

Xiaoting Li, Tsuyoshi Ono (Eds.)  
**Multimodality in Chinese Interaction**

# **Applications of Cognitive Linguistics**



Editors

Gitte Kristiansen

Francisco J. Ruiz de Mendoza Ibáñez

Honorary editor

René Dirven

## **Volume 34**

# Multimodality in Chinese Interaction

---

Edited by  
Xiaoting Li  
Tsuyoshi Ono

**DE GRUYTER**  
MOUTON

ISBN 978-3-11-046033-9  
e-ISBN (PDF) 978-3-11-046239-5  
e-ISBN (EPUB) 978-3-11-046051-3  
ISSN 1861-4078

Names: Li, Xiaoting, editor. | Ono, Tsuyoshi (Linguist), editor.  
Title: Multimodality in Chinese interaction / edited by Xiaoting Li,  
Tsuyoshi Ono.

Description: Berlin ; Boston : De Gruyter, [2019] | Series: Applications of  
cognitive linguistics (ACL) ; band/volume 34 | Includes bibliographical  
references and index.

Identifiers: LCCN 2018025229 (print) | LCCN 2018049299 (ebook) | ISBN  
9783110462395 (pdf) | ISBN 9783110460339 (alk. paper) | ISBN 9783110460513  
(epub)

Subjects: LCSH: Chinese language–Discourse analysis. | Chinese  
language–Grammar. | Modality (Linguistics) | Nonverbal communication. |  
Social interaction–China. | Semiotics.

Classification: LCC P302.15.C4 (ebook) | LCC P302.15.C4 M85 2019 (print) |  
DDC 495.101/41–dc23

LC record available at <https://lccn.loc.gov/2018025229>

**Bibliographic information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2019 Walter de Gruyter GmbH, Berlin/Boston  
Typesetting: Integra Software Services Pvt. Ltd.  
Printing and binding: CPI books GmbH, Leck

[www.degruyter.com](http://www.degruyter.com)

# Contents

Xiaoting Li, Tsuyoshi Ono

**Introduction: A multimodal approach to Chinese interaction — 1**

## Part I: Theory and methodology

Sandra A. Thompson

**Multimodality and the study of Chinese talk-in-interaction — 13**

Xiaoting Li

**Researching multimodality in Chinese interaction: a methodological account — 24**

## Part II: Multimodal practices

Hongyin Tao

**List gestures in Mandarin conversation and their implications for understanding multimodal interaction — 65**

K.K. Luke and Xiaoling He

**Hand gestures and emergent speakership: A study of turn competition and gesticulation in Cantonese conversation — 99**

Kawai Chui

**Grounding and gestural repetition in Chinese conversational interaction — 119**

Tomoko Endo

**Embodying stance: *wo juede* ‘I feel/think’ and gaze — 148**

## Part III: Multimodal organization of talk and interaction

Xiaoting Li

**Multimodal turn construction in Mandarin conversation – Verbal, vocal, and visual practices in the construction of syntactically incomplete turns — 181**

Ni-Eng Lim

**On co-operative modalities in the formulation of Mandarin Chinese  
turn-continuations — 213**

Liang Tao

**Self-repair in Mandarin Chinese: The multimodality of conversation — 255**

I-Ni Tsai

**A multimodal analysis of tag questions in Mandarin Chinese multi-party  
conversation — 300**

**Index — 333**

Xiaoting Li, Tsuyoshi Ono

# Introduction: A multimodal approach to Chinese interaction

The past few years have seen a dramatic increase of interest in the study of multimodality and Chinese interaction. This is partly due to the advancement of technology that allows us to examine social interaction as temporally unfolding multimodal events. In addition, the increasing awareness of the multimodal nature of social interaction has prepared the theoretical ground for studying Chinese interaction from a multimodal perspective. However, in contrast to the heightened interest, empirical multimodal research on Chinese interaction is still scarce. This volume stems from the theme session “Language and the Body in Chinese Spoken Discourse” at the 12th International Cognitive Linguistics Conference and the Workshop “Multimodality and Chinese Interaction” in 2013. It brings together scholars working on the methodology of multimodal analysis and the utilization of multimodal approach in studying Chinese interaction. Some papers in this volume draw on a conversation analytic approach to multimodal interaction, whereas others come from a cognitive linguistic perspective. But they all use data of unscripted Chinese conversational interaction and share the goal to explore the inherently multimodal property of Chinese interaction. In this introduction, we first offer an account of the multimodal approach this volume adopts and some issues related to multimodality. Then we discuss some unique features of Chinese grammar that contribute to the multimodal study of social interaction from a cross-linguistic perspective. Finally, we provide a summary of the chapters in this volume.

## 1 Multimodality

Multimodality refers to a collection of approaches that perceives and examines human communication as multimodal events involving multiple means of communication including (but not limited to) language, image, gesture, gaze, posture, objects in the environment etc. Multimodality does not prioritize any particular means of communication or modality such as language, but considers communication as a multimodal achievement. These approaches to multimodality are rather heterogeneous ranging from computer science and robotics to linguistics, sociology, and communication studies. Papers in this volume are from

---

**Xiaoting Li and Tsuyoshi Ono**, University of Alberta

<https://doi.org/10.1515/9783110462395-001>

a linguistic approach (broadly construed) to multimodality. Within the field of linguistics, multimodality is approached from several research traditions such as social semiotics (Halliday 1978; Zlatev 2016) and systemic functional grammar (Halliday 1985), mediated discourse (Scollon 1998, 2001) and multimodal (inter) actional studies (Norris 2004), gesture studies (Kendon 1980; Cienki and Müller 2008; Müller and Cienki 2009; McNeill 2013) and semiotics (Fricke 2012, 2013), corpus linguistics (Cameron and Deignan 2003; Allwood 2008; Bateman 2008; Kipp et al. 2009), and interactional linguistics (Couper-Kuhlen and Selting 2001; Selting 2013) and conversation analysis (Mondada 2007). Li (this volume a) provides a fuller account of these traditions and multimodal approaches.

Although these approaches differ in their research subjects (written or spoken interaction) and views of how to analyze the multimodal resources, they share two interrelated theoretical assumptions. One assumption is multimodal resources work together in expressing meaning and accomplishing action, and language is only one of the multimodal resources. These resources may mutually elaborate each other (convergence) or play off one another (divergence) in interaction (Li 2014, this volume b). The other assumption is each multimodal resource or semiotic system has its own organization. For example, language structure is governed by syntactic rules (Chomsky 1957) and built on various structural units (Bloomfield 1933), whereas gesture is described based on its kinesic properties (McNeill 1992) and can be delimited into gesture phrases and gesture units (Kendon 2004; Schröder 2017). Unveiling the complex working together of these multimodal resources in meaning-making and action-formation in interaction is a central task for students of multimodality and interaction.

## 2 Chinese interaction

The study of multimodality in Chinese interaction contributes to our understanding of the multimodal nature of human interaction from a cross-linguistic perspective. That is, the Chinese language provides specific resources and sets unique constraints for the organization of interaction in Chinese.

Three typological features of the language are of particular relevance to the organization of Chinese interaction: topic-comment structure, prevalent phrasal structure, and lexical tone. Topic-comment structure is one of the most striking features of Mandarin Chinese sentence structure in contrast to English which is a subject-prominent language (Li and Thompson 1976, 1981). A topic is usually followed by a pause and/or particles in natural speech (Li and Thompson 1981:15). This structure is relevant to turn organization in Mandarin



conversation. That is, participants orient to the complete topic-comment structure as relevant to possible turn completion, and refrain from taking the turn at the juncture between the topic and comment (Li 2014). Second, the common unit type for turn construction in Chinese is not necessarily clause, but rather phrasal structure (Tao 1996). It is argued that Mandarin syntax mainly exists at the phrase level (Zhu 1985; Lü 1979). A series of phrasal constructions or elliptical clauses that share a topic (commonly referred to as *liushuiju* ‘flowing water sentences’ [Lü 1979] in Chinese) are pervasive in Chinese spoken discourse (Tsao 1990; Wang and Li 2014). For example, in Excerpt 1 the phrasal constructions and elliptical clauses in lines 4 to 7 have the same topic: those college applicants with four hundred marks (lines 1–3).

(1) Tao (1996:82) (orthography slightly modified, line numbers added)

- 01 *sì bǎi fēn yǐshàng de*,  
four hundred mark above NOM  
‘those with four hundred marks,’
- 02 *jiùshì kǎoshēng ā*,  
indeed examinee PRT  
‘those college applicants,’
- 03 *dádào sì bǎi fēn yǐshàng de*.  
reach four hundred score above NOM  
‘those who have achieved four hundred marks and above,’
- 04 *bào zhíyè gāozhōng de*,  
apply vocation high school NOM  
‘those who have applied to a vocational school,
- 05 *hái yǒu hǎo duō*,  
still have very many  
‘there are still a lot of them,’
- 06 *jiù méi yǒu*.  
still NEG have  
‘(they) haven’t,’
- 07 *gēnběn jiù tóudàng bùchūqù*,  
basically somehow accept NEG out  
‘(they) are basically accepted by nobody.’

How these phrases or “fragments” of sentences build speech is not governed by grammar, but pragmatic or discourse factors in Chinese (Qian 1997; Liao 1992). Thus, Chinese grammar is called “discourse grammar” which is not concerned with sentence, but discourse structure (Qian 1997; Liao 1992; Chu 1999; Jiang 2005). Despite the differing opinions regarding the structural level that

Chinese grammar is mainly concerned with, it is generally acknowledged that the position and temporal production of the “fragmented” speech units is fluid and flexible. One implication of the fluid production of syntactic structure is that prosody becomes particularly significant in indicating the possible completion of a unit and a turn. For example, in the midst of a flow of “fragmented” phrasal constructions, prosody at the end of each “fragmented” construction may help indicate if it may (or may not) be the possible completion of a speaker’s turn. The importance of prosody in demarcating the basic discourse unit in spoken Chinese has long been acknowledged (Chao 1968; Wang and Li 2014). The third typological feature relevant to multimodal analysis of Chinese interaction is lexical tone. In tone languages such as Chinese, pitch movement or pitch contour is primarily used to distinguish lexical meanings rather than conveying pragmatic information (Liao 1994). The unit-final pitch movement or pitch contour may be affected by lexical tones of the unit-final syllables, and thus may not be very much relevant to turn organization. The projection of possible turn completion may be accomplished through other prosodic/phonetic parameters such as pitch register, pitch range, duration, pause etc. (Li 2014).

From a cross-linguistic perspective, these distinctive features of Chinese grammar and prosody expand our understanding of the practices used in interaction that may not exist in interaction conducted in other languages. However, we are not unaware of the methodological and theoretical issues raised by a cross-linguistic perspective in conversation analysis. Schegloff and Sacks (1974:234–235) refrains from linking findings about interaction to characterizations of data such as participant attributes and ethnicities, unless “warrant can be offered for the relevance of such characterizations of the data from the data themselves.” Sidnell (2010:10) also inquires “how can we, for instance, warrant a description of the data as ‘requests in Polish’, ‘French compliments’ or whatever where, typically, the participants do not display any overt orientation to the relevance of the language being used or their ethnicity?” As is also pointed out by Sidnell (2010:10), one solution to this challenge is to establish the connection between practices used in interaction to particular linguistic features of the language. In Mandarin Chinese (and other Chinese dialects), the topic-comment structure, pervasive “fragmented” syntactic construction, and lexical tone provide specific resources and set unique constraints for the organization of turns and interaction. An examination of these resources and practices unique to Chinese interaction offers insights into the diversity of conduct used in the organization of human interaction. We use “Chinese interaction” in the title of this volume, as the papers speak about the resources and practices distinctive to the language.

Generally speaking, multimodal analysis of Chinese face-to-face interaction is 1) to explore the interactional function of particular multimodal practices (such as a type of body movement, a lexico-syntactic construction, and a prosodic feature) in Chinese interaction, and 2) to examine how multimodal resources (such as lexico-syntax, prosody, the body, objects in the surround, sequential position) are deployed together to organize talk and accomplish social actions in Chinese interaction. Chapters 3–5 (H. Tao, Luke, and Endo) belong to the first type of research, and Chapters 6–10 (Li, Lim, Chui, L. Tao, and Tsai) represent the second one.

### 3 The chapters

The subsequent ten chapters are divided into three parts. Part One contains two articles providing overviews of the historical and theoretical backgrounds and a methodological framework in the study of Chinese multimodal interaction:

**Sandra Thompson** contributes a general overview article examining the intersection between functional linguistics and the study of grammar-in-interaction. She surveys key developments in the areas of conversation analysis and multimodality, and then provides comments on multimodality in the study of Chinese interaction. It highlights theoretical and historical contexts in which scholars recently started to focus on bodily-visual behavior in Chinese interaction.

**Xiaoting Li's** article is divided into two main parts where she first reviews various approaches to multimodality in general and then lays out the specific steps to make multimodal analysis of Chinese interaction particularly focusing on transcription and analysis. These methodological foci are illustrated using actual Mandarin conversational segments.

Part Two contains three chapters, each focuses on particular linguistic and bodily-visual practices:

**Hongyin Tao** investigates the role of gesture in constructing list sequences in that a distinction between composite gestures (more or less fixed gestures with distinct listing qualities) and reiterative gestures (a series of gestures of various types produced intermittently with different items of a list) is proposed. These gesture types are shown to correspond to two broader types of discourse functions: to enhance the rhetorical effect for persuasion, exemplification, and clarification, and to do with discourse structuring, tracking, and interlocutor meta-interaction.

**Tomoko Endo** focuses on how gaze shift and the epistemic marker *wo juede* 'I think/feel' in the initial and final positions of a turn function to facilitate stance

expression and turn organization. She shows that, using these resources, speakers take two kinds of stance, one of (dis)affiliation and one of participation. In particular, speakers are found to look away from participants with conflicting opinions, and toward participants who they wish to select as the next speaker.

The previous articles all focus on Mandarin Chinese. **K.K. Luke and Xiaoling He** add another Chinese dialect to our discussion by focusing on two gesture-types in Cantonese multi-party interaction. In particular, Luke and He show that the raised hand/finger and the arm-tap can be used by participants to make a bid for speakership, and to make one in such a way as to display their orientation to their own status as non-current speaker and non-selected next speaker.

Part Three contains five chapters each of which explores how multimodal resources are deployed to organize turns, sequences, repairs, and accomplish social actions in Chinese interaction:

**Xiaoting Li's** article examines the multimodal construction of syntactically incomplete turns. Specifically, Li examines cases where a bodily-visual behavior such as iconic gesture and head shake completes a syntactically incomplete turn. She also discusses cases where bodily-visual behavior may not play any constitutive role in the incomplete syntactic structure, but are deployed along with prosody and sequential position to indicate and anticipate possible completion of the syntactically incomplete turns.

**Ni-Eng Lim** highlights the multimodal mechanics of constructing further talk past a possibly complete turn. He finds that syntactically discontinuous constituents after possible completion are still marked as “continuing” by the combination of prosodic cues (i.e. rush-throughs, lack of pitch reset, and declining intonation contour). He also finds that where further talk is prompted by recipient's lack of uptake or intervening talk by others, thereby resulting in prosodic separation between the host-TCU and its turn-continuation, the addition is still syntactically continuous due to the use of grammatical structures available in the language (e.g., serial verb construction, topic-comment structure, verb-resultative complement, etc.).

**Kawai Chui** examines two grounding sequence types where participants collaborate moment-by-moment to establish mutual understanding in progress. In sequences without gestural repetition, addressees typically manifest their understanding of the speaker's utterance by offering positive responses explicitly while the latter largely express implicit acknowledgements. In sequences with gestural repetition, which provides evidence for the bilateral and interactive nature of speaking, addressees mimic the speaker's previous gesture while expressing understanding mainly in an explicit way and often provide additional information, which is typically followed by the speaker's explicit acknowledgement.

**Liang Tao** examines same-turn self-repairs where speakers “stop an utterance in progress and then abort, recast or redo that utterance” (Fox et al. 2009: 60) by highlighting language specific tendencies in that a) repair initiation falls on single-syllable words, and dual-syllable ‘units’ which might not be a grammatical unit, b) on or after the last segment of the syllable (thus rarely on the onset consonant), and c) a single syllable or a dual-syllable phonological unit is recycled to adjust to the tone sandhi rules. She also examines the connection between repair and bodily-visual practices such as head turns, gaze shift from the recipients, and hand gestures.

**I-Ni Tsai** focuses on tag questions formulated by the two most frequent question tags in spoken Mandarin: *shi bu shi* ‘COP not COP’ and *dui bu dui* ‘right not right’. Based on the study of turn design partly through the inspection of body movement, body position, eye gaze, etc., Tsai finds that the tag questions are commonly deployed in multi-party conversations to embody speaker’s differential states of knowledge in relation to different co-participants, and to seek affiliation in a disaffiliating move.

## 4 Closing remarks

This volume contributes to the growing field of multimodality and embodied interaction by offering new insights into Chinese interaction from a multimodal perspective and proposing methodological framework for multimodal analysis of Chinese interaction. It presents the most recent findings on multimodal practices in aspects of Chinese interaction.

The interdisciplinary nature of multimodal research determines the studies in this volume are necessarily interdisciplinary. The papers are from a range of disciplinary backgrounds including CA/ethnomethodology, interactional linguistics/discourse-functional linguistics, cognitive linguistics, and psycholinguistics to name just three. Regardless of our disciplinary backgrounds, we have the common goal to explore the linguistic and embodied underpinnings of social action and interaction in Chinese. However, the field of multimodality in Chinese interaction is a new, less-studied territory. Few researchers have been trained from the outset to conduct microanalysis of linguistic structure, bodily movement, and action sequence in Chinese interaction. This volume is a small step towards an understanding of Chinese interaction that is closer to the reality of how people interact in their everyday life. We hope that this volume will stimulate interchange between Chinese linguists and researchers on multimodal interaction, and that more researchers will participate in this rapidly-developing conversation on multimodality and Chinese interaction.

## References

- Allwood, Jens. 2008. Multimodal corpora. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, 207–25. Berlin: Mouton de Gruyter.
- Bateman, John. 2008. *Multimodality and Genre*. Basingstoke: Palgrave Macmillan.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Cameron, Lynne & Alice Deignan. 2003. Combining large and small corpora to investigate tuning devices around metaphor in spoken discourse. *Metaphor and Symbol* 18(3). 149–160.
- Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Cienki, Alan & Cornelia Müller. 2008. Metaphor, gesture, and thought. In Raymond W. Gibbs Jr (ed.), *The Cambridge handbook of metaphor and thought*, 483–501. Cambridge: Cambridge University Press.
- Couper-Kuhlen, Elizabeth & Margret Selting. 2001. Introducing interactional linguistics. In Margret Selting & Elizabeth Couper-Kuhlen (eds.), *Studies in interactional linguistics*, 1–24. Amsterdam/Philadelphia: John Benjamins.
- Chu, Chauncey C. 1999. *A Discourse Grammar of Mandarin Chinese*. New York: Peter Lang.
- Fox, Barbara A., Wouk, Fay, Hayashi, Makoto, Fincke, Steven, Tao, Liang, Sorjonen, Marja-Leena, Laakso, Minna and Hernandez, Wilfrido Flores. 2009. A Cross-linguistic investigation of the site of initiation in same-turn self-repair. In Jack Sidnell (ed.), *Conversation analysis: Comparative perspectives*, 60–103. Cambridge: Cambridge University Press.
- Fricke, Ellen. 2012. *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken*. Berlin/Boston: Mouton de Gruyter.
- Fricke, Ellen. 2013. Towards a unified grammar of gesture and speech: A multimodal approach. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill & Sedinha Teßendorf (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.1)*, 733–754. Berlin/Boston: Mouton de Gruyter.
- Kendon, Adam. 1980. Gesture and speech: two aspects of the process of utterance. In Mary R. Key, (ed.), *Nonverbal communication and language*, 207–227. The Hague: Mouton.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kipp, Michael, Jean-Claude Martin, Patrizia Paggio & Dirk Heylen. (eds.). 2009. *Multimodal corpora*. Berlin, Germany: Springer.
- Halliday, Michael A K. 1978. *Language as social semiotic: The social interpretation of language and meaning*. Baltimore: University Park Press.
- Halliday, Michael A K. 1985. *An introduction to functional grammar*. London: Edward Arnold.
- Jiang, Wangqi. 2005. Hanyu de ‘Juzi’ yu yingyu de sentence [Chinese “sentence” and English sentence]. *Journal of PLA University of Foreign Languages* 28(1). 10–15.
- Li, Charles N. & Sandra A. Thompson. 1976. Subject and topic: A new typology of language. In Charles N. Li (ed.), *Subject and topic: A new typology of language*, 457–489. New York: Academic Press.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.
- Li, Xiaoting. 2014. *Multimodality, interaction and turn-taking in Mandarin conversation*. John Benjamins: Amsterdam/Philadelphia.

- Liao, Qiuzhong. 1992. *Liao Qiuzhong wenji [Essays of Liao Qiuzhong]*. Beijing: Beijing Yuyan Xueyuan Chubanshe.
- Liao, Rongrong. 1994. Pitch contour formation in Mandarin Chinese: A Study of tone and intonation. Ph.D. dissertation, Ohio State University.
- Lü, Shuxiang. 1979. *Hanyu Yufa Fenxi Wenti [Analysis of Chinese grammar problems]*. Beijing: Shangwu Yinshuguan.
- McNeill, David. 1992. *Hand and mind*. Chicago: University of Chicago Press.
- McNeill, David. 2013. Gesture as a window onto mind and brain, and the relationship to linguistic relativity and ontogenesis. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silvia Ladewig, David McNeill & Sedinha Teßendorf (eds.), *Body – language – communication. An international handbook on multimodality in human interaction*. Volume 1, 28–54. Berlin & Boston: De Gruyter Mouton.
- Mondada, Lorenza. 2007 Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies* 9(2). 195–226.
- Müller, Cornelia & Alan Cienki. 2009. Words, gestures, and beyond: Forms of multimodal metaphor in the use of spoken language. In Charles Forceville & Eduardo Urios-Aparisi (eds.), *Multimodal Metaphor*, 297–328. Berlin & New York: Mouton de Gruyter.
- Norris, Sigrid. 2004. *Analyzing multimodal interaction: A methodological framework*. London: Routledge.
- Qian, Guanlian. 1997. *Hanyu wenhua yuyongxue. [Pragmatics in Chinese culture]*. Beijing: Tsinghua University Press.
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50. 696–735.
- Schröder, Ulrike. 2017. Multimodal metaphors as cognitive pivots for the construction of cultural otherness in talk. *Intercultural Pragmatics* 14(4). 493–524.
- Scollon, Ron. 1998. *Mediated discourse as social interaction: A study of news discourse*. London: Longman.
- Scollon, Ron. 2001. *Mediated discourse: The nexus of practice*. London: Routledge.
- Selting, Margret. 2013. Verbal, vocal, and visual practices in conversational interaction. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill & Sedinha Teßendorf (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.1)*, 589–609. Berlin, Boston: Mouton de Gruyter.
- Sidnell, Jack. 2010. Comparative perspectives in conversation analysis. In Jack Sidnell (ed.), *Conversation analysis: Comparative perspectives*, 3–27. Cambridge: Cambridge University Press.
- Tao, Hongyin. 1996. *Units in Mandarin conversation: Prosody, discourse, and grammar*. Amsterdam: John Benjamins.
- Tsao, Fengfu. 1990. *Sentence and clause structure in Chinese: a functional perspective*. Taipei: Student Book.
- Wang, Hongjun & Rong Li. 2015. Lun hanyu yupian de jiben danwei he liushuiju de chengyin. [On the basic unit of Chinese discourse and the formation of rundown sentences]. *Yuyanxue Luncong* 49. 11–40.
- Zhu, Dexi. 1982. *Yufa Jiangyi [Lectures on Chinese grammar]*. Beijing: Shangwu Yinshuguan.
- Zlatev, J. 2016. Mimesis : The role of bodily mimesis for the evolution of human culture and language. In David Dunér & Göran Sonesson (eds.), *Human Lifeworlds: The Cognitive Semiotics of Cultural Evolution*, 63–82. Peter Lang Publishing Group.





---

## Part I: **Theory and methodology**



Sandra A. Thompson

# Multimodality and the study of Chinese talk-in-interaction

## 1 Introduction

Research in conversation analysis and interactional linguistics has long recognized that language and bodily-visual behavior must be considered together in understanding how conversation is organized.

In this brief survey, I offer an overview of the intersection between functional linguistics and the study of interactional linguistics. I survey key developments in the areas of conversation analysis (CA) and multimodality, and then provide a few comments on multimodality in the study of Chinese interaction.

## 2 Functional linguistics, conversation analysis, and grammar-in-interaction

‘Functionalism’ as a vibrant subfield of linguistics can be said to have arisen in the 1970’s, partially as a reaction to the burgeoning interest at the time in approaching grammar as a formal system. The primary assumption shared by functional linguists has consistently been that the forms and structures of language are adapted to, and shaped by, their communicative functions. Functionalists take the internal organization of language to be a complex adaptive response to the ecological settings in which language is found, the interactional functions which it serves, and the full cognitive, social, and physiological properties of the human user. Functional linguistic research aims to clarify the relationship between linguistic form and function, and to determine the nature of the functions which appear to shape linguistic structure.<sup>1</sup>

At the same time, the field of conversation analysis (CA) was beginning to emerge with the work of Harvey Sacks (cf. Sacks 1974, 1995, Schegloff 1968) and his colleagues and students. Although Sacks originally had no special interest in language in his early explorations of social order (Heritage 1984), his growing

---

<sup>1</sup> For representative references, see Comrie (1978a, 1978b), Dixon (1977, 1979), Givón (1979), Greenberg (1978), Langacker & Munro (1975), Li (1976), Lord (1973), and Schachter (1976).

interest in the organization of everyday talk in his work with Emanuel Schegloff eventually drew him to the details of language. When Sacks, Schegloff & Jefferson (1974) appeared, although the field of linguistics at that time was fairly thoroughly dominated by the generative paradigm with its emphasis on language as a ‘mental’ construct, this article came to have a substantial impact on linguistics.<sup>2</sup>

As early as the late 1970’s, some functional linguists were already beginning to realize the need to study linguistic function by examining language in everyday use, working with monologic narratives (e.g., Chafe 1980; Hopper 1979; Hopper & Thompson 1980), but also with dialogic conversation and other everyday interactions (e.g., Duranti & Ochs 1979; Keenan & Schieffelin 1976).

The 1980’s saw the beginning of a merger between these two research traditions, with ‘discourse-functional’ research being more influenced by research in CA, as linguists began to shift their attention away from strictly cognitive considerations, realizing that language may be best understood as a resource for the accomplishment of actions in social interaction, and that by studying people actually talking, we gain a deeper appreciation of the role of grammar in carrying out social actions.<sup>3,4</sup> Three major contributions to our understanding of grammar have arisen from this focus on grammar ‘at work’.

The first of these is the understanding that grammar cannot be a fixed property of human brains, but must be seen as emergent, constantly undergoing revision as it is deployed and redesigned in everyday talk, where ‘emergent structure’ is understood as a set of recurrent patterns in a given language that emerge from humans pursuing their ordinary interactional business of communicating information, needs, identities, attitudes, and desires (Auer & Pfänder 2011; Bybee 2010; Fauconnier and Turner 2008, Hopper 1987).

A second contribution of examining grammar at work is the discovery that if linguistics is to account for language in everyday use, then its perspective on the nature of grammar must be both cognitively realistic as well as interactionally sensible. In other words, as recurrent solutions to speakers’ social needs become habitual, these habitually used linguistic forms become cognitively ent-

---

<sup>2</sup> As a primary indication of its influence, it is the most widely cited article ever to have appeared in *Language*; as of October 2016, Google Scholar counts 13,469 citations to this article. For a discussion of the influence of CA on linguistics, and of linguistics on CA, see Fox et al. (2013).

<sup>3</sup> See Langacker (1990) and Taylor (2002) for similar claims from a cognitive perspective.

<sup>4</sup> For discussions of grammar and social action, see, e.g., Clark & Wilkes-Gibbs (1986), Couper-Kuhlen (2014), Evans (2007), Ford (1993), Fox (2007), Goodwin & Goodwin (1987, 1992a), Hopper (2004), Kärkkäinen (2012), Fox & Thompson (2010), Laury, Ono & Suzuki (forthcoming), Curl (2006), Curl & Drew (2008), Thompson & Couper-Kuhlen (2005), Thompson (forthcoming), and Thompson et al. (2015).

renched as the patterns we call the grammar of a language. We could say, then, that ‘grammar’ is what has been ritualized from social interactions, resulting in a very loosely organized set of richly and complexly categorized memories people have of how they and fellow speakers have resolved recurrent interactional problems (Bybee 2006, 2010; Tomasello 1998, 1999, 2002, 2003).

A third contribution to our understanding of grammar at work is the discovery that grammar is tightly intertwined with the social activities that people are engaged in (Auer 1992, 1996; Clark 1996; Goodwin & Goodwin 1987, 1992a, 1992b; Schegloff 1996). One way in which these activities implicate the nature of grammar is that certain kinds of activities give rise to certain recurrent kinds of grammar. For example, the activities motivating posing questions and giving answers have a number of grammatical consequences (Fox & Thompson 2010; Heritage & Roth 1995; Raymond 2003; Schegloff 1996; Weber 1993; Thompson, Fox & Couper-Kuhlen 2015). Thompson, Fox & Couper-Kuhlen (2015) show, for instance, that English speakers respond to question-word questions such as *What did you have for lunch?* with either a lexical/phrasal response (e.g., *soup*) or a clausal response (e.g., *I had soup (for lunch)*), and that the choice seems to be determined by whether the responder takes the question, or responding to it, as problematic in some way.

### 3 Multimodality and the study of language and the body

An emphasis on the significance of multiple modalities in joint meaning-making can be traced to the groundbreaking and foundational work of Charles Goodwin, particularly Goodwin (1979, 1981). Among the first to analyze social interaction in terms of video data, Goodwin (1979) forged an entirely new area of research within the study of grammar in talk in interaction by showing how the construction of an ordinary ‘sentence’ depended on the gaze behavior of the speaker and his recipients. Goodwin (1981) further revealed the extent to which bodies are involved in the way humans use language to interact, and, strikingly, the way the very grammatical shape of their utterances can depend on the use of their bodies. Since that time, a research tradition focusing on the skilled use of multiple modalities by interactants has emerged within studies of language in interaction. For exemplary analyses, see Fox (2001), Goodwin (2000), Hayashi (2003), Mondada (2006, 2011), and Streeck (1993, 1994, 2009); for state-of-the-art treatises see Depperman (2013), Sidnell & Stivers (2005), and Streeck et al. (2011).

## 4 Multimodality in Chinese

### 4.1 Chinese discourse-functional linguistics

Just as elsewhere in linguistics, in the late 1970's and early 1980's, the field of Chinese linguistics began to see a growing interest in functional approaches to linguistic structure. Cheng, Li & Tang (1979), Chu (1983), Huang (1982, 1983), Li & Thompson (1981), Lü (1979), Paris (1981), and Teng (1975) capture the essence of this development.

And as research on conversation began to flourish and attract scholarly attention in the late 70's, so the study of Chinese grammar and interaction came into its own at about this time. Tsao (1979) is arguably the first major linguistic work based on Mandarin conversational data, followed by a rich range of studies on such topics as word order, the *ba*-construction, and final particles in everyday conversation. Many important contributions have appeared since then (e.g., Tao & Thompson 1994; for a state-of-the-art discussion, see Biq, Tai & Thompson (1996). More recent contributions include Biq (2001, 2004a, 2004b), Huang (1999), Luke & Zhang (2007), H. Tao (1999, 2001, 2003), L. Tao (2001, 2006), Wu (2002, 2004, 2005, 2006, 2011), Zhang (2012), and papers in *Chinese Language and Discourse* (see just below). An important stimulus to the growth of Chinese interactional linguistics was the founding by K.K. Luke, Hongyin Tao, and Li Wei, in 2010, of a new journal, *Chinese Language and Discourse* (<https://benjamins.com/#catalog/journals/cld/main>). A number of significant contributions in both interactional linguistics and CA have appeared in the pages of this journal since its inception. Innovative and equally important is the appearance of Huang's *Chinese Grammar at Work* (2013), a masterful discussion of a wide variety of frequently found constructions in Mandarin conversation.

### 4.2 Chinese CA

Early work in CA arose out of the field of sociology; its earliest practitioners were American and British sociologists. Their ground-breaking work in founding and establishing an entirely new field of endeavor in the study of language notwithstanding, it is also the case that these scholars were neither trained in linguistics nor were they fluent in, or particularly interested in, languages other than English. In the 1980's, however, scholars speaking languages other than English began to study CA and extend its findings to their own languages, and this is when CA scholarship on Chinese was initiated.

The beginning of Chinese CA can be traced directly, then, to the University of York, where K.K. Luke was studying; his 1990 book, *Utterance particles in Cantonese conversation*, is not only the first major contribution to Chinese CA, but Luke courageously takes on the issue of final particles, notoriously one of the most difficult areas of study for any language. Importantly, he also makes explicit the crucial link between CA, that is, the study of the organization of interaction, and grammar, that is, the study of the organization of linguistic patterns: he argues that applying CA techniques to Cantonese final particles reveals both a) what kinds of interactional problems speakers use these particles to address, and b) how solutions to these problems have impinged upon the structure of the Cantonese language.

Since then, more speakers of Chinese languages have joined the community of scholars working on CA and working with Chinese conversation. Following Luke, notable examples include Zhang (1998), the first in-depth study of repair in Chinese, and Wu (2004), the first book-length treatise on the Mandarin final particles *ou* and *a* in conversation, with special consideration of prosody and stance. These scholars have typically had training in both linguistics and CA, and their research, in focusing on grammar as social action happening in real time, has contributed greatly to our appreciation of aspects of Chinese grammar that have resisted analysis with methods utilizing context-free constructed example sentences. At the same time, along with much current research on conversational data in languages other than English (e.g., Cha'palaa, German, Finnish, Japanese, Korean, Lao, Tzeltal, Yéllî Dnye),<sup>5</sup> this work on Chinese conversation has in turn helped to change the field of CA, bringing strong awareness of language diversity to the study of the organization of everyday interactions.

### 4.3 Multimodality in the study of Chinese interactions

As was happening elsewhere in the field of CA, scholars working on Chinese conversation soon began analyzing the use of the body in everyday interactions. Arguably the first such works were Wu (1997) and H. Tao (1999); several articles on gesture followed soon thereafter, e.g., Chui (2003, 2005a, 2005b). Yang (2011) relates nonverbal bodily behavior with taking, yielding, and maintaining turns in everyday conversation.

---

<sup>5</sup> For influential examples of a comparative approach to CA, see Sidnell (2009) and Haakana, Laakso & Lindström (2009).

A major step forward in the study of bodily-visual behavior in Chinese interactions was the appearance of Li (2011), which attracted ‘underground’ attention even before it appeared as a book (Li 2014). This was the first book-length study of multimodality in Chinese conversation, and its findings on body movements and the construction of turns opened the way for the launching of a lively new research area among students and scholars of Chinese talk-in-interaction, with several conference panels, and now the appearance of the volume of papers you have before you.

## References

- Auer, Peter. 1992. The neverending sentence: Rightward expansion in spoken language. In Miklós Kontra & Tamás Váradi (eds.), *Studies in spoken languages: English, German, Finno-Ugric*, 41–59. Budapest: Linguistics Institute, Hungarian Academy of Sciences.
- Auer, Peter. 1996. On the prosody and syntax of turn-continuations. In Elizabeth Couper-Kuhlen & Margret Selting (eds.), *Prosody in conversation*, 57–100. Cambridge: Cambridge University Press.
- Auer, Peter & Stefan Pfänder (eds.). 2011. *Constructions: Emerging and emergent*. Berlin: de Gruyter.
- Biq, Yung-O. 2001. The grammaticalization of *jiushi* and *jiushishuo* in Mandarin Chinese. *Concentric: Studies in English Literature and Linguistics* 27(2). 103–124.
- Biq, Yung-O. 2004a. Construction, reanalysis, and stance: ‘V yi ge N’ and variations in Mandarin Chinese. *Journal of Pragmatics* 36(9). 1655–1672.
- Biq, Yung-O. 2004b. People, things and stuff: General nouns in spoken Mandarin. *Concentric: Studies in Linguistics* 30(1). 41–64.
- Biq, Yung-O, James H-Y Tai, & Sandra A. Thompson. 1996. Recent developments in functional and discourse approaches to Chinese. In James C.T. Huang & Audrey Li (eds.), *New Horizons in Chinese Linguistics*, 97–140. Boston: Kluwer.
- Bybee, Joan. 2006. From usage to grammar: The mind’s response to repetition. *Language* 82(4). 529–551.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Chafe, Wallace (ed.). 1980. The pear stories: Cognitive, cultural, and linguistic aspects of narrative production. Norwood, NJ: Ablex.
- Cheng, Robert L., Ying-che Li, & Ting-chi Tang (eds.). 1979. *Proceedings of symposium of Chinese linguistics*. Taipei: Student Book Co.
- Chu, Chauncey C. 1983. A reference grammar of Mandarin Chinese for English speakers. New York: Peter Lang.
- Chui, Kawai. 2003. Categorization of gestures in communication. In Lily I-Wen Su, Chinfa Lien, & Kawai Chui (eds.), *Form and function: Linguistic studies in honor of Shuanfan Huang*, 105–129. Taipei: Crane.
- Chui, Kawai. 2005a. Topicality and gesture in Chinese conversational discourse. *Language and Linguistics* 6(4). 635–654.
- Chui, Kawai. 2005b. Temporal patterning of speech and iconic gestures in conversational discourse. *Journal of pragmatics* 37(6). 871–887.
- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.



- Clark, Herbert H. & Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22. 1–39.
- Comrie, Bernard. 1978a. *Aspect*. Cambridge: Cambridge University Press.
- Comrie, Bernard. 1978b. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin: University of Texas Press.
- Couper-Kuhlen, Elizabeth. 2014. What does grammar tell us about action? In Ritva Laury, Marja Etelämäki & Elizabeth Couper-Kuhlen (eds.), *Approaches to grammar for Interactional Linguistics*. [Special issue]. *Pragmatics* 24(3). 623–647.
- Curl, Traci S. 2006. Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics* 38. 1257–1280.
- Curl, Traci S. & Paul Drew. 2008. Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction* 41(2). 1–25.
- Deppermann, Arnulf (ed.). 2013. Conversation analytic studies of multimodal interaction. [Special issue]. *Journal of Pragmatics* 46(1).
- Dixon, R. M. W. 1977. Where have all the adjectives gone? *Studies in Language* 1. 1–80.
- Dixon, R. M. W. 1979. Ergativity. *Language* 55(1). 59–74.
- Duranti, Alessandro & Elinor Ochs. 1979. Left-dislocation in Italian conversation. In Talmy Givón (ed.), *Discourse and syntax*, 377–416. New York: Academic Press.
- Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness*, 366–431. Oxford: Oxford University Press.
- Fauconnier, Gilles and Mark Turner. 2003. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Ford, Cecilia E. 1993. Grammar in interaction: Adverbial clauses in American English conversations. Cambridge: Cambridge University Press.
- Fox, Barbara A. 2001. On the embodied nature of grammar: Embodied being-in-the-world. In Joan Bybee & Michael Noonan (eds.), *Complex sentences in grammar and discourse*, 79–100. Amsterdam: John Benjamins.
- Fox, Barbara A. 2007. Principles shaping grammatical practices: An exploration. *Discourse Studies* 9. 299–318.
- Fox, Barbara A. & Sandra A. Thompson. 2010. Responses to WH-questions in English conversation. *Research on Language and Social Interaction* 43(2). 133–156.
- Fox, Barbara A., Sandra A. Thompson, Cecilia E. Ford, and Elizabeth Couper-Kuhlen. 2013. Conversation analysis and linguistics. In Jack Sidnell and Tanya Stivers (eds.), *Handbook of Conversation Analysis*, 726–740. London: Blackwell-Wiley.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Goodwin, Charles. 1979. The interactive construction of a sentence in natural conversation. In George Psathas (ed.), *Everyday language: Studies in ethnomethodology*, 97–121. New York: Irvington.
- Goodwin, Charles. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32. 1489–1522.
- Goodwin, Charles & Marjorie H. Goodwin. 1987. Concurrent operations on talk: Notes on the interactive organization of assessments. *IPRA Papers in Pragmatics* 1(1). 1–54.
- Goodwin, Charles & Marjorie H. Goodwin. 1992a. Assessments and the construction of context. In Charles Goodwin & Alessandro Duranti (eds.), *Rethinking context*, 147–189. Cambridge: Cambridge University Press.

- Goodwin, Charles & Marjorie H. Goodwin. 1992b. Context, activity and participation. In Peter Auer & Aldo di Luzo (eds.), *The contextualization of language*, 77–99. Amsterdam: John Benjamins.
- Greenberg, Joseph H. 1978. Universals of human language. Volumes 1–4. Stanford: Stanford University Press.
- Haakana, Markku, Minna Laakso & Jan Lindström (eds.). 2009. *Talk in interaction – comparative dimensions*. Helsinki: Finnish Literature Society.
- Hayashi, Makoto. 2003. Language and the body as resources for collaborative action: A study of word searches in Japanese conversation. *Research on Language and Social Interaction* 36(2). 109–141.
- Heritage, John. 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Heritage, John & Andrew L. Roth. 1995. Grammar and institution: Questions and questioning in the broadcast news interview. *Research on language and social interaction* 28(1). 1–60.
- Hopper, Paul J. 1979. Aspect and foregrounding in discourse. *Syntax and Semantics* 12. 213–241.
- Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistics Society* 13. 139–157.
- Hopper, Paul J. 2004. The openness of grammatical constructions. *Chicago Linguistic Society* 40. 239–256.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251–299.
- Huang, Shuan-fan. 1982. *Papers in syntax*. Taipei: Crane.
- Huang, Shuan-fan. 1983. *Yuyan zhexue* [Philosophy of Language]. Taipei: Crane.
- Huang, Shuanfan. 1999. The emergence of a grammatical category definite article in spoken Chinese. *Journal of Pragmatics* 31(1). 77–94.
- Huang, Shuan-fan. 2013. *Chinese grammar at work*. Amsterdam: John Benjamins.
- Kärkkäinen, Elise. 2012. I thought it was very interesting: Conversational formats for taking a stance. *Journal of Pragmatics* 44(15). 2194–2210.
- Keenan, Elinor Ochs & Bambi Schieffelin. 1976. Topic as a discourse notion: A study of topic in the conversation of children and adults. In Charles N. Li (ed.), *Subject and topic*, 335–384. New York: Academic Press.
- Langacker, Ronald W. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. (Cognitive Linguistics Research 1.) Berlin/New York: Mouton de Gruyter.
- Langacker, Ronald W. & Pamela Munro. 1975. Passives and their meaning. *Language* 51(4). 789–830.
- Laury, Ritva, Tsuyoshi Ono & Ryoko Suzuki. Forthcoming. Questioning the clause as a crosslinguistic unit in grammar and interaction. [Special issue]. *Studies in Language*.
- Li, Charles. 1976. *Subject and Topic*. New York: Academic Press.
- Li, Charles & Sandra Thompson. 1981. *A functional reference grammar of Mandarin Chinese*. Berkeley: University of California Press.
- Li, Xiaoting. 2011. Syntax, prosody, and body movements in turn organization in everyday Mandarin face-to-face conversation. Beijing: Peking University dissertation.
- Li, Xiaoting. 2014. Multimodality, interaction, and turn-taking in Mandarin conversation. Amsterdam: John Benjamins.
- Lord, Carol. 1973. Serial verbs in transition. *Studies in African Linguistics* 4. 269–297.
- Lü, Shuxiang. 1979. *Hanyu yufa fenxi wenti* [Issues in the analysis of Chinese grammar]. Beijing: Shanwu Yinshuguan.

- Luke, Kang Kwong. 1990. *Utterance particles in Cantonese conversation*. Amsterdam: John Benjamins.
- Luke, Kang Kwong & Wei Zhang. 2007. Retrospective turn continuations in Mandarin Chinese conversation. *Pragmatics* 13(4). 605–635.
- Mondada, Lorenza. 2006. Participants' online analysis and multimodal practices: Projecting the end of the turn and the closing of the sequence. *Discourse Studies* 8(1). 117–129.
- Mondada, Lorenza. 2011. Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics* 43. 542–552.
- Paris, Marie-Claude. 1981. *Problèmes de syntaxe et de sémantique en linguistique chinoise*. Paris: Presses Universitaires de France.
- Raymond, Geoffrey. 2003. Grammar and social organization: Yes/no interrogatives and the structure of responding. *American Sociological Review* 68. 939–967.
- Sacks, Harvey. 1974. An analysis of the course of a joke's telling in conversation. In Richard Bauman & Joel Sherzer (eds.), *Explorations in the ethnography of speaking*, 337–353. Cambridge: Cambridge University Press.
- Sacks, Harvey. 1995. *Lectures on Conversation*. Oxford: Blackwell.
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.
- Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic or none of the above? In Charles N. Li (ed.), *Subject and topic*, 493–518. New York: Academic Press.
- Schegloff, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist* 70(6). 1075–1095.
- Schegloff, Emanuel A. 1996. Turn organization: One intersection of grammar and interaction. In Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and grammar*, 52–133. Cambridge: Cambridge University Press.
- Sidnell, Jack (ed.). 2009. *Conversation Analysis: Comparative perspectives*. Cambridge: Cambridge University Press.
- Sidnell, Jack & Tanya Stivers (eds.). 2005. Multimodal interaction. [Special issue]. *Semiotica* 156(1/4).
- Streeck, Jürgen. 1993. Gesture as communication I: Its coordination with gaze and speech. *Communication monographs* 60. 275–299.
- Streeck, Jürgen. 1994. Gesture as communication II: The audience as co-author. *Research on language and social interaction* 27(3). 239–267.
- Streeck, Jürgen. 2009. *Gesturecraft: The manufacture of meaning*. Amsterdam: John Benjamins.
- Streeck, Jürgen, Charles Goodwin & Curtis LeBaron (eds.). 2011. *Embodied Interaction: Language and body in the material world*. Cambridge: Cambridge University Press.
- Tao, Hongyin. 1999. Body movement and participant alignment in Mandarin conversational organization. In *Papers from the 35th Regional Meeting of the Chicago Linguistic Society, Vol. II: The Panels*, 125–139. Chicago: The Chicago Linguistic Society.
- Tao, Hongyin. 2001. Some Interactive Functions of Topic Constructions in Mandarin Conversation. In Proceedings of the Joint Meetings of the 10th International Association for Chinese Linguistics and the 13th North American Conference on Chinese Linguistics (Graduate Students in Linguistics Publications), 317–331. Los Angeles: University of Southern California. [Chinese translation by Yue Yao in Yuyanxue Luncong [Studies in Linguistics] 36. 363–376. Beijing: Commercial Press, 2007.]

- Tao, Hongyin. 2003. Toward an emergent view of lexical semantics. *Language and Linguistics* 4(4). 837–856.
- Tao, Hongyin & Sandra A. Thompson. 1994. The discourse and grammar interface: Preferred clause structure in Mandarin conversation. *Journal of the Chinese Language Teachers Association* 29(3). 1–34.
- Tao, Liang. 2001. Switch reference and zero anaphora: Emergent reference in discourse processing. In Alan Cienki, Barbara J. Luka & Michael B. Smith (eds.), *Conceptual and discourse factors in linguistic structure*, 253–269. Stanford: CSLI Publications.
- Tao, Liang. 2006. Classifier loss and frozen tone in spoken Beijing Mandarin: The YI+GE phono-syntactic conspiracy. *Linguistics* 44(1). 91–133.
- Taylor, John R. 2002. *Cognitive Grammar*. Oxford Textbooks in Linguistics. Oxford: Oxford University Press.
- Teng, Shou-hsin. 1975. *A semantic study of transitivity relations in Chinese*. Berkeley and Los Angeles: University of California Press.
- Thompson, Sandra A. Forthcoming. Understanding ‘clause’ as an emergent ‘unit’. [Special issue]. *Studies in Language*.
- Thompson, Sandra A. & Elizabeth Couper-Kuhlen. 2005. The clause as a locus of grammar and interaction. *Discourse Studies* 7(4–5). 481–505.
- Thompson, Sandra A., Barbara A. Fox & Elizabeth Couper-Kuhlen. 2015. *Grammar in everyday talk: Building responsive actions*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 1999. *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael (ed.). 1998. *The new psychology of language: Cognitive and functional approaches to language structure, volume I*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tomasello, Michael (ed.). 2002. *The new psychology of language: Cognitive and functional approaches to language structure, volume II*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tsao, Feng-fu. 1979. *A functional study of topic in Chinese: The first step toward discourse analysis*. Taipei: Student Book Co.
- Weber, Elizabeth. 1993. *Varieties of questions in English conversation*. Amsterdam: John Benjamins.
- Wu, Ruey-Jiuan Regina. 1997. “Transforming Participation Frameworks in Multi-Party Mandarin Conversation: The Use of Discourse Particles and Body Behavior.” *Issues in Applied Linguistics* 8(2): 97–118.
- Wu, Ruey-Jiuan Regina. 2002. Discourse-pragmatic principles for temporal reference in Mandarin Chinese conversation. *Studies in Language* 26(3). 513–541.
- Wu, Ruey-Jiuan Regina. 2004. *Stance in talk: A conversation analysis of Mandarin final particles*. Amsterdam: John Benjamins.
- Wu, Ruey-Jiuan Regina. 2005. “There is more here than meets the eye!”: The use of final *ou* in two sequential positions in Mandarin Chinese conversation. *Journal of Pragmatics* 37(7). 967–995.
- Wu, Ruey-Jiuan Regina. 2006. Initiating repair and beyond: The use of two repeat-formatted repair initiations in Mandarin conversation. *Discourse Processes* 41(1). 67–109.
- Wu, Ruey-Jiuan Regina. 2011. A conversation analysis of self-praising in everyday Mandarin interaction. *Journal of Pragmatics* 43(13). 3152–3176.

- Yang, Ping. 2011. "Nonverbal aspects of turn taking in Mandarin Chinese interaction." *Chinese Language & Discourse* 2(1). 99–130.
- Zhang, Wei. 1998. *Repair in Chinese conversation*. Hong Kong: University of Hong Kong dissertation.
- Zhang, Wei. 2012. Latching/Rush-Through as a Turn-Holding Device and Its Functions in Retrospectively Oriented Pre-Emptive Turn Continuation: Findings from Mandarin Conversation. *Discourse Processes* 49(3–4). 163–191.

Xiaoting Li

# Researching multimodality in Chinese interaction: a methodological account

Verbal and nonverbal activity is a unified whole, and theory and methodology should be organized or created to treat it as such.

(Pike, 1976:26)

## 1 Introduction

People utilize a variety of semiotic systems or modalities to produce talk and perform action in interaction. Sometimes, verbal and vocal behavior is the primary mode of communication such as in telephone conversation, while at other times bodily-visual conduct such as gesture, gaze, and posture may take on primacy; and perhaps even more often, verbal, vocal, and visual practices are orchestrated in nuanced ways to accomplish action in interaction. In order to gain a deeper understanding of how Chinese interaction is conducted and organized, a theoretically grounded methodological framework that deals with multimodal interaction is essential. Chinese linguistic research has advanced our knowledge of the organization, production, and comprehension of the language (though mostly in contrived settings) in numerous ways. But language is just one of the resources necessary for recognizing others' communicative and informative intentions. When looking at real-life interaction, it is apparent that we are still very much at the beginning of multimodal research on interaction.

This chapter is an attempt to propose a way to conduct multimodal analysis of Chinese interaction. Multimodal analysis is immensely multidisciplinary. The methodological framework proposed here is rooted in the research tradition of microanalysis of social interaction represented by conversation analysis and interactional linguistics. Before explicating the methods in conducting multimodal analysis of Chinese interaction in Section 3, I discuss the main approaches to multimodal analysis in linguistics and semiotics in Section 2.

## 2 Multimodal approaches

In this section, four main approaches to multimodality are briefly discussed. These approaches to multimodality are developed from different fields and disciplines, ranging from social semiotics, systemic functional linguistics, mediated discourse to conversation analysis. Due to the difference in their theoretical orientation, these approaches may have different analytical foci and methods in analyzing multimodal phenomena.

### 2.1 Multimodal discourse analysis

The first approach to multimodality is multimodal discourse analysis developed from Halliday's theories of social semiotics (Halliday 1978) and systemic functional grammar (Halliday 1985) with an initial focus on the semiotic resources of texts and images (Kress and Van Leeuwen 1996; Forceville 1996; Forceville and Urios-Aparisi 2009). Kress and Van Leeuwen (1996) describes that meaning is realized not only through language, but visually through dynamic selection of texts, images, and other locally available semiotic resources in communicating ideology and discourse. One of the important contributions of the social semiotic approach to multimodality is its proposal of the concept "semiotic resources". According to Van Leeuwen (2005: 285),

Semiotic resources are the actions, materials and artifacts we use for communicative purposes, whether produced physiologically – for example, with our vocal apparatus, the muscles we use to make facial expressions and gestures – or technologically – for example, with pen and ink, or computer hardware and software – together with the ways in which these resources can be organized.

Although this approach acknowledges the diversity of semiotic resources including both the signs of visual representations (such as text, image, music, space, objects, and mathematical symbols) and embodied actions (such as gesture, gaze, and posture), majority of the research in this approach concentrates on the former, while embodied actions are the focus of microanalysis of social interaction (Li, 2016) and meaning construction (Kappelhoff and Müller 2011; Forceville 2011).

Drawing on Halliday's systemic functional linguistics (SFL), O'Halloran (2000, 2004a, 2004b) explore the metafunctional systems of semiotic resources and how semiotic choices integrate in multimodal discourse such as mathematics texts and film (Müller and Cienki 2009; Forceville and Renckens 2013). Semiotic resources are systems of meaning integrated in multimodal phenomena in specific situational and cultural context. The SFL-oriented multimodal discourse

analysis is closely related to the social semiotic theory in that it strives to develop a multimodal social semiotic theory in meaning-making in multimodal objects and events in a culture (Jewitt 2009:33). Thus, the social semiotic and SFL approaches to multimodality are subsumed under the approach of multimodal discourse analysis.

## 2.2 Multimodal communication

The second approach to multimodality is multimodal communication which is also referred to as “multimodal interactional analysis” by Norris (2004). As this framework addresses communicative awareness and attention and centers on analyzing communicative mode, I will use the term “multimodal communication” to refer to this approach, and “multimodal interaction” to refer to the sequence-oriented microanalysis of social interaction in Section 2.4.

Multimodal communication studies draw on theories of mediated discourse (nexus of practice), interactional sociolinguistics (the ethnographic study of language use and identity construction), and social semiotic approach to multimodality (its attention to other semiotic resources such as gesture, music, and color) (Norris 2004; Norris and Jones 2005). The first step towards “multimodal interactional analysis” is communicative mode (Zima and Brône 2015). Norris (2004:12) adopts the social semiotic theory in viewing communicative modes as semiotic systems with rules and regularities (Kress and Van Leeuwen 2001) that have communicative function in interaction. The basic unit of analysis in this multimodal methodology is mediated action (Scollon 1998, 2001) which is further categorized into higher-level and lower-level actions. Through the analysis of mediated action, Norris (2004) explores how communicative modes are constitutive of action, identities, relations, as well as social interaction. By combining elements from multiple disciplines such as social semiotics, mediated discourse, and interactional sociolinguistics, “multimodal interactional studies” provides a methodological perspective to the complexity of the semiotic systems involved in human interaction.

## 2.3 Multimodal grammar

The third approach to multimodality is to incorporate gesture as a grammar system called multimodal grammar. It derives from the field of gesture studies. The importance of gesture in meaning-making has long been recognized and underlined by the pioneers in gesture studies (Kendon 1980, 1988; McNeill 1985). Gesture and speech have been considered as “two sides of one process of



utterance” (Kendon 1980, 2004). But the use of the term multimodality in the field of gesture studies and the proposal of a multimodal grammar is a relatively recent endeavor (Fricke 2012, 2013).

A body of research in gesture studies has documented the structure of gesture based on four formal parameters of sign language: hand shape, orientation, movement, and position (Stokoe 1960; Cienki 2013; Ladewig and Bressemer 2013), and described different types of gestures such as palm up open hand gesture (Müller 1998, 2004), recurrent gesture (Ladewig 2014), and away gesture (Bressemer and Müller 2014). Through the systematic documentation of the gesture forms, their meaning, and their syntagmatic and paradigmatic structure from a form-based perspective (what is called “a grammar of gestures”, Müller, Bressemer, and Ladewig 2013; Zima 2017), gesture is argued to have “potential for language” (Müller 2009) and “emerging linguistic structures” (Müller, Bressemer, and Ladewig 2013). That it is being used in conjunction with speech shows that spoken language is inherently multimodal, which points towards a multimodal grammar or multimodal construction grammar (Cienki 2017; Zima & Bergs 2017). Closely related to the linguistic documentation of gestures is the observation that gestures occupy syntactic positions and adopt syntactic functions such as NP, VP, adjectives and adverbs in a syntactic structure (Fricke 2012; Ladewig 2012). Such observations of the integration of gesture into syntactic structure of spoken language prepare the ground for the proposal of “multimodal grammar” (Fricke 2012; Müller, Bressemer, and Ladewig 2013:709). Fricke (2012, 2013) argues that gesture goes through the same processes of typification and semantization as spoken language, as is evidenced by two typified forms of pointing gesture in German. Gestural structures also exhibit the same features of constituency and recursion as language. The same principles underlying gesture and language allow the integration of these two (among other sign systems) into a multimodal grammar.

The approach of multimodal grammar broadens the scope of multimodality in linguistics from the study of gesture-speech relations or language-image relations to the theorizing of integrated grammar systems of language and “human faculty of language” (Fricke 2013:751).

## 2.4 Multimodal interaction

The fourth approach to multimodality is multimodal interaction.<sup>1</sup> It is also the approach adopted in this chapter. Thus, it will be discussed in more detail.

---

<sup>1</sup> Streeck, Goodwin, and LeBaron (2011) uses the term “embodied interaction in the material world” to emphasize the importance of viewing human interaction as situated in and co-constructed by the

Multimodal interaction is characterized by the rigorous microanalysis of the formation of action sequences by taking into account a full array of verbal, vocal, and visual resources in human face-to-face interaction. In contrast to the multimodal discourse analysts' interest in extending social semiotic analysis to text-image-artefact combinations, research in multimodal interaction is concerned with naturally-occurring face-to-face interaction. In this approach, face-to-face interaction is "by definition, multimodal interaction in which participants encounter a steady stream of meaningful facial expressions, gestures, body postures, head movements, words, grammatical constructions, and prosodic contours" (Stivers and Sidnell 2005:1), and located in the material world (Streeck, Goodwin, and LeBaron 2011).

To better understand multimodal interaction, reviewing its intellectual roots is useful. Multimodal interaction is formed through four streams of work: conversation analysis, interactional linguistics, linguistic anthropology, and workplace studies.

Conversation analysis (CA) was developed by American sociologists Harvey Sacks and Emanuel Schegloff in studying recordings of telephone calls for their broader inquiry of establishing an observational science of social action. CA aims to identify and explicate sequential structures and practices in the formation and organization of action and turn in human interaction. It achieves this aim through the microanalysis of the design and the sequential position of each turn and action in sequences of turns and actions in temporally unfolding interaction. The systematic attention to the minutiae of interaction distinguishes it from other multimodal approaches discussed in this section. CA contributes to the development of multimodal interaction by providing a reproducible method of analyzing actions (based on "position" and "composition") applicable to all forms of face-to-face interaction. However, although some early CA work addresses the visual aspects of social interaction (Sacks and Schegloff 2002[1975]), it tends to view bodily-visual conduct as subordinate to verbal conduct and insignificant to the structures and practices identified (Drew 2004:78). Adopting the CA method, research on multimodal interaction does not prioritize verbal conduct, but rather studies the ways in which multimodal practices such as talk, gaze, gesture, and posture are brought together to build coherent courses of actions in interaction (Stivers and Sidnell 2005). An increasing number of studies on multimodal interaction have shown that bodily-visual practices are not only relevant to the formation of action in situated interaction (Goodwin 2000a, 2000b, 2003;

---

material surround. Since material objects and structures in the environments are also modalities, we call this approach "multimodal interaction", which includes both embodied actions involving voice, mouth, hands, face and body, and the material world.

Hayashi 2003, 2005), but also transformative to the (re-) conceptualization of the unit of interaction (Keevallik 2013), and collaborative turn construction and organization (Goodwin 1979; Iwasaki 2009, 2011; Mondada 2007; Li 2014). In particular, Charles Goodwin's work connects the CA method to multimodal analysis of interaction, and illuminatingly demonstrates how action and interaction is constructed and (re-)configured through the mutual elaboration of talk, body, and the material surround (Goodwin 2000a, 2000b, 2003).

The point of departure for CA inquiry is action, rather than specific semiotic system or modality. Language as a semiotic system and its relevance to interaction is the central inquiry of interactional linguistics.

International linguistics is the second stream of work contributing to the formation of multimodal interaction. It is concerned with "how linguistic structures and patterns of use are shaped by, and themselves shape, interaction" (Couper-Kuhlen and Selting 2001:1). It begins with the study of prosody in interaction (Couper-Kuhlen and Selting 1996) by explicating the relevance of different prosodic features such as pitch (Couper-Kuhlen 2001, 2004), rhythm (Auer Couper-Kuhlen, and Müller 1999), and voice quality (Ogden 2001) to action formation and turn/sequence organization in interaction. Another focus of interactional linguistics is the role of lexico-syntactic structures in interaction (Ochs, Schegloff, and Thompson, 1996; Hakulinen and Selting 2005; Auer 2009; Couper-Kuhlen and Thompson 2005). Recently, multimodality begins to gain currency in interactional linguistics with a growing interest in bodily-visual practices and their interplay with linguistic structures in interaction (Selting 2013; Ford, Thompson, and Drake 2012; Walker 2012). A striking methodological feature of CA and interactional linguistics is participants' orientation. This methodology is implemented by demonstrating how a practice is deployed, and how it is treated by participants through their observable behavior displaying such an orientation (Walker 2004). This interactional perspective to prosodic (and linguistic) practices also applies to the analysis of other multimodal practices such as gesture, gaze, and posture. I will return to a fuller account of the concept of participants' orientation and its application to analyzing Chinese interaction in Section 3.3.

The third field of study that contributes to the formation of multimodal interaction is linguistic anthropology. Linguistic anthropology investigates the ways in which linguistic forms (e.g., prosody in Gumperz 1982) and embodied practices (Duranti 1992, 2016; M. H. Goodwin 2006) construct and organize culturally defined events and social life in and through interaction. In linguistic anthropology, taking into account of the human body and the built environment is crucial for the analysis of any situated and embodied interaction (Duranti 1997:322), and for understanding the meaning of composite utterances that are inherently composed of complex semiotics (Enfield 2009). For example, language, gesture, gaze,