



'I think' is what I mean: Prosody as a signifier of speaker attitude across cultures and communication contexts

Lucy Kind^{1,2}, Victoria A. Murphy²

¹University of Oxford, Oxford, UK. ²Google, San Francisco, USA.
kind@google.com, victoria.murphy@education.ox.ac.uk

Abstract

[I + verb] belief constructs before expressions of opinion may appear superfluous, as opinions carry inherent notions of speaker belief. However, past research suggests that [I + verb] forms may actually fulfill various pragmatic functions based on prosodic variation, which in turn empirically correlate to different levels of speaker confidence [1]. This paper examines prosody as a pragmatic cue for speaker confidence of US vs Chinese speakers in communication contexts of personal belief vs judgment expressions. Production task results support predictions that frequency of functional [I + verb] usage corresponded to culturally specific attitudes of each group: Based on confidence rating values for each [I + verb] variation calculated from a previous proof-of-concept perception task, Native US individuals were more confident in expressing self-opinions but less confident in expressing opinions of others, whilst Native Chinese individuals were more confident in expressing opinions of others and less confident in expressing self-opinion. These results correspond to theories of identity and speaker style in US vs. Chinese cultures. Through investigating pragmatic-prosodic mappings of [I + verb] forms vs. functions, this paper demonstrates both the use of prosody as a tool for pragmatic communication, as well as the effect of culture on speaker attitude.

Index Terms: prosodic-pragmatic interface, speaker style and communication context, cross-cultural pragmatics

1. Introduction

Two widely discussed themes of cultural patterns are Hofstede's model of cultural dimensions [2, 3] and Markus and Kitayama's theory of self-construal [4]. Specifically, more recent empirical research has focused on the comparison between US and Chinese cultures in terms of independence and interdependence, and how this relates to identity [5].

For US culture, the individual is seen as the primary unit of society, therefore placing greatest value on the unique, singular self (Hofstede, 1980; Triandis, 1991); this appears to result in a high degree of modality (commitment) to express self-assuredness, causing a speaker to appear assertive or direct [4].

While US culture emphasizes the individual, Chinese culture may place more significance on the group or society [2], which may cause individuals to define the self relationally [6, 4, 7]. Such a high dependence on listener judgment may lead to a wariness to commit to an opinion [8, 9]. Consequently, Mandarin speakers may use indirectness or low modality to distance themselves from their statements [10].

These two cultures may therefore vary in speaker style depending on communication context, which may be examined through the lens of prosody as a pragmatic cue.

2. Prosodic-pragmatic [I+verb] framework

Zhao, Dehe, and Murphy (2017) [1] developed a conceptual framework for examining prosodic-pragmatic [I + verb] variations. This framework draws upon focus-marking and alternative semantics theory. Each [I + verb] form generates a conversational implicature and resulting speaker attitude that is consistent across Mandarin and English. These forms, described below, vary based on focus-marking placement (prosody) and syntax, which is used as a cue of prominence to create certain conversational implicatures.

The discourse marker (DM) form, with no focus-marking, expresses speaker reluctance and appears least confident (labeled as "Reluctant"). The epistemic marker (EM) form, with verbal focus-marking, appears more confident than the DM form, but still expresses uncertainty (labeled as "Uncertain"). The Main clause (MC) form, with pronoun focus-marking, expresses deliberate speaker stance and thus appears even more confident (labeled as "Deliberate"). In addition, simply stating an utterance without a [I + verb] presupposition appears most confident in its brevity and directness (labeled "Deliberate").

In summary, prosodic [I + verb] variations express different pragmatic attitudes, which are linked to different levels of speaker confidence. However, even with focus-marking as an apparent universal cue for prominence, the expectations of how these pragmatic [I + verb] forms are used across different communicative contexts may differ cross-culturally. Using this prosodic-pragmatic framework, the next section outlines an investigation of speaker attitude across cultures and communicative frequency of different pragmatic [I + verb] forms.

3. Methodology

A mixed-method, open-ended production task examined the influence of cultural background on communication style across two communicative contexts. These contexts are expressions of self-opinion and expressions of opinions about others, being the two cognitive processes of primary focus in most cross-cultural research on self-construal [7].

3.1. Participants

The sample size was 50 individuals. This comprises 25 native monolingual English speakers in the US and 25 native monolingual Chinese speakers in China. Monolingual status was determined by an online language-screening questionnaire and was defined as not being conversationally proficient in any other language.

Participants were aged 18 to 71. While a one-way ANOVA did not show a statistically significant main effect of group for age of participants ($F(1,3) = 7.37, p < .001$), age was run as a covariate for robustness. A one-way ANOVA showed no significant gender effect ($p = .71$).

3.2. Materials

An oral discourse completion task (DCT)[11] was created and validated through pilot testing to ensure successful elicitation of all variations of [I + verb] constructs.

This task had an English and Mandarin language version. Each version 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; 3 sets of 3 questions per scenario) to elicit responses of social judgment. For each set of Opinion of others prompts, a story is provided in which a fictional character made a mistake, although the nature of the mistake is purposefully vague to allow for interpretation. The “other” whom the participant is to appraise is assumed as a stranger and therefore an out-group member, designating high social distance between the character and the participant [12].

Prompts specifically do not include any of the verbs included in [I + verb] constructs (e.g. think, feel, believe, guess) to avoid priming. While prompts all express the sentiment of requesting the participant’s opinion, they are worded in various ways to avoid being repetitive. Prompts may be found at <https://osf.io/z8dq>.

3.3. Procedure

Participants were provided with a written explanation of the study in their respective languages and gave written consent before taking online language questionnaires. All participants were administered the task by the lead researcher in a quiet, undisturbed location, including empty classrooms in university buildings, quiet café settings, and home sitting rooms.

Participants were instructed to respond as if having a normal conversation and to continue talking until stopped by the researcher. They were informed that any inquiry may only be answered with “However you wish to interpret it.” 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 misheard, participants could request to rehear it once.

Order of prompts within each set was randomized and order of sets counterbalanced. Prompts were recorded on a Macbook Pro via Praat at 96000 hz.

3.4. Coding scheme for [I + verb] constructs

As per the conceptual framework, [I + verb] constructs were coded under 5 categories: MC, EM, and three DM categories (DMF for syntactic-front/initial position, DMM for medial position, and DME for end/final position), plus one for direct opinion utterances (DIR) [1]. Note that Zhao, Dehé, and Murphy’s previous proof-of-concept task shows that pitch accent in English equivalent to maximum pitch range in Mandarin to signify focus marking, with perception task results providing evidence that this coding system functions for both languages (For examples of conceptual framework prosodic analysis, see their Appendix 2)[1].

Table 1 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.

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 Man-

Table 1: [I + verb] construct intonation markings

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

darin [13, 1].

Interrating and intrarating agreement, done by 3 native speakers for English and Mandarin respectively, were calculated with Cohen’s kappa and conducted on 10% of total constructs. The raters completed a pilot before conducting the ratings calculated below.

For English coding, interrater reliability was 89% (std. error = .02, $p < .001$), and intrarater reliability was 96% (std. error = .01, $p < .001$). For Chinese coding, interrater reliability was 90% (std. error = .04, $p < .001$), and intrarater reliability was 94% (std. error = .03, $p < .001$).

4. Results

Frequency counts of each [I + verb] construct variation were tabulated from transcriptions of audio recordings using AntConc 2 software for English transcriptions and AntPConc 2 software for Mandarin transcriptions [14]. Prosodic and syntactic analyses were then performed to determine specific [I + verb] variation, as per the previous section.

The production task was quantitatively analyzed for frequency of pragmatic forms of [I + verb] usage, as validated in Zhao, Dehe, and Murphy’s (2017) study. To reiterate, based on the pragmatic-prosodic mapping for [I + verb], Direct (DIR) and Main clause (MC) were considered as ‘Deliberate’, Epistemic marker (EM) as ‘Uncertain,’ and no Discourse marker (DM) at sentence-front (DMF), mid (DMM) and end (DME) positions as ‘Reluctant’.

A one-way ANOVA on number of utterances per response showed no significant mean differences between groups for either prompt set, indicating that length of audio recording should not affect results.

4.1. Percentage pragmatic function of [I + verb] to total utterances produced

Results are presented below for each prompt set.

4.1.1. Self-opinion prompt set

Means for percentage of pragmatic function to total utterances are presented in Table 2.

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 two levels (Native US, Native CHI), with age as a covariate, and with a dependent variable of percentage of pragmatic function of [I + verb] to total utterances produced (being calculated akin to Lin, Ren and Woodfield’s (2012) DCT study regarding frequency percentages of pragmatic speech acts).

ANCOVA results (see Figure 1) revealed a between-group effect of sociolinguistic background ($F(1,3) = 22.26, p < .001, \eta_p^2 = .41$) and within-group effect of pragmatic function ($F(1,2)$

Table 2: Means and SDs for percentage of pragmatic function to total utterances

Pragmatic function	Group	Self-opinion: Mean (SD)	Opinion-of-other: Mean (SD)
Deliberate	Native US	25.27 (9.93)	5.36 (3.06)
	Native CHI	3.43 (2.17)	27.23 (8.23)
Uncertain	Native US	20.47 (10.25)	11.56 (8.65)
	Native CHI	5.74 (4.05)	5.28 (4.72)
Reluctant	Native US	4.47 (3.74)	7.00 (5.51)
	Native CHI	15.32 (9.68)	4.88 (3.80)

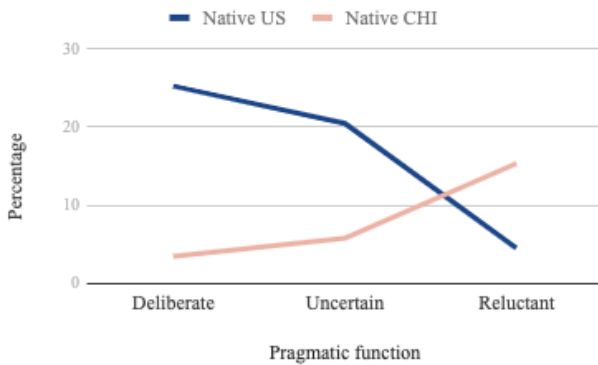


Figure 1: Mean % pragmatic function to total utterances for Self-opinion prompts.

= 3.10, $p = .05$, $\eta_p^2 = .03$). Moreover, results showed a large interactional effect between pragmatic function and sociolinguistic background ($F(1,6) = 30.80$, $p < .001$, $\eta_p^2 = .49$). There was not 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 ($F(1,2) = 2.67$, $p = .07$, $\eta_p^2 = .03$).

Post-hoc paired t-tests on pragmatic function conditions revealed significant between-group differences. Deliberate and Uncertain usage were significantly higher for the Native US group than the Native CHI group. For Reluctant condition, frequency of [I + verb] was significantly higher for Native CHI group than the Native US group.

4.1.2. Opinion of others prompt set

Means for percentage of pragmatic function to total utterances are presented in Table 2.

Results (see Figure 2) revealed a small within-subject effect of pragmatic function ($F(1,2) = 7.92$, $p < .01$, $\eta_p^2 = .08$) and 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 ($F(1,6) = 24.93$, $p < .001$, $\eta_p^2 = .44$).

Post-hoc paired t-testing was conducted between-groups on pragmatic function. Deliberate [I + verb] usage was significantly lower for Native US vs. Native CHI groups, Uncertain [I + verb] usage was significantly higher for Native US vs Native CHI groups, and Reluctant condition showed no significant between-group differences.

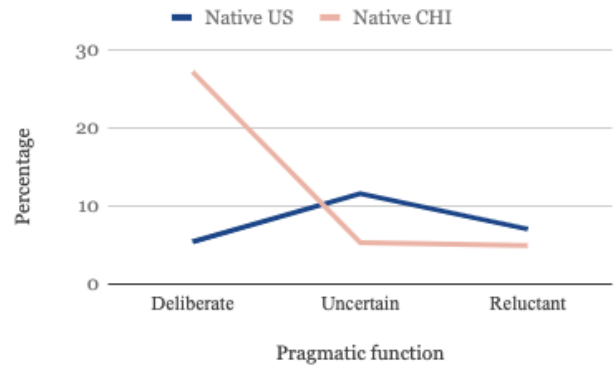


Figure 2: Mean % pragmatic function to total utterances for Opinion of others prompts.

4.2. Calculated Confidence Rating (CR) of speaker confidence

CR values were calculated for each participant based on CR values of prosodic-syntactic variations from Zhao, Dehe, and Murphy's (2017) proof-of-concept study [1]. Each prosodic-syntactic variation count was multiplied by % of total constructs per prosodic variation per participant for an average CR.

4.2.1. Self-opinion prompt set

Means for CR of the Native US group were 5.30 (SD = .33). Means for the Native CHI group were 4.36 (SD = .34).

ANCOVA results revealed a large significant between-group effect of sociolinguistic background ($F(1,3) = 54.28$, $p < .001$, $\eta_p^2 = .63$). Furthermore, post-hoc paired t-tests showed significant mean differences between groups ($p < .001$). CR was higher for Native US vs Native CHI group, paralleling the first analysis.

4.2.2. Opinion of others prompt set

Means for CR of the Native US group were 4.70 (SD = .32). Means for the Native CHI group were 5.61 (SD = .34).

ANCOVA results revealed a large significant between-group effect of sociolinguistic background ($F(1,3) = 36.92$, $p < .001$, $\eta_p^2 = .53$). Post-hoc paired t-tests showed significant mean differences between groups ($p < .001$). CR was higher for Native CHI vs Native US group, showing an inverse trend from the previous prompt set.

4.3. Contextual comparison

To identify whether frequency of functional [I + verb] usage varies across groups regarding Self-opinion vs. Opinion of others, a dependent samples t-test was conducted per group, comparing Deliberate, Uncertain, and Reluctant [I + verb] usage frequency between prompt sets. These comparisons were conducted for each dependent variable in the two analyses above (see Table 3).

Mean difference of Deliberate [I + verb] usage between Self-opinion vs. Opinion of others prompts was higher for Native US vs. Native CHI. Conversely, mean difference for reluctant usage was lower for Native US vs. Native CHI.

For Calculated CR, there once again appears a trend: Mean difference between Self-opinion vs Opinion of others is higher for Native US vs Native CHI, with significant differences within

Table 3: Paired *t*-test results between *Self-opinion* vs. *Opinion of others* prompt set

% Condition to total utterances	Group	Mean (SD)	T(1,24) Value
Deliberate	Native US	19.88 (10.55)	9.42
	Native CHI	-23.72 (9.62)	-12.33
Uncertain	Native US	8.95 (10.86)	4.12
	Native CHI	.36 (4.54)	.40
Reluctant	Native US	-2.60 (4.26)	-3.05
	Native CHI	10.60(7.32)	7.24
Calculated CR	Native US	.60 (.42)	7.13
	Native CHI	-1.25(.37)	-16.91

both groups.

5. Discussion and Conclusion

Findings from both prompt sets suggest that sociolinguistic background does influence frequency of functional [I + verb] usage when expressing an opinion. For *Self-opinion* prompts, more assertive and less reluctant frequency usage was linked to higher levels of US sociolinguistic background whereas more reluctant and less assertive frequency usage was linked to CHI sociolinguistic background; these trends reversed for *Opinion of others* prompts. These trends were further supported by calculated CRs for both prompt sets.

Furthermore, significant mean differences in calculated CR were found between prompt sets as a function of sociolinguistic background, with higher confidence in *Self-opinion* prompts associated with increased US sociolinguistic background, and increased mean CR difference in *Opinion of others* prompts associated with CHI sociolinguistic background. These results mirror ANCOVA results above, suggesting that US sociolinguistic tendencies may lead to more confidence when expressing one’s own preferences, whilst CHI sociolinguistic tendencies may lead to more confidence when passing social judgment.

Results revealed that the trends in significant mean differences at *Deliberate* and *Reluctant* conditions reflect each culture’s attitude toward expressing one’s own opinion vs. passing social judgment. Given that the characters under appraisal in the *Opinion of others* prompts are fictional, abstract entities, they should be processed as out-group members with no real connection to the participant; this should therefore lead to Native CHI participants’ more assertive language in expressing their opinion of an out-group member, especially regarding a socially unacceptable mistake. Results exemplify this prediction, thereby coinciding with notions that Chinese face-saving “indirectness” applies only to in-group communication and with more aggressive behavior toward the out-group [15, 16].

Alternatively, in the US where one’s identity is self-dependent and the individual self is highly valued, speakers may feel less inclined to have the right to passing judgment and may therefore use less direct language when evaluating on another’s mistake [4, 6]. Results further coincide with the notion that the opposite of positive evaluation in the US is not negative evaluation, but instead neutral evaluation, which is associated with less assertive language usage [17]. Combined with the notion that US individuals dispel negative information about the self in order to maintain positive self-regard [18], these findings may ex-

plain why US sociolinguistic background coincides with more hesitant and less assertive [I + verb] constructs when appraising others.

Conversely, this self-defined, high positive self-regard and absence of fear for negative appraisal may have led those with US sociolinguistic background to use more confident [I + verb] constructs when expressing personal beliefs. Results correspond to research that US individuals give compliments more strongly and frequently than CHI individuals [19], thereby fostering a comfortable and forgiving conversational floor. Further support comes from US speaker’s assumption of receiving positive evaluation when expressing oneself [17].

Moreover, US self-confidence when expressing personal preference coincides with findings regarding US individuals viewing disposition as defined internally, meaning that one’s own opinion is steadfast to one’s own beliefs and accordingly not easily swayed by others [20]. Thus, an internally defined US disposition would lead to more confident and direct communication style when expressing personal preferences. Hence, it appears that a highly tolerant conversational floor, combined with a high self-regard and focus placed on the “unique and precious individual” [17], fosters 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 identity as determined by the judgment of those around them [4, 6]. Indeed, for *Self-opinion* prompts, Native CHI group’s lower frequency of assertive [I + verb] and higher frequency of reluctant [I + verb] coincides with findings of less perceived tolerance for individual opinions and subsequent reluctance from fear of negative appraisal [8]. This negative appraisal would thereby arise when presenting an opinion maligned with cultural norms; accordingly, individuals may develop a more careful, indirect communication style, as results show. Results further correspond with research of a listener-focused communicative perspective of Chinese culture leading to caution when determining how another may react to one’s statement [21].

Whilst previous research on [I + verb] has investigated focus-marked [I + verb] variation within one language (e.g. [22]) or the particular DM [I + verb] across two languages (e.g. [13]), there is no work conducted on the full range of prosodic [I + verb] variations in a cross-linguistic design. This paper thereby contributes to 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.

Furthermore, the significantly different results between frequency of functional [I + verb] variations in *Self-opinion* and *Opinion of others* prompts highlight the importance of communicative context when generalizing about cultural behavior. Theoretical assumptions about homogeneity of language use across situations have led to unanimous notions of “directness” and “high modality” being associated with individualistic societies like the US, [4, 6] and notions of “indirectness” and “low modality” being associated with relational societies like China [8]. However, this study’s genre-specific results show these generalizations do not hold across contexts. Thus, future research on cultural theory should consider not only the issue of homogenizing a culture of individuals, but also the issue of homogenizing a culture’s behavior across different social contexts.

6. Acknowledgements

Research was supported by the Clarendon Fund and the Jesus College Old Members Fund.

7. References

- [1] L. Zhao, N. Dehe, and V. Murphy, "From pitch to purpose: The prosodic-pragmatic mapping of [i + verb] belief constructions in english and mandarin," *Journal of Pragmatics*, vol. 123, pp. 57–77, 2017.
- [2] G. Hofstede, *Culture's Consequences: International Differences in Work-Related Values*. London and Beverly Hills: Sage Publications, 1980.
- [3] —, "Dimensionalizing cultures: The hofstede model in context," *Online Readings in Psychology and Culture*, vol. 2, no. 1, pp. 57–77, 2011.
- [4] H. R. Markus and S. Kitayama, "Culture and the self: Implications for cognition, emotion, and motivation," *Psychological Review*, vol. 98, no. 2, pp. 224–253, 1991.
- [5] D. Oyserman and S. W. S. Lee, "Does culture influence what and how we think? effects of priming individualism and collectivism," *Psychol. Bull.*, vol. 134, pp. 311–342, 2008.
- [6] C. Sedikides, "A three-tier hierarchy of self-potency: Individual self, relational self, collective self," *Advances in Experimental Social Psychology*, vol. 48, pp. 235–295, 2013.
- [7] S. E. Cross, E. E. Hardin, and B. Gercek-Swing, "The what, how, why, and where of self-construal," *Personality and Social Psychology Review*, vol. 15, no. 2, pp. 142–179, 2011.
- [8] X. Zou, K. Tam, W. M. Morris, L. Lee, I. Lau, and C. Y. Chiu, "Culture as common sense: Perceived consensus vs. personal beliefs as mechanisms of cultural influence," *Journal of Personality and Social Psychology*, vol. 97, no. 4, pp. 579–597, 2010.
- [9] G. Hofstede, G. J. Hofstede, and M. Minkov, *Cultures and Organizations: Software of the Mind*. USA: McGraw-Hill, 2010.
- [10] J. E. F. Tree and J. C. Schrock, "Basic meanings of you know and i mean," *Journal of Pragmatics*, vol. 34, pp. 727–747, 2002.
- [11] M. del Mar Vanrell, I. Feldhausen, and L. Astruc, "The discourse completion task: status quo and outlook," in *Methods in prosody: A Romance language perspective (Studies in Laboratory Phonology (SILP))*, M. del Mar Vanrell, Ed. Berlin: Language Science Press, 2018, pp. 191–228.
- [12] P. Brown and S. C. Levinson, *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press, 1987.
- [13] L. Huang, *Discourse markers in spoken English: a corpus study of native speakers and Chinese non-native speakers (Ch. 8: Analysis of "I think")*. Birmingham: University of Birmingham, 2011.
- [14] L. Anthony, *AntConc (Version 3.2. 2)[Computer Software]*. Tokyo, Japan: Waseda University, 2011.
- [15] M. Brewer and Y. Chen, "Where (who) are collectives in collectivism? toward conceptual clarification of individualism and collectivism," *Psychol Rev*, vol. 114, no. 1, pp. 133–151, 2007.
- [16] M. Bond and S. Wang, "Aggressive behavior in chinese society: The problem of maintaining order and harmony," *Bulletin of the Hong Kong Psychological Society*, vol. 8, pp. 5–25, 1983.
- [17] T. Katriel and G. Philipsen, "'what we need is communication': 'communication' as a cultural category in some american speech," *Communication Monographs*, vol. 48, no. 4, pp. 301–317, 2002.
- [18] S. J. Heine, D. R. Lehman, H. R. Markus, and S. Kitayama, "Is there a universal need for positive self-regard?" *Psychological Review*, vol. 106, no. 4, pp. 766–794, 1999.
- [19] —, "Chinese immigrants' perceptions of semantic dimensions of direct/indirect communication in inter cultural compliment interactions with north americans," *The Howard Journal of Communication*, vol. 9, pp. 245–262, 1998.
- [20] I. Choi, R. E. Nisbett, and A. Norenzayan, "Causal attribution across cultures: variation and universality," *Psychol. Bull.*, vol. 125, pp. 47–63, 1999.
- [21] J. O. Yum, "The impact of confucianism on interpersonal relationships and communication patterns in east asia," *Communication Monographs*, vol. 55, pp. 374–388, 1988.
- [22] N. Dehé and A. Wichmann, "Sentence-initial i think (that) and i believe (that): Prosodic evidence for use as main clause, comment clause and discourse marker," *Studies in Language*, vol. 34, no. 1, pp. 36–74, 2010.
- [23] H. C. Triandis, *Culture and social behavior*. Boston: McGraw-Hill, 1994.