

Contrastive Analysis of Interpersonal Metafunction in Chinese and English TED Talks

Junning Li*

Guilin University, Guilin, 541003, China

*Corresponding author: yisa1234@126.com

Abstract: As a form of knowledge dissemination aimed at the public, TED talks demonstrate a strong demand for interpersonal interaction within the framework of one-way output. This study, grounded in the interpersonal metafunction theory of Systemic Functional Linguistics, conducts a comparative analysis of TED talks in both Chinese and English versions. It focuses on three progressive levels-the Mood system, the Modality system, and the Appraisal system-examining the stance differentiation between Subject and Finite, the participatory mobilization of interrogative and imperative structures, the graduation features of modalization and modulation, and the attitudinal presentation of affect, judgement, and engagement resources. The results indicate that English talks rely more on explicit grammatical markers and clear graduation scales, with clear stance boundaries and stronger contraction of dialogic space; Chinese talks, on the other hand, convey interpersonal meaning through modal particles, subject ellipsis, and contextual inference, exhibiting higher stance fluidity and more inclusive and mitigating engagement strategies. The differences between the two languages in interpersonal metafunctional encoding are not a matter of superiority or inferiority, but rather different realization paths of their respective linguistic systems and cultural conventions under the same communicative goal.

Keywords: interpersonal metafunction, TED talks, Chinese-English comparison, Mood system, Appraisal system.

Introduction

TED talks have achieved widespread global dissemination through their brief, concise, and emotionally engaging mode of expression, with a core feature being that speakers need to secure the audience's emotional identification and cognitive acceptance within a limited time. As one of the three metafunctions in Systemic Functional Linguistics, the interpersonal metafunction is specifically responsible for the mechanism by which language constructs social relations and expresses attitudinal stances, thus providing an appropriate theoretical tool for analyzing interactive strategies in this unidirectional communication genre. Chinese and English exhibit systematic structural differences in mood markers, modal resources, and evaluative lexis, and how these differences affect speakers' modulation of distance, authority, and empathy with the audience merits in-depth investigation^[1]. Existing studies have largely focused on the speech features of a single language or on macro-level rhetorical strategies, lacking a fine-grained cross-linguistic comparison grounded in the internal systems of the interpersonal metafunction. Drawing on a corpus of 30 TED talks with corresponding themes in Chinese and English (covering four major topics-technology, education, psychology, and social innovation-with 15 talks in each language per topic), this study compares the specific operational patterns of mood, modality, and appraisal resources, aiming to reveal the different linguistic pathways that the two languages rely on in achieving similar communicative intentions. In doing so, it deepens our understanding of cross-cultural verbal interaction patterns and provides new empirical evidence for comparative studies within Systemic Functional Linguistics.

1. Theoretical Foundations of the Interpersonal Metafunction and Cross-Linguistic Contrastive Dimensions

1.1 Core Components of the Interpersonal Metafunction and Its Interactive Negotiation Mechanisms

Within the framework of Systemic Functional Linguistics, the interpersonal metafunction is primarily realized through the two subsystems of Mood and Modality: the former defines the dialogic relations between communicative roles, while the latter regulates the negotiability of propositions. In the genre of TED talks—characterized by one-way dissemination yet a strong pursuit of interaction—speakers employ the Mood system to presuppose the audience as potential respondents, and through the graduation configuration of modalization and modulation, they exercise flexible control over the certainty of viewpoints and the force of action recommendations. Such interactive negotiation constitutes a dynamic regulatory process based on the axis of semantic negotiation, in which speakers must both maintain dominance in information delivery and leave room for the audience to identify with or question the content, thereby constructing a communicative zone that lies between informing and conversing^[1].

From the perspective of the operation of negotiation mechanisms, the core of the interpersonal metafunction lies in the continuous alternation of Subject and Finite to switch among speech functions such as statements, questions, and proposals. Each clause, while conveying experiential content, also delineates the power distance and psychological distance between the speaker and the audience. The frequent use of first-person plural subjects helps to create a sense of group belonging with shared responsibility, while a high proportion of modal auxiliaries alters the audience's gradient of trust in the truth value of propositions. Understanding the operational logic of this mechanism serves as the prerequisite foundation for the subsequent discussion of the Mood system and the Appraisal system. The choice of person in the Subject directly affects the degree of explicitness or implicitness of responsibility attribution, while the tense and modality markers in the Finite regulate the proximity between the proposition and reality; the two jointly determine the keynote of each round of semantic exchange^[2].

1.2 Similarities and Differences in Linguistic Resources for Encoding Interpersonal Meaning in Chinese and English Contexts

The grammatical and lexical resources that Chinese and English rely on for encoding interpersonal meaning differ significantly, resulting in distinct interactional styles in TED talks across the two languages. English, through its strict subject-verb agreement and a rich system of modal auxiliaries (e.g., can, might, should), is able to precisely grade different degrees of possibility, obligation, and willingness. Chinese, by contrast, lacks strict morphosyntactic marking, and its interpersonal meanings are conveyed more through modal adverbs (e.g., *yexu* 'perhaps', *yinggai* 'should'), sentence-final particles (e.g., *ba*, *ne*, *ma*), and lexical modal expressions. Such resource disparities determine that speakers of the two languages adopt different pathways when adjusting their stances and eliciting audience responses.

In addition to grammatical differentiation, the encoding of interpersonal meaning in Chinese and English contexts also reflects different pragmatic conventions. English TED talks tend to simulate a dialogic atmosphere through explicit modality markers and frequent interrogative structures, thereby reducing the sense of authority inherent in one-way statements. In Chinese talks, speakers often convey attitudinal nuances such as euphemism, deliberation, or assertion through the stacking of modal particles and intonational variations, relying more heavily on contextual cues and the audience's inferential abilities. These similarities and differences are not a matter of superiority or inferiority, but rather different optional pathways that the two linguistic systems offer when facing the same communicative intention; accurately identifying their specific manifestations is a crucial step in cross-linguistic comparison.

1.3 Analytical Levels and Operational Pathways for Cross-Linguistic Comparison of the Interpersonal Metafunction

Cross-linguistic comparative analysis requires the establishment of clear analytical levels to ensure comparability across different language materials. Within the domain of the interpersonal metafunction, this study sets three progressive levels: the first level is the Mood system, which examines the

distribution frequencies and shifting patterns of speech functions such as declarative, interrogative, and imperative; the second level is the Modality system, which analyzes the linguistic forms and graduation distributions of both modalization and modulation; and the third level is the Appraisal system, which focuses on the attribution, intensity, and accommodation of attitudinal resources. At each level, the analysis must start from linguistic forms and move upward to relate to the constructive effects of interpersonal meaning.

In terms of operational procedures, this study adopts a combination of functional annotation and qualitative comparison. Following the annotation framework of Systemic Functional Linguistics, this study identifies and categorizes the mood features, modal expressions, and evaluative lexis in the Chinese and English TED talk texts respectively, and subsequently counts the occurrence frequencies, combination patterns, and distribution patterns of each subcategory of resources across the three levels. The key point of the comparison is to reveal how different linguistic resources serve similar interactional goals, namely, to secure the audience's emotional identification and cognitive acceptance within a limited time. Through this multi-level, function-driven comparative approach, this study concentrates on the systematic differences and commonalities in the construction of interpersonal meaning^[2].

2. Mood System as a Regulatory Device in Chinese and English TED Talks

2.1 Role Differentiation of Subject and Finite in Constructing Speaker Stance

The pairing relation between Subject and Finite directly determines the speech function type of a clause, thereby shaping a specific stance distance between the speaker and the audience in TED talks. English relies on strict inversion rules between Subject and Finite to distinguish declaratives from interrogatives; speakers convey determinate information through regular subject-predicate structures while embedding tense and modality markers in the Finite to adjust the reliability of propositions. Chinese, by contrast, lacks overt morphological features of the Finite, and the choice of Subject depends more on topic-chain coherence and contextual factors; speakers blur the boundaries of responsibility by omitting subjects or adopting unmarked word order, thus creating a looser yet more inclusive stance space. Such structural differences make the stance of English speakers relatively clear and identifiable, whereas the stance of Chinese speakers exhibits greater fluidity and contextual dependency.

From the functional perspective of stance construction, the person choice of the Subject and the graduation configuration of the Finite jointly regulate the speaker's discursive authority and group-belonging implications. The first-person singular Subject combined with low-value modal Finite elements (e.g., might, could) tends to weaken the assertive force of personal claims, conveying a tentative observational conclusion; the first-person plural Subject, by contrast, narrows the psychological distance with the audience through implicit collective identification. In Chinese, the Subject is frequently omitted, and speakers rely on contextual cues to imply responsibility attribution, a practice that reduces the coerciveness of explicit stance declarations. The different operational patterns of Subject and Finite in the two languages reflect their respective strategies for coping with the demands of stance regulation in unidirectional communication.

2.2 Differentiated Mobilization of Audience Participation through Interrogatives and Imperatives

In TED talks, interrogative sentences are not primarily used for obtaining unknown information; instead, they serve as rhetorical devices to guide the audience along the thinking path set by the speaker. In English, interrogatives often take the form of tag questions or self-posed and self-answered questions, which activate the audience's desire to fill in information gaps by creating a brief sense of suspense, thereby simulating an interactive trajectory of dialogue in the listeners' minds. Chinese interrogative sentences, on the other hand, more frequently end with modal particles that indicate a question or a tentative tone, and are accompanied by a rising intonation, to convey an invitation for response—even though the speaker does not expect an actual verbal reply. Both languages exploit interrogatives to reduce the one-way nature of information delivery, but English tends toward logical-deductive intellectual guidance, whereas Chinese leans more toward emotionally resonant participatory invitations^[3].

Imperatives serve a different functional role in mobilizing audience participation, as their purpose lies in stimulating willingness to act rather than issuing commands. In English, imperatives often

employ base verb forms (e.g., Imagine, Consider) to guide the audience in cognitive operations, with a relatively direct tone but with content directed toward abstract thinking. In Chinese, imperatives frequently add mitigating elements such as qing 'please' or rang 'let' at the beginning, or adopt suggestive expressions like bufang... 'might as well...' to soften the imperative force. In comparison, interrogatives focus more on cognitive-level participation mobilization, whereas imperatives tend toward guidance at the level of behavioral intention. English speakers use the two sentence types with relatively balanced frequency, while Chinese speakers show a stronger preference for interrogatives to create a sense of interaction, and their use of imperatives is more restrained and accompanied by more politeness markers.

3. Gradational Features of Modalization and Modulation in Expressing Certainty and Obligation

3.1 Gradational Features of Modalization: Cross-Linguistic Calibration Strategies for Certainty Expressions

Modalization concerns the assessment of the probability of a proposition, ranging from high to low. In TED talks, it is manifested as the speaker's fine-grained regulation of the degree of certainty with which they hold their own viewpoints. English, relying on modal auxiliaries, constructs a systematic modal gradient from high-value items (must, will) to low-value items (may, might), allowing speakers to achieve continuous and quantifiable gradational control between strong assertion and weak speculation. This structural feature enables English-speaking speakers to clearly convey their self-assessment of the evidential weight of a proposition without resorting to paralinguistic means.

In contrast, Chinese lacks a morphologically consistent hierarchical system of modal auxiliaries and instead relies on a combination of modal adverbs (such as "definitely," "probably," and "perhaps"), sentence-final modal particles (which mark tentativeness or emphasis), and context-coordinated strategies to achieve similar gradational expressions. The gradational features in Chinese exhibit stronger pragmatic dependence and subjective fuzziness: high certainty is often realized by lexical items like "of course" and "without doubt," while low certainty is expressed through "possibly" and the tentative particle, supplemented by intonational downgrading. This difference leads Chinese speakers, when expressing non-absolute propositions, to tend overall to be more euphemistic and implicit than English speakers, and their gradational boundaries are also more susceptible to the influence of interactive context and audience inference.

3.2 Gradational Features of Modulation: Cross-Cultural Pragmatic Differences in Obligation Regulation

Modulation concerns the regulation of obligation and inclination in action recommendations, directly determining the strength of the action-imposing force that the speaker exerts on the audience. English constructs a clear strength gradient through the obligation-value hierarchy of modal auxiliaries (must > should > could > might): high-value expressions convey binding directives, while mid- and low-value expressions reserve room for the audience's autonomous choice. In TED talks, English speakers tend to give explicit recommendations directly at high-value modulation levels in order to quickly establish cognitive guidance, with their logic approximating "explicit instruction."

Chinese, by contrast, exhibits different modulation strategies. Chinese grammar builds a continuum from strong constraint to weak suggestion through a combination of obligatory modal auxiliaries (such as those meaning 'should,' 'ought to,' and 'may'), modal particles, and syntactic ellipsis. It is particularly noteworthy that Chinese speakers often significantly weaken the imperative force by omitting the subject (e.g., "should note..." rather than "you should note...") or by attaching mitigating strategies (e.g., "may try..."), thereby maintaining an equal dialogue posture with the audience. Comparative analysis shows that English TED talks prefer direct, mid-to-high-value obligation resources in modulation expressions, while Chinese speakers more frequently employ low-to-mid-value modulation and implicit mitigating mechanisms. Such differences in gradational orientation profoundly reflect the different acceptances of the two language cultures toward authority expression, interpersonal distance in public discourse, and listener autonomy^[4].

4. Attitudinal Presentation of the Appraisal System in Chinese and English TED Talks

4.1 Chinese-English Preferences and Intensity Differences in Affect Resources for Expressing Personal Experience

Affect resources attend to the speaker's emotional reactions to events or phenomena, and are frequently employed in TED talks to establish emotional bonds with the audience. English-speaking speakers tend to choose explicit affective lexis (e.g., happy, sad, frustrated) to directly convey inner feelings, and such expressions are relatively evenly distributed in intensity, ranging from low intensity (interested) to high intensity (terrified). Chinese-speaking speakers, in contrast, more often convey affect indirectly through metaphorical expressions or situational descriptions, for example, using somatic formulations such as "feeling very unsettled in the heart" in place of abstract emotional adjectives. This difference results in English affect transmission being more direct and identifiable, whereas Chinese affect expression requires the audience to decode it with reference to the surrounding context. Thus, the two languages form a contrast in the degree of explicitness versus implicitness in affect resources.

In terms of intensity distribution, English-speaking speakers tend to use high-intensity affect resources with relatively higher frequency, especially when recounting personal setbacks or breakthrough moments, where they are inclined to employ strong emotional markers. Chinese-speaking speakers, even when expressing similar experiences, often adopt intensity-modulating devices, such as adding mitigating adverbs like "somewhat" or "a bit," or appending a softening particle after emotional terms to lower the assertive force. This preference difference is not a matter of incompetence, but rather reflects different conventions between the two languages regarding the degree of emotional exposure. In English TED talks, the intensity peaks of affect resources typically occur in narrative climactic passages, whereas in Chinese talks they are more dispersed and accompanied by multiple mitigating markers, forming a relatively flattened emotional intensity curve.

4.2 Functional Differentiation between Social Dignity Orientation and Capacity Orientation in Judgement Resources

Judgement resources are used to evaluate people's behavior and character, and in TED talks they mainly manifest in two evaluative dimensions: social dignity orientation and capacity orientation. The social dignity orientation involves whether the behavior conforms to group norms and value criteria; English-speaking speakers, when evaluating others, often use terms such as "integrity" and "honesty," focusing on the consistency between individual behavior and social expectations. Chinese-speaking speakers also touch upon dignity-oriented evaluations, but they tend to embed them in specific situations or relational networks, avoiding isolated labeling of individuals with fixed tags. The capacity orientation focuses on whether an individual possesses the skills or talents required to accomplish specific tasks; in English talks, capacity evaluation is often presented through explicit adjectives (smart, talented), whereas Chinese talks tend to convey capacity judgments indirectly through behavioral descriptions (e.g., "this person can come up with such an idea")^[5].

The functional differentiation between the two orientations in TED talks is manifested in how speakers construct credibility. Social dignity judgement is mostly used to shape a positive image or to tactfully point out behavioral deviations, while capacity judgement serves the purpose of persuading the audience to adopt a certain viewpoint or method. English-speaking speakers show relatively little frequency difference in their use of social dignity and capacity-oriented resources, often deploying them alternately in their narratives. Chinese-speaking speakers exhibit a stronger orientation preference: the frequency of capacity judgement resources is notably higher than that of social dignity judgement, and the targets of capacity evaluation are mostly the speakers themselves or collective representatives, rather than third parties. This differentiation reflects the different paths of constructing "trustworthiness" under the two linguistic-cultural backgrounds, with one emphasizing moral integrity and norm conformity, and the other placing greater value on practical competence and outcome orientation.

Conclusion

This study systematically compares the realization of the interpersonal metafunction in Chinese and English TED talks from three levels, namely the Mood system, the Modality system, and the Appraisal

system. The findings reveal that English-speaking speakers tend to employ explicit grammatical markers (such as modal auxiliaries and inversion structures) and clear graduation scales to regulate stance and participation, with relatively clear-cut boundaries for the contraction and expansion of dialogic space; Chinese-speaking speakers, on the other hand, rely more on implicit linguistic devices such as modal particles, subject ellipsis, and contextual inference, and their stance expression exhibits higher fluidity, with engagement strategies tending toward inclusivity and mitigation. The two languages also display systematic differences in intensity distribution and orientation preferences in their use of affect and judgement resources. These findings suggest that cross-linguistic comparisons of the interpersonal metafunction should not remain at the level of formal enumeration, but should instead focus on how different linguistic resources serve similar interactional goals. Future research may extend the analytical scope to other oral communication genres (e.g., academic lectures and podcast conversations), or introduce empirical data on audience comprehension and feedback, so as to further verify the impact of interpersonal metafunctional encoding differences on actual communicative effects.

References

- [1] Ding, C., & Dong, J. (2023). *A contrastive study of modality in Chinese and English technology talks: Taking "Yixi" and TED Talk as examples*. *Journal of University of Shanghai for Science and Technology (Social Sciences Edition)*, 45(4), 382-388.
- [2] Guo, X. (2022). *A comparative analysis of attitudinal resources in Chinese and English inspirational speeches from the perspective of appraisal theory*. MA thesis, Xi'an International Studies University.
- [3] Zhang, S. (2022). *A comparative study of the use of hedges by Chinese and English TED speakers from the perspective of adaptation theory*. MA thesis, Jilin University.
- [4] Wang, W. (2021). *A multidimensional analysis of register features in English TED talk discourse*. *Foreign Language Education*, 42(2), 23-28.
- [5] Ma, X. (2021). *A comparative study of stance markers in the narrative parts of Chinese and English speeches*. MA thesis, Hebei Normal University.