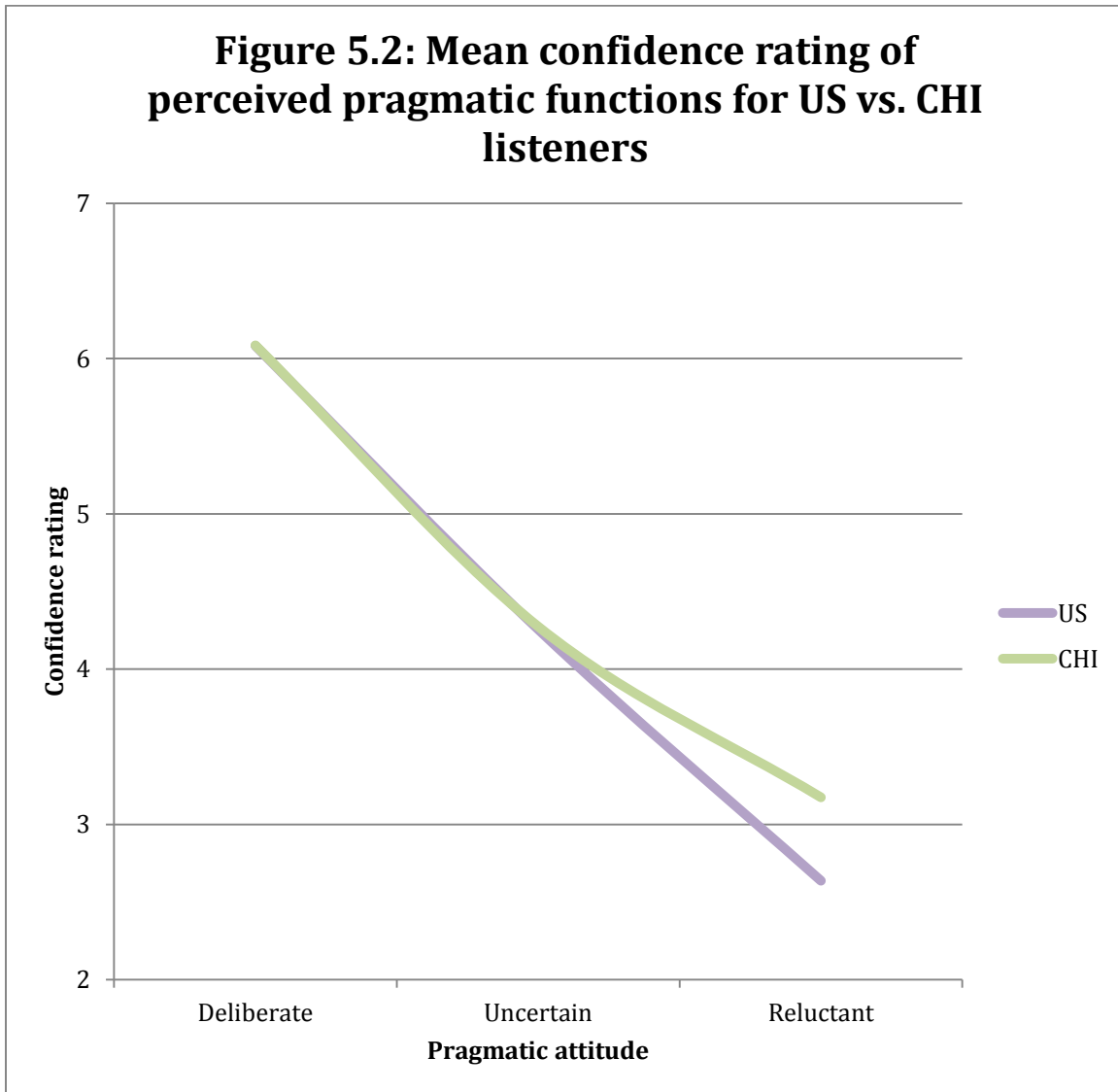
We first present here a brief overview of the main findings followed by detailed analyses in section 5.2. Results from Analysis 1 indicated that pragmatic function had a significant effect on CR, thereby supporting Hypothesis 1's prediction of a pragmatic-prosodic link between CR and different prosodic-syntactic [I + verb] variations based on pragmatic interpretation. As illustrated in Figure 5.1, CR was highest for DIR variation, then second highest for PRN, then followed by V, with no focus-marking variations possessing the lowest CR scores; this same trend of CR appeared between both US and CHI groups.

Figure 5.1: Mean confidence rating of prosodic-syntactic [I + verb] variations for US vs. CHI listeners



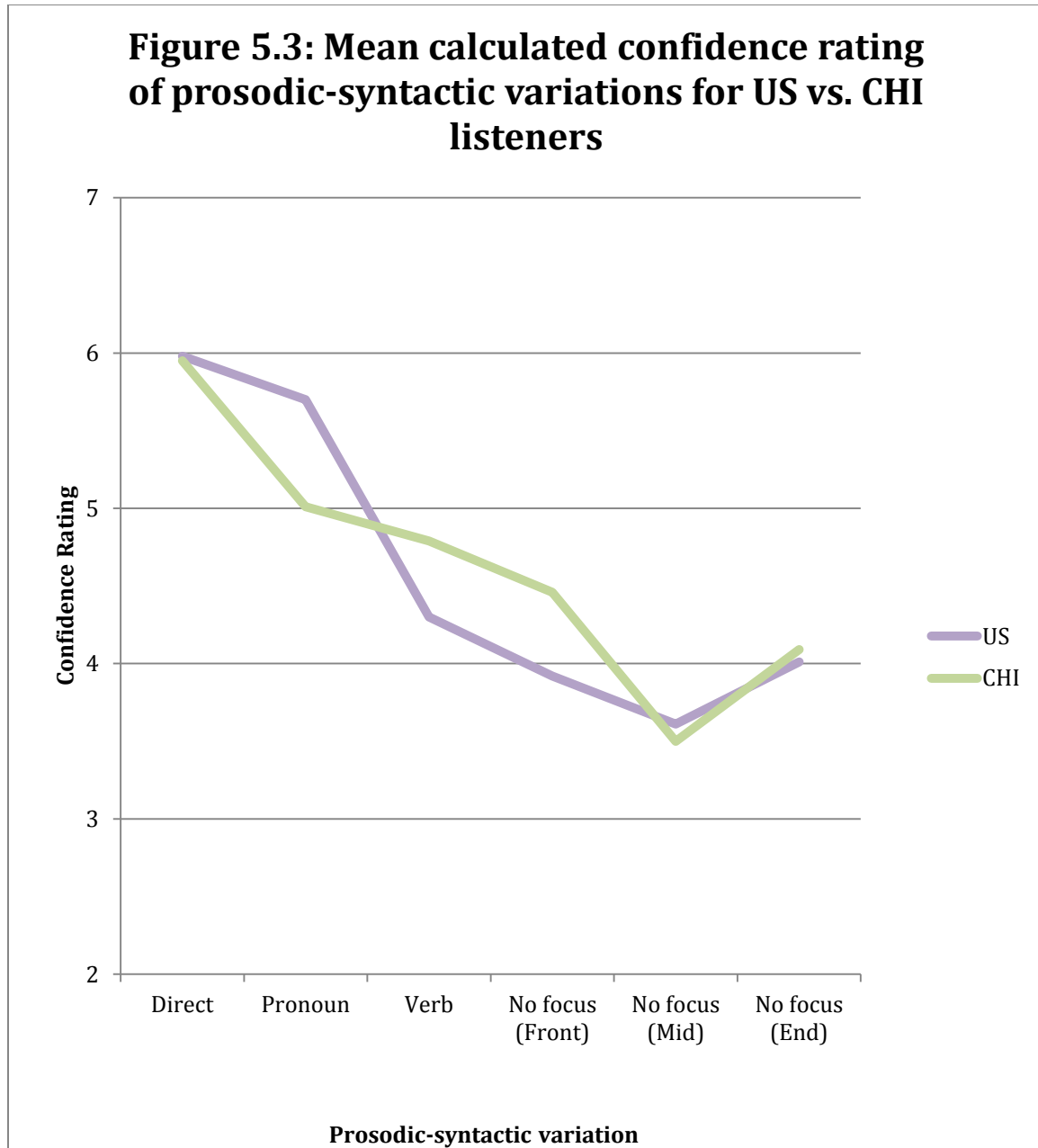
Additionally, Analysis 2's examination on CR of perceived pragmatic function found a significant effect of perceived pragmatic function on CR of [I + verb] and thus also align with Hypothesis 1, with highest CR for Deliberate second highest

CR for Uncertain variations, and lowest CR for Reluctant variations across both groups (see Figure 5.2).



Furthermore, a third analysis establishing a CR correlation between pragmatic attitude and prosodic variation via mean CR for each participant mirrored findings from Analysis 1 by confirming a significant within-subject effect

of prosodic variation on CR. Examining Figure 5.3, it is clear that the same shape and trend of CR appears between Analysis 1 and Analysis 3, thereby further strengthening validity regarding a universal pragmatic-prosodic mapping of [I + verb].



Finally, because all three analyses found no significant effect of sociolinguistic background on CR performance, results strongly support Hypothesis 2's prediction of the same CR trend of [I + verb] across languages.

In sum, results coincide with the notion of a cross-linguistic pragmatic-prosodic mapping of [I + verb]: It appears that prosodic focus-marking, and the alternative semantic connotations plus modality it invokes, may determine perceived confidence of [I + verb] statements across typologically different languages. The following section describes these findings in statistical detail.

5.2 In-depth statistics

Each of the three analyses is now presented in detail below.

5.2.1 Confidence Rating of [I + verb] prosodic variations

The first ANCOVA investigated whether prosodic variation and pragmatic attitude of [I + verb] statements are correlated in terms of CR, and whether sociolinguistic background influences the CR of [I + verb] statements. A mean CR value⁹ 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

⁹ Whilst the task consisted of Likert items, the dependent variable of confidence rating was calculated using each participant's mean score of their repeated measures ratings for each [I + verb] variation, thereby being analysed as scalar data and benefitting from this parametric analysis' robustness and ability to handle multiple dependent variables.

conceptual framework (see section 2.2) as well as between-groups. Table 5.1 presents the mean CRs of prosodic variations by group.

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

Focus placement	Socio-linguistic background	Mean	SD
Direct (DIR)	US	6.16	.73
	CHI	6.19	.87
	Total	6.17	.80
Pronoun (PRN)	US	5.94	.87
	CHI	5.22	.89
	Total	5.59	.96
Verb (V)	US	4.47	.82
	CHI	4.89	.83
	Total	4.67	.86
No focus-marking: Sentence initial (FRONT)	US	3.61	1.12
	CHI	4.25	1.38
	Total	3.93	1.30
No focus-marking: Within-sentence (MID)	US	3.20	1.07
	CHI	3.37	1.10
	Total	3.29	1.09
No focus-marking: Sentence final (END)	US	3.41	.82
	CHI	3.50	.96
	Total	3.45	.89

Examination of histograms and Shapiro-Wilk testing suggested normality of data distribution. Whilst sphericity cannot necessarily be assumed ($X^2(1,14) = 25.43, p = .03$), results calculated with Mauchly's sphericity vs. Huynh-Feldt measurements (.96) did not differ regarding significance. Furthermore, 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 ($F(1,4.81) = 24.74, p < .001, \eta_p^2 = .31$). Conversely, no between-subject effect of sociolinguistic background ($F = .32, p = .58, \eta_p^2 = .01$) or covariate of age ($F = .71, p = .40, \eta_p^2 = .01$) was found. Additionally, age did not interact significantly with prosodic variation ($p = .107$). There was, however, a moderate interaction between sociolinguistic background and prosodic variation ($F(1,4.81) = 4.76, p < .001, \eta_p^2 = .08$). To examine this effect, post-hoc paired t-tests were run with a Bonferroni adjustment for multiple comparisons. Results confirmed 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 this study predicts that CR differs only for variation in focus placement; thus, given that MID and END are both no focus conditions which differ only in syntactic placement, this result does not refute the hypothesis. However, FRONT, also a no focus 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, 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) is highly consistent between US and CHI groups, with CR trending from high-to-low as levels significantly shift from Direct to Pronoun, to verb, to no focus-marking conditions.

Because there was no main effect of sociolinguistic background, results support the notion of focus-marking as a pragmatic tool, as well as supporting the predicted cross-linguistic congruence of pragmatic-prosodic mappings¹⁰.

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 ($p = .58$), suggesting a cross-linguistic pragmatic-prosodic mapping between Mandarin and American English.

¹⁰ Note that analysis regarding the effect of sociolinguistic background only relates to Proof of Concept (i.e. a widespread phenomenon 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.

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

The second ANCOVA investigated whether pragmatic attitude and perceived pragmatic function of [I + verb] statements are correlated in terms of CR, and whether sociolinguistic background affects trend of CR of [I + verb] statements. A 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.

Although only 6% of data was incomplete for the main data set, 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, multiple imputation (MI) was implemented for the repeated measures analysis.

Modern approaches tailor the number of imputations to be equal to at least the percentage of missing data cases (White, Royston & Wood, 2011). Thus, given 19% missing cases, $M = 20$ was used with a random number generator set at 2,000,000 for the Mersenne Twister.

Examination of histograms and Shapiro-Wilk testing suggested normality of data distribution. 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 5.4¹¹. 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

¹¹ 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 ($F(2,106.46) = 307.58, p < .001, \eta_p^2 = .85$). 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-significant¹². 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. Results revealed significant differences between all levels of perceived pragmatic function ($p < .001$). As predicted, a visible difference of means appeared between each level with highest CR for Deliberate, second highest for Uncertain, and lowest for Reluctant. For Deliberate and Uncertain levels, US and CHI mean CRs are nearly identical. While US participants rated 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 Hypothesis 1 and indicating a universal perception of high confidence for deliberate statements, medium confidence for uncertain statements, and low confidence for reluctant statements.

Similar to results of the first analysis, 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

¹² 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.

groups, which furthermore supports Hypothesis 2's prediction of a cross-linguistic similarity in confidence trend of [I + verb] pragmatic attitudes.

5.2.3 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. 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). A one-tailed Spearman rank-order correlation test was also conducted to further establish the link between prosodic variation and pragmatic attitude.

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. 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.

Levene's test indicated homogeneity of variance at all conditions ($p > .05$) except V ($p < .01$) and P ($p = .04$). 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. Shapiro-Wilk testing suggested potential non-parametric distribution within 8 out of 12 levels between both groups.

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 identify 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

Focus placement	Sociolinguistic background	Mean	SD
Direct (DIR)	US	5.98	.31
	CHI	5.95	.39
	Total	5.97	.35
Pronoun (PRN)	US	5.70	.57

	CHI	5.01	.81
	Total	5.35	.81
Verb (V)	US	4.30	.42
	CHI	4.80	.78
	Total	4.55	.67
No focus-marking: Sentence initial (FRONT)	US	3.92	.72
	CHI	4.46	.85
	Total	4.19	.83
No focus-marking: Within-sentence (MID)	US	3.61	.64
	CHI	3.50	.50
	Total	3.56	.57
No focus-marking: Sentence final (END)	US	4.01	.73
	CHI	4.09	.72
	Total	4.05	.72

Results appeared to exhibit the same trend of CR between prosodic variation as in the first analysis: descriptively, CR shifts from high-to-low from DIR to PRN to V, with the no focus-marking variations at the bottom. This trend will be discussed more thoroughly after presenting ANCOVA results below.

ANCOVA 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 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 focus-marking conditions vs. the FRONT and END no focus-marking conditions; however, given that the hypothesis only predicts a CR trend from DIR down to PRN to V to no focus-marking conditions, what appears meaningful is that all three no focus-marking 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 focus-marking 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 (ANCOVA 1) and pragmatic variation with pragmatic attitude-calculated CR ($p < .01$ for DIR, $p < .001$ for rest).

In summary, results from all analyses appear to substantiate hypotheses (see table 5.9).

Table 5.9: Summary of variable effects from each analysis

Analysis	Significant variables	Not significant variables
ANCOVA 1	Prosodic variation	Sociolinguistic background
ANCOVA 2	Perceived pragmatic function	
ANCOVA 3	Calculated CR	

*Significant variables were the same across both US and CHI groups

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 focus-marking 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 served both to establish a pragmatic-prosodic link as well as to strengthen the existence of a cross-linguistic trend of CR in syntactic-prosodic variations of [I + verb] constructs.

6. Discussion

This perception task served to investigate the existence of a cross-linguistic pragmatic-prosodic mapping for [I + verb] belief constructions by examining individuals' confidence ratings and perceived pragmatic function ratings of different [I+ verb] variations. Implications of our findings are discussed below.

Prosodic variation as a pragmatic tool

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, thereby establishing validity for the CR scale presented in Section 3.1. 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, findings suggest a significant link between pragmatic attitude and prosodic variation, wherein the same specific confidence rating trend decreasing from for DIR to PRN to V to no focus-marking (DMF, DMM, DME) conditions was

found in both Mandarin and English. Furthermore, post-hoc paired t-testing from all three analyses showed significant mean differences between DIR to PRN to V to no focus-marking conditions thereby confirming a CR trend which varies based on pragmatic-prosodic mapping of [I + verb] variations.

Whilst mean differences between no focus-marking 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: Hypothesis 1 predicts a trend of high-to-low CR from DIR to PRN to V to no focus-marking conditions, on the basis of the same prosodic variation for all no focus-marking conditions (with variation being purely syntactic). Thus, results from both pilot and main phase testing support predictions, establishing a proof of the concept regarding a specific pragmatic-prosodic link for perceived confidence of [I + verb] belief constructs.

These results therefore reinforce the notion that focus-marking does create systematic and consistent conversational implicatures; such findings align with theories of alternative semantics as presented by Rooth (1992/1996) (refer to section 2.1), 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). Thus, different [I + verb] variations possess semantic contrasts created through variation of prosodic focus-marking in [I+ verb], coinciding with Dehé 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 focus-

marking 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 focus-marking placement (equating to “Epistemic marker” [I + verb] constructs), 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 focus-marking of [I + verb] (equating to “Discourse marker” [I + verb] constructs), indicating less semantic value than the focus-marked variations above (Dehé & 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 prosodic focus-marking via local pitch events, with the consistent CR trend in pragmatic-prosodic mapping of [I + verb] decreasing from DIR to PRN to V to no focus conditions, thereby supporting predictions in line with past [I + verb] research and alternative semantics theory.

Cross-linguistic similarity of CR trends

As predicted, 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 focus-marking 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 focus-marking 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 Hypothesis 2's predicted cross-linguistic congruence of pragmatic-prosodic mappings. These findings across two such typologically different languages thereby reinforce the notion of focus-marking as a cross-linguistic cue for prominence as expressed by Zimmermann and Onea (2011) and Rooth (1985/1992) (see section 2.3).

7. Conclusion and Implications

The [I + verb] perception task discussed in this paper revealed that perceived confidence of [I + verb] constructs corresponds to alternative semantic interpretations created by prosodic variation of focus-marking, with direct statements appearing most confident, followed by Main clause, Epistemic marker, and finally Discourse marker [I + verb] variations.

Given the typological difference between Mandarin and English, this study's results are a promising indicator of a widespread cross-linguistic similarity for [I + verb] mapping, especially given consistency with theories of alternative semantics (see section 1.3) and focus-marking (see section 1.2). However, it must be noted that this congruent trend in form-function interpretation of [I + verb] in Mandarin/English does not essentially necessitate universality of pragmatic-prosodic mapping. Future research may choose to include a third typologically different language from English and Mandarin to examine whether the same trend of confidence rating is found, 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). 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.

Given the notion of prosodic focus-marking as a cue for pragmatic prominence, combined with fruitful results of cross-linguistic similarity from such typologically

diverse languages in this study, research can be optimistic that other languages should also possess a congruent pragmatic-prosodic mapping to that of English and Mandarin.

Additionally, whilst due to pragmatic constraints this study was limited in its use of only the most common cognitive verb, “think,” in the [I + verb] perception task, there should be no reason that using another cognitive belief verb (e.g. feel, believe, suppose) should elicit 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.).

Given less pragmatic constraints, one may also consider including an investigation of whether pro-drop within languages like Mandarin affects this mapping. This can be implemented by adding another [I + verb] variation with no pronoun for EM and DM prosodic-syntactic variations (since DIR condition contains no pronoun and the pronoun must be present for focus-marking in MC condition). We would thereby posit that this would not affect the CR trend, as pro-drop does not affect focus-marking. Nevertheless, whilst increasing the amount of conditions via multiple verbs or pro-drop variations would require a larger and therefore longer task, cross-linguistic mapping would be further substantiated if the CR trend remained consistent with these inclusions.

Alternatively, future research could also choose to conduct the perception task using the strong veridical 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.

With no previous research having conducted such an investigation, results of this study provide the proof of concept evidence for a cross-linguistic similarity in pragmatic-prosodic mapping of [I + verb]. This finding offers a novel contribution toward better understanding the pragmatic-prosodic interface: Whilst previous research on [I + verb] has focused on theories regarding how prosodic prominence may pragmatically affect the meaning behind [I + verb] variations (e.g. Dehé & Wichmann, 2010; Wichmann, Dehé, & 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 claims regarding assertiveness of various [I + verb] conditions created through focus-marking, and the speaker attitude they invoke regarding confidence, strengthen the notion of a widespread phenomenon of prosodic focus-marking as a cue for pragmatic prominence. Furthermore, given the different natures of focus marking in English vs. Mandarin, these findings also serve to strengthen the notion that pitch and pitch range both function successfully as a cue of prominence

This study further benefits future research by providing calculated notions of confidence rating for each [I + verb] variation, which no previous studies have as of yet produced. Utilizing this cross-linguistic confidence rating scale for [I + verb], future studies now have the ability to investigate whether speaker confidence as expressed via frequency of [I + verb] usage differs across cultures or contexts, in a manner that explores whether sociolinguistic groups behave differently when expressing opinion in various conversational or social situations.

8. Appendices

Appendix A: Language questionnaires

[US group:]

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? _____

[CHI group]

姓名: _____

出生日期: 年_____月____日____

这份问卷是无记名试卷。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。

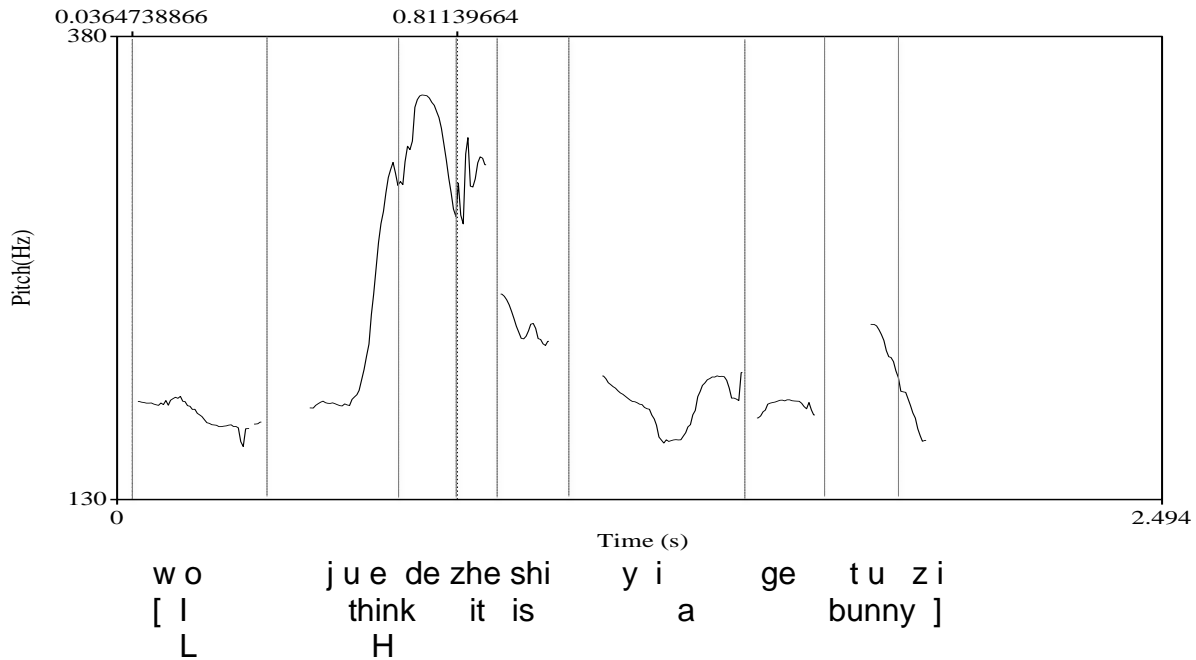
1. 你能同时说流利的中文和吗? (例如: 可以进行日常对话)

☐ 是 ☐ 不是 (不可以参加)

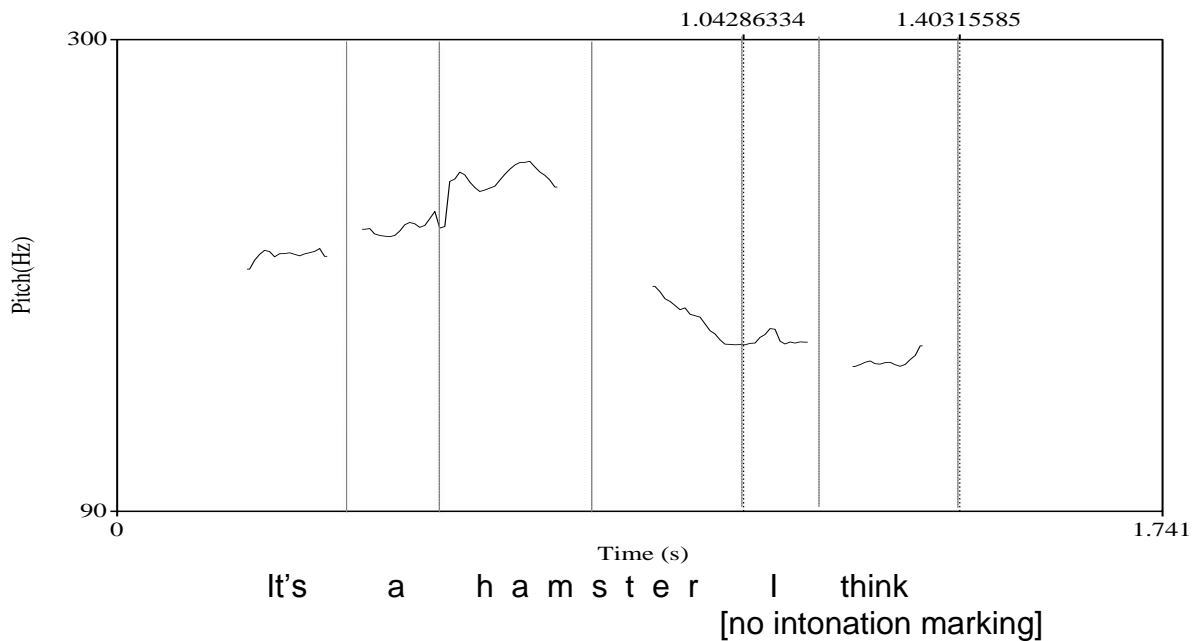
2. 你是不是在中国以外的国家和地区居住过 (不包括去国外度假和旅游)? 如果有, 请列出时间, 地点, 停留天数: _____

Appendix B: Examples of PRAAT analysis of prosodic-syntactic [I + verb] variations, separated by syllable

Example 1: Focus-marking on the verb, sentence-initial position = V condition.



Example 2: No focus-marking, sentence-final position = DME condition.



Appendix C: Perception task example trial page

[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)

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
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)

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	2	3	4	5	6	7
最低分			中等分			最高分

[Screenshot of example prompt:]

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
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Appendix D: Perception task condition distribution

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

Appendix E: Non-parametric Kruskal-Wallis test for 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

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.
- 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.
- Boersma, P. & Weenink, D. (2016). Praat: Doing phonetics by computer. [Computer program]. Version 6.0.21.
- 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.
- 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.
- Chen, Y. & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36(2008), 724-746.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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.
- Dehé, 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.
- Endo, T.K. (2010). University of California, Los Angeles, ProQuest Dissertations Publishing.
- 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.
- 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.
- 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.

- Guanwu, F. (2008). Pragmatic markers in Chinese. *Journal of Pragmatics*, 40, 1687-1718.
- 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.
- Jin, S. (1996) *An Acoustic Study of Sentence Stress in Mandarin Chinese*. Ph.D. dissertation, Columbus, Ohio: Ohio State University.
- Kaltenböck, G. (2009). *Pragmatic functions of parenthetical I think*. In Paper Presented at the 11th International Pragmatics Conference, Melbourne.
- Kasper, G. & Rose, K. (1999). Pragmatics and SLA. *Annual Review of Applied Linguistics*, 19, 81-104.
- Katz, J. & Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 87(4), 771-816.
- 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.
- 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.

- 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.
- Nikula, T. (1996). *Pragmatic Force Modifiers: A Study in Interlanguage Pragmatics*. Jyvaskyla: University of Jyvaskyla.
- 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.
- Qualtrics Company, Inc (2015). Qualtrics software, Version June/July 2015 of Qualtrics. Provo, UT, USA. <http://www.qualtrics.com>
- 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.
- 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.
- Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition*. 2nd edition. Oxford: Blackwell.

- 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.
- Van Ginkel, J.R. & Kroonenberg, P.M. (2014). Analysis of Variance of Multiply Imputed Data. *Multivariate Behavioral Research*, 49(1), 78-91.
- Wang, V. (2011). *Making requests by Chinese EFL Learners*. Amsterdam: John Benjamins Publishing.
- 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.
- Zhao, L. (2013). *The effect of bilingualism on young children's metaphor comprehension*. Masters Thesis, London: University College London.
- 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*.

[dataset] [Authors] (2017). Data for: From pitch to purpose: The prosodic-pragmatic mapping of [I + verb] belief constructions in English and Mandarin. *Mendeley Data*.