

THE UNIVERSITY OF CHICAGO

GRAMMATICAL FORM AND 'THINKING-FOR-SPEAKING' IN
MANDARIN CHINESE AND ENGLISH:
AN ANALYSIS BASED ON SPEECH-ACCOMPANYING GESTURES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY
SUSAN DARDEN DUNCAN

CHICAGO, ILLINOIS

DECEMBER 1996

ACKNOWLEDGEMENTS

I have many people to thank for their help, intellectual guidance, encouragement, humor, patience, and forbearance during the extended period that was needed to complete this dissertation. Chief among them are my advisors, Professors Susan Goldin-Meadow and David McNeill, as well as my readers, Professors Boaz Keysar and Michael Silverstein.

Many other people have helped in different ways and I am very grateful to each of them. Particular thanks are reserved for four people. I thank my husband David Klein and my mother Anne Grimes Eberle. This work would not have come into being without their unending willingness to assist me at any time and in any way. Their encouragement, love and support made it possible.

I thank Karl-Erik McCullough for passing on to me his enthusiasm for this area of research. The smattering of footnotes in this dissertation is no catalog of his contributions to the development of my understanding.

I also particularly thank George Chao, instructor of Mandarin Chinese at the University of Chicago. His example shows why it is wonderful to do one's work, whatever it is, with integrity, focused attention, and creativity. It illuminates by contrast the pointlessness of doing work in which one is not invested in a heartfelt way. In the time since I studied Mandarin with Teacher Chao, I have encountered people all over the world who know and admire him.

During the revisions period, I benefitted, more than I know how to characterize, from conversations with S. Kita and Adam Kendon, Jan-Peter de Ruiter, Stephen Levinson, David Wilkins, Melissa Bowerman, Dan Slobin, Eric Pedersen, Elizabeth

Keating, Giovanni Bennardo, Christel Stolz, Juergen Bohnemeyer, Eva Schultz-Berndt, and others at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. These individuals all further extended and shaped my understanding of a broad range of linguistic and gestural phenomena. I am greatly indebted to them.

My friends Susan Alberts, Rick King, Lisa Michelbrink, Cornelia Mueller, Liz Stevenson and Robert Zimmerman, and all the members of both the Goldin-Meadow and McNeill labs helped and inspired me. They kept me in mind of the fact that I would have a real life to return to, peopled with wonderful friends.

Xiao-lei Wang, Lai Mei-yi, Hui-fang Hong, and Jane Y. Hsu shared their Chinese culture, sensibilities, time, and native-language expertise very generously. I am proud and grateful to have had their friendship.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	xi
Chapter	
1. INTRODUCTION	1
The Paradigm for the Analysis of Speech-Accompanying Gesture	1
Toward a Model of Language Production	2
Gesture and ‘Thinking-for-Speaking’	4
Speech-Gesture Timing and the ‘Growth Point’ of Utterance	9
2. TWO STUDIES OF THE RELATIONSHIP BETWEEN LINGUISTIC FORM AND CONCURRENT CONCEPTUALIZATION	14
General Methods	17
Participants	17
Stimuli	18
Procedure	19
Elicitation	19
Speech Transcription and Gesture Coding	20
Types of Gesture	21
The Mechanics of Unreflective, Representational Gesture	25
The Analytic Significance of Gesture Phases	27

3. STUDY ONE: GESTURAL INDICES OF VERB ASPECT AND <i>AKTIONSART</i>	31
Introduction	31
Linguistic Means of Conveying Aspectual View	34
Summary	37
Methods for Study One	38
Participants	38
Speech Sampling	38
Gesture Sampling	39
Sampling Totals	41
Gesture Coding	42
Coding Summary	47
Analysis	48
Coding Reliability	51
Results of Study One	51
Outcomes of the Speech Sampling Process	51
Gesture stroke durations	54
The Complexity of Gesture Forms	57
Spoken Perfective Aspect	62
Spoken Progressive Aspect	66
Spoken Durative Aspect	70
Summary of Gesture Profiles	74
Layering of aspectual views	74
Aspect and metaphor	75
Aspect-linked Gesture Features Occuring in Non-aspect-marked Contexts	76
Speaker choice of aspectual view	77
Discussion	78
Stroke durations	79
The Durative Difference	80
The Nature of Iconicity in Gesture	82
The Possibility of Indexing View of Event in Gesture Alone	84
Linguistic Aspect as a Heterogenous Conceptual Category	86

4. STUDY TWO: UTTERANCE STRUCTURE AND GESTURE	91
Introduction	91
Points From Current Analyses of Mandarin Syntax	97
The Nature of Mandarin Topicalized Utterances	100
Topic-Comment Constructions	104
The <i>Ba</i> -construction	109
Summary	112
Methods for Study Two	113
Participants	113
Elicitation	114
Speech Sampling	115
Gesture Sampling	117
Sampling Totals	117
Gesture Coding	118
Coding Summary	123
Analysis	123
Coding Reliability	127
Results of Study Two	128
Outcomes of the Speech Sampling Process	128
Gesture-to-Utterance Ratio	132
Gestural Encoding of Event Components	135
Single-Gesture versus Multi-Gesture Utterance Accompaniments	145
Encoding of Event Components and Case Roles in Gesture.	150
Summary of Findings on Rate of Gesture and Gesture Semantic	
Content	172
Other Findings	176
Speech-Gesture Synchrony	177
Speech-Gesture Timing and Semantic Content in Mandarin	
non-SV(O) utterances	177
The Position of Gesture within the Production Envelope	180
Repetition in Gesture	182
Discussion	186
Summary and Conclusions	191

The Status of the Results of Study Two	194
Future Research: An Alternative Use of Narration Data	198
5. SUMMARY AND CONCLUSIONS	201
REFERENCES	205
APPENDIX A — Aspect Sample	210
Mandarin Aspect Sample	210
English Aspect Sample	223
APPENDIX B — Utterances Contributed by Each Speaker to Each Target	
Aspect Category	228
APPENDIX C — Sentence Structures	229
Mandarin Sentence Structures Sample	229
English Sentence Structures Sample	239
APPENDIX D — Utterances Contributed by Each Speaker to Each Target	
Utterance Category	244
APPENDIX E — Number of Utterances with Specific Gestural Accompaniments:	
Single versus Multiple Gestures; Isolating versus Incorporating Gestures	245

LIST OF FIGURES

Figure

1. Comparison of aspectual viewpoints. 34
2. Individual 95 Percent Confidence Intervals for Mean Stroke Duration,
Based on Pooled Standard Deviation. 57
3. Individual 95 Percent Confidence Intervals for Mean Gestures per Utterance,
Based on Pooled Standard Deviation. 134

LIST OF TABLES

Table

1. Summary of coding categories in speech and gesture — Study One	48
2. Mean durations in seconds, and standard deviations, of gesture strokes accompanying utterances with perfective versus imperfective aspect; Mandarin versus English	55
2a. ANOVA table for gesture stroke duration; Mandarin versus English, perfective versus imperfective aspects	55
3. Mean gesture stroke durations in seconds and standard deviations: progressive versus durative aspect; Mandarin versus English	56
3a. ANOVA table for gesture stroke duration; Mandarin versus English, the imperfective aspects: progressive versus durative	56
3b. ANOVA table for gesture stroke duration; Mandarin and English together: imperfective progressive versus imperfective durative aspects	57
4. Gesture durations and complexity across different spoken aspect contexts; Mandarin versus English	59
4a. Hand(s) used to gesture across different spoken aspect contexts; Mandarin versus English	63
4b. Expressions of motion events only — gestural encoding of event compo- nents across different spoken aspect contexts; Mandarin versus English . .	65
5. Summary of coding categories in speech and gesture — Study Two	123
6. Mean number of gestures per utterance (and standard deviation), all utterance structures; Mandarin versus English	132
6a. Mean number of gestures per utterance (and standard deviation), by utterance structure; Mandarin versus English	133
6b. ANOVA table for gestures per utterance by utterance structure in Mandarin. .	134
7. Proportion of utterances accompanied by one gesture versus multiple gestures; Mandarin versus English	135
8. Isolating versus Incorporating gestures as a proportion of total gestures by language; Mandarin versus English, all utterance structures	136

8a. Element Only, Action Only and Incorporating gestures as a proportion of gestures by language; Mandarin versus English, all utterance structures. .	137
8b. Isolating (Element Only, Action Only) and Incorporating (distinguishing pantomimic enactment) gestures as a proportion of total gestures within utterance structure category; Mandarin versus English	138
8c. The utterance structure categories of Mandarin, in descending order, the proportions of Element Only gestures within the category of Isolating gestures	142
8d. The utterance structure categories of English: In descending order, the proportions of Element Only gestures within the category of Isolating gestures	144
9. The Isolating and Incorporating gestural accompaniments of all 80 utterances of each language, enactments included, all utterance categories; Mandarin versus English	146
9a. The Isolating versus Incorporating gestural accompaniments of all 80 utterances of each language in single versus multi-gesture accompaniments, enactments included, all utterance categories; Mandarin versus English	147
9b. Action Only, Element Only and Incorporating gestural accompaniments of all 80 utterances of each language, in single versus multi-gesture accompaniments, enactments included, all utterance categories; Mandarin versus English	148
9c. The Isolating versus Incorporating gestural accompaniments of all 80 utterances of each language in single versus multi-gesture accompaniments, enactments included, by utterance category; Mandarin versus English	149
10. Proportions of gestures marked for particular event components, by utterance category; Mandarin versus English; enactments and non-enactments separated	152
10a. Proportions of Isolating versus Incorporating gestures marked for particular event components, by utterance category; Mandarin versus English; pantomimic enactments excluded	157
11. Proportions of gestural event component encoding, including the contributions of enactments, by utterance category; Mandarin versus English . . .	161
12. Main points from the results of the planned analyses of Study Two; Mandarin versus English	173
13. The semantic contents of the TOPIC statements and their accompanying gestures in Mandarin TOPIC-COMMENT utterances	177

LIST OF ABBREVIATIONS

BA	Topic-marking particle/discourse particle (“ba”/吧)
-CL	Noun classifier, or ‘measure word’
CMPL	Completive verb particle
DUR	Durative aspect marker (“-zhe”/著)
GP	‘Growth Point’
IP	Information Processing
MA	Topic-marking particle/discourse particle (“ma”/嘛)
-MOD	Adjectival, possessive, or adverbial modifier (“de”/的 or 得)
NE	Topic-marking particle/discourse particle (“ne”/呢)
OBJ	Object-fronting particle (“ba”/把)
ONOM	Onomatopoetic form
PAS	‘Passive voice’ marker (“bei”/被)
PRF	Perfective aspect marker (“-le”/了)
PRG	Progressive aspect marker (“zai”/在)
TOP	Topic marker: any of the ‘pause particles’

CHAPTER 1

INTRODUCTION

A Paradigm for the Analysis of Speech-Accompanying Gesture

The psycholinguistic studies of the spontaneous gestures that accompany speech summarized in McNeill (1992) have shown that analysis of such gesture gives an expanded view of the conceptual representations that participate in speech production; one that complements, enlarges upon, and can in some particulars contradict a view based on speech alone. These studies support the thesis, given in McNeill, that gesture emerges directly from the same underlying unit of thinking as speech, without undergoing the linear segmentation required by the socially-constituted linguistic code. Analytic study of speech-accompanying gesture leads to the view that conceptual representations have idiosyncratic and holistic dimensions of patterning just as gestures do. For, when we compare the gestures of different speakers attempting to convey similar meanings, although their gestures will have sufficient features in common to allow the analyst to be confident that the speakers are representing similar event content, these gestures will also reliably differ from speaker to speaker at least somewhat as a result of individual speakers' contextualized handling of the meanings in their own narrative creations.

Research (McNeill, 1992) has shown that speakers of diverse languages and cultures, for example, English, Italian, Japanese, Georgian, Spanish, Russian, Swahili, Turkana, Chinese, and more, spontaneously gesture when speaking. All gesture systems across languages appear to accomodate both intentionally-produced, as well as unwitting,

gesture. Speakers of these languages perform gestures that combine with speech in broadly similar ways for the same referential content. Further, such gesture has been found to function similarly across these languages in important ways. It serves, together with speech, to propel communication forward by marking those elements in evolving conceptual representations that most contrast with preceding elements.

Toward a Model of Language Production

Traditionally, the vocal speech stream has been the focus of studies of human communication. Although sequential, segmented speech code is but one facet of spoken language production (i.e., that part of language that may be transcribed in written form), yet it has been the primary, often the only, input to the development of models of the human communicative competence that supports language comprehension and language production (Levelt, 1989; Bates and McWhinney, 1991). Yet human communicative exchanges are complexly structured, contextualized, evolving, backward- and forward-referring events. Models of human communicative competence must capture patterning at other levels of linguistic and conceptual analysis if they are adequately to describe the way communicative events evolve, and what parameters are in play from moment to moment in real-time language production. The speech stream alone, given its many conventionalized and selective features, necessarily does not reflect many dimensions of whole communicative events. The gestures that accompany speech demonstrably have dimensions of patterning that are more or less unencumbered by categorial constraints on form and reference. In contrast, the speech code, as analyzed within the paradigm established by Saussure (1959 [1916]), is a system whose architecture depends essentially upon categorially organized contrasts that have their own existence separate from their use at the moment of speaking.

In contrast to socially-constituted speech code, unwitting gestures are ‘locally created.’¹ They are what McNeill has termed ‘idiosyncratic’ to the speaker, in that they have dimensions of meaning whose values are established within context, and only at the moment of speaking. This does not exclude the fact that, across cultures, varying norms for gesture are attested that give wider or narrower scope to the non-Saussurian dimensions of gestural patterning (Havilland, 1993; Kendon, 1995a); that is, there are clearly socially-constituted dimensions of gesture patterning. However, some dimensions of gesture production differ from speech code in having a non-Saussurian semiotic architecture. This means that gesture can provide an often densely-encoded source of data on linguistic and conceptual patterning at multiple levels, that is different in kind from that provided by the speech code.²

Bates and MacWhinney (1991) note that a model of speech production must “account for the process by which native speakers select a set of expressions to convey meaning” in the context of real-time language use, and that the model “must have crosslinguistic generality.” Speech-accompanying gestures richly encode many facets of the representations that participate in spoken communication in such a way that the representations are transposed to the hands and gesture space of speakers. Dimensions of representations not encoded by speech thus become visible and analyzable. By comparing the communicative speech-gesture productions of native speakers of different

¹This is Adam Kendon’s term, although his own use of it may emphasize the socially-constituted dimensions that such local creations can manifest (Kendon, 1995).

²The attempt here is only to distinguish two frameworks of semiotic representation, one that is analytic and categorial (‘Saussurian’); one that is synthetic and holistic. The intention is not to identify the vocal modality exclusively with the former and the gestural modality exclusively with the latter. Such a distinction would be artificial in an important sense. I thank Karl-Erik McCullough for a large part of my awareness of the fact that much in spoken expression patterns according to the synthetic/holistic framework. Adam Kendon’s work (1995b and elsewhere) makes clear that much in the gestural modality can pattern within the analytic/categorial dimension. These facts are noted, but will not be pursued further here.

languages, it becomes possible to determine whether thinking is structured differently for speech production in these languages, or if thinking differs in terms of which features of events dominate in conception. Including speech-accompanying gesture in the analysis gives access to patterning in these dimensions, whether or not they achieve expression in speech itself. The assumption underlying the methods of elicitation and analysis used in the studies presented here is that gesture gives evidence of how speakers of different languages negotiate with the formal structures of their languages, when they are attempting to convey similar meanings. As such, the analytic study of speech-accompanying gesture is a tool to investigate the conceptual processes and representations that interact with linguistic representations on-line during real-time language production, cross-language.

Gesture and ‘Thinking-for-Speaking’

This is a method for examining Slobin's (1987; in press, b) proposed zone of interaction between thought and linguistic patterning that he calls ‘thinking-for-speaking.’ This is the structuring of thought into forms appropriate for spoken language production. “‘Thinking for speaking’ involves picking those characteristics that (a) fit some conceptualization of [an] event and (b) are readily encodable in the language.” Languages differ in how they encode, or do not encode, the components of events, and in how they distribute such components of meaning among the constituents of utterances. To the extent that this is so, it is conceivable that the conceptual representations that participate in real-time speech production are differently configured across different languages.

Slobin’s own formulation for discovering such differing configurations, if they exist, carries with it a risk of circular reasoning, in that it attempts to capture different patterns of thinking-for-speaking through examination only of speech. In cross-language

research of this type, the risk lies in depending solely upon the categories and mechanisms of the speech code to infer the character of speech-accompanying conceptualization. Since gesture has dimensions of patterning that are not constrained in the same ways speech code is, incorporating the analysis of gesture into the exercise is a means to circumvent this limitation. To see why this is true, we can consider the differences between the Spanish and English languages that Slobin (in press, a) has examined to find evidence of different patterns of thinking-for-speaking in those languages.

The basis for Slobin's comparison derives from Talmy's (1975, 1985, 1991) lexical semantic analysis of motion verbs in the two languages. According to Talmy, (see also Aske, 1989; Choi and Bowerman, 1991) the components of motion events, for example, MOTION, PATH, MANNER, FIGURE, and GROUND, are differently conflated in verbs of the two languages. Talmy proposes that in the verbs of Spanish, MOTION is typically conflated with PATH, as in the verb 'entrar'/'to enter.' In Spanish sentences, this leaves MANNER of motion to be specified, if at all, in an adjunct to the verb. According to Slobin, the Spanish language possesses a "tiny lexicon" of verbs that encode MANNER of motion. Further, in many syntactic contexts it is cumbersome to express MANNER of motion within the restrictions imposed by that language's system of contrasts. English, in contrast, has a large inventory of MANNER-encoding verbs. According to Talmy's analysis, speakers of English typically express the PATH component of motion events, not in the main verb itself, but in 'satellites' to the verb; for example, 'roll down,' where the 'down' is such a particle satellite. These features of English make it easy to build expressions that encode complex PATHs and MANNER. Slobin compares Spanish and English story narrations, as well as motion event

descriptions drawn from translations of novels from each of these languages into the other. He shows that Spanish narrative tends overall to show less encoding of MANNER and less encoding of the segments of complex PATHs.

On the basis of such ‘speech code’-based evidence, Slobin is led to conclude that thinking-for-speaking is different in the two languages, and as well that learning to think-for-speaking in different languages likely has a developmental trajectory. He speculates that the end state of the developmental process in mature native speakers is likely to be some kind of linguistic relativity of thinking, if only in *speaking*-related thinking. He states, “learning a language as a child constrains one’s sensitivity to what Sapir called ‘the possible contents of experience as expressed in linguistic terms’” (in press, b).

McNeill and Duncan (in preparation), however, examine the gestures that accompany the spoken expression of motion events in Spanish and demonstrate that Slobin’s assertions about thinking-for-speaking, and the development of it in the individual, must be refined. They give evidence from videotaped cartoon narrations by adult native Spanish speakers.³ In these narrations, one finds instances where *gesture* clearly expresses MANNER and complex PATHs that are nowhere encoded in the accompanying spoken Spanish utterances. For example, one speaker says in Spanish only that, “(the cat) gets in through the drainpipe, right? enters (through) the drainpipe,” while gesturally expressing that the cartoon cat enters and ascends with a climbing MANNER of motion. The speaker expresses this MANNER with hand movements that

³The example described is from a narration by a native Spanish speaker who had little competence in English. The speaker is re-telling the same cartoon as described in the studies presented here, but as part of another study by Gale Grubman Stam.

resemble a cat's paws climbing upward.⁴

It thus cannot be fully correct to say that the nature of the impact of long-term experience with the possibilities for expression in one's native language on one's acquired *habitual* patterns of thinking-for-speaking is that it constrains one's sensitivity to different parts of experience; that it reduces a speaker's attentiveness to those features of events that are underspecified or difficult to accommodate in the linguistic code. At least this is not always the case, as the example just given shows. The Spanish speaker of the example was clearly sensitive to, and evidently encoded, the motion event component of MANNER as it was presented in the eliciting cartoon stimulus. The repetitive, MANNER-expressive movement in the speaker's gesture permeated the gestural side of her speech-gesture production. The MANNER-expressive gesture began prior to the excerpted speech portion, and continued after it. Throughout the extended spoken production, there were no speech constituents co-referential with the MANNER expressed in gesture.

This Spanish example makes clear that native speakers can attend to and recall components of motion events that they tend not to encode in their speech. In English language production, the components of MANNER as well as PATH are regularly encoded in speech, whereas in Spanish, MANNER may be represented routinely instead in the holistic-synthetic framework, expressed via gesture in the cited example. If we rely solely on the units of speech code and the syntactic structures chosen by speakers in such cases to infer the character of the thinking-for-speaking in different languages, we will certainly err in determining whether and how, "each native language trains its native

⁴The complete Spanish speech segment is, "<eh> entonces busca la manera / de entrar / se mete por el desagüe / si? / desagüe / entra" ("um / so he looks for the way to get in / he gets in through the water pipe, right? / water pipe / he enters").

speakers to pay different kinds of attention to events and experiences when talking about them” (Slobin, in press, b).

Incorporating gesture into such analyses has the effect of reframing the linguistic relativity issue that Slobin’s proposal of thinking-for-speaking brings along with it, in the following way. What may be the effect, if any, of long term thinking-for-speaking across languages whose analytic-categorial frameworks of linguistic representation (their speech codes) target, ignore, or frame particular components of motion events in different ways? The example given shows that MANNER of motion, while clearly attended to by the Spanish speaker, was not linguistically *categorized* as such in the course of speaking about the event. Because the language neither requires, nor apparently makes it easy for her to do so, a Spanish speaker will not routinely bring MANNER-of-motion information into the analytic-categorial framework of representation when speaking (Slobin, in press, a). It requires empirical investigation, but it may be that in speaking situations in Spanish, the majority of representations of MANNER will be in the holistic-synthetic framework. These representations could be evidenced in gestures or in onomatopoeic vocal forms. In regard to the development of habitual thinking-for-speaking over a lifetime of speaking a given language, the issue thus becomes one of the long-term effects of having a particular portion of the ‘contents of experience’ routinely represented within a holistic-synthetic framework of representation, while other portions are routinely represented in an analytic-categorial framework.

Examining speech together with its co-occurring spontaneous gesture is a means to study the outcome in adult speakers of this long-term partitioning of experience within a significant cognitive-behavioral domain, speaking. Comparative cross-language analysis of speech-accompanying gesture thus is one means to examine the interface between thought and language; between thinking and speaking. Its advantage over

examination of speech alone is that it permits a different angle of view on the structure, content, and real-time evolution of conceptual representations that is distinct from, but closely coordinated with, the angle given by the various possibilities for spoken expression across languages.

Speech-Gesture Timing and the ‘Growth Point’ of Utterance

Gesture is “so intimately integrated into the production of the spoken utterance that its planning and organization must proceed simultaneously with, if not in advance of, the production of speech itself” (Kendon, 1980). In fact, it has long been recognized that unreflective, representational gestures of the sort analyzed for the studies reported here, almost always have their onsets a bit prior to the onsets of their co-expressive speech (Butterworth and Beattie, 1978; Butterworth and Hadar, 1989; Morrel-Samuels and Krauss, 1992). This advance interval is typically extremely short, but is instrumentally determinable. The exact timing relationship between speech and gesture onsets is not generally perceptible when viewing real-time language production. It is necessary to watch video recordings of these productions in slow motion, with sound, to see exactly how gestures time with their accompanying speech. The preparation phase of a gesture may precede the segment of speech with which the gesture is co-expressive by a fair margin. The margin by which the ‘stroke’ phase⁵ of such gestures precedes the speech with which it shares a semantic relationship, though, is typically so small that the timing relationship between productions in the two modalities is virtually synchronous. It is also frequently the case in contextualized, motivated discourse, that the stroke phase of gesture occurs at or very near the portion of the speech stream that is intonationally emphasized (Kendon, 1972, 1980). In this way, one portion of a spoken utterance is

⁵The phases of gesture production are explained in the General Methods section of Chapter Two.

typically singled out by both gesture and intonational emphasis.

According to McNeill (1992), this synchrony among the elements of speech and gesture production in continuous, motivated discourse reveals the nature and location of the ‘growth points’ (GP) of utterances. The GP is an hypothesized unit of analysis in language production that comprises both non-discrete conceptual representations of a sort involved in holistic, or visuo-spatial thinking and discourse planning, together with discrete, categorial units of speech code. In the GP, by hypothesis, expressive material patterned according to these two frameworks of representation exists in an integrated state in which each is completed by the other.

McNeill has hypothesized that GPs are where the ‘new information’ in discourse is differentiated from a background of established information. The proposed GP is similar to Vygotsky's (1987 [1930's]) ‘psychological predicate’ (see also Zinchenko, 1985). That is, the GP constitutes, “the novel departure of thinking from the presupposed background, and the minimal unit of verbal thinking that retains the essential properties of the linguistic and imagistic whole of a thought” (McNeill, 1992:220). This is expressive material that contrasts with what has gone before, or that contrasts with what is presupposed in the discourse, and that will itself, in turn, become part of the background against which subsequent content is differentiated. McNeill and Duncan (in preparation) extend this such that GPs may be any element of discourse that achieves definition against such a background. The examination presented here of speech-gesture timing as it relates to topicalization processes in Mandarin, will show that speech-gesture combinations may index *non*-novel information in the discourse as well; that is, there is potential for the GPs of utterances cohesively to index ‘given’ information in discourse, in the service of setting up a domain of reference for utterances to follow.

The stroke phase of gesture is its semiotically-valued portion. Production in the speech and gesture modalities is very precisely coordinated. A gesture's stroke phase, together with the unit or units in the speech stream with which it shares a semantic and a timing relationship, are taken to display the GP of a communicative 'production envelope.' The temporally extended surface production envelope takes shape in time as linearly-segmented speech code is generated through a process conceived of as 'unpacking' the material contained in the irreducibly dual-natured GP. The facts of coordinated speech-gesture co-occurrence, and the expression of related content across the two modalities, are evidence for a fused duality of representational frameworks driving communicative productions. Such facts justify the examination of speech-accompanying gesture for clues as to how communicative productions evolve. This evolution may be viewed only imperfectly via examination of the set of forced choices prompted by the speech code. As McNeill (1992:251) points out:

"Language provides conventional packages of meaning that one cannot avoid. Marking one feature of meaning often obliges us to mark another. Precisely because gestures are not constrained by language-like conventions, the gesture needs to differentiate only those aspects of meaning that are significant. ... Language is a system of regulated contrasts; gestures are outside this regulation."⁶

⁶Compare this with Kendon (1995b), who clearly demonstrates that not all categories of gesture, nor all dimensions of gesture, nor all gesturing across cultures, fall outside a 'system of regulated contrasts.' That is, within particular cultures, speech-gesture dialect groups, gesture registers, and/or particular 'gesture systems' there exist gestural forms or 'morphemes' that pattern according to systems of categorial oppositions, such that ill-formed or non-meaningful versions of these forms are possible; as well there are dimensions of gesture, for example deictic/orientational dimensions (see also Havilland, 1993) that in some speech-gesture dialect groups are highly communicatively loaded. Gesture production in such dimensions is received by a listener in the speech-gesture dialect group as part of the explicit information meant to be communicated by the speaker. Such forms, regardless of where they appear in speech-gesture production, or whether they are produced singly or in combination with other forms, have stable meanings, recognizable and meaningful to conversation participants within that system.

A model of speech production that is founded on the notion like the hypothesized GP will in theory be able to constrain and account for the direction taken from one production unit to the next in discourse, thus satisfying Bates and MacWhinney's requirement for an account of "the process by which native speakers select a set of expressions to convey meaning" in the context of real-time language use. Analysis of the GPs that actually form in different languages has the potential to show how variation cross-language in lexico-syntactic resources interacts with the on-line creation of communicative packages in language production, and so fulfills these authors' further requirement that such an account have cross-linguistic generality. As language production progresses from GP to GP, at each step of progress, the number of narrative moves is limited that: (i) differentiate new from given information in a sufficiently contextualized fashion, (ii) satisfy a given language's requirements for linguistic form, and (iii) shorten the distance between the current narrative location and the narrative goal.

Current information-processing (IP) theoretic models of the speech production process such as that set forth in the Levelt (1989) volume entitled *Speaking* are being extended (Hadar and Butterworth, 1992; Krauss, et al., 1991) in an attempt to model the production of speech-accompanying gesture. Such models, however, cannot account for many phenomena of speech-gesture timing and gesture form that will be described below, including gestural holds, self-interruptions of both speech and its accompanying gesture (Kita, 1993). Neither can such models account for the way discourse context, on the one hand, influences the semantic and timing relationships among gesture, prosody, and units of speech code, and on the other hand promotes or constrains speakers' choices among structural options for expression. Gigerenzer and Goldstein (1994) call such weaknesses the "Achilles heel" of modularist IP models of language production. Within

the IP theoretic framework, phenomena such as gestural holds, for instance, are puzzling and it is not possible to treat their potential significance as clues to the organization of thinking-for-speaking. The descriptive framework of GP theory, on the other hand, does provide a means to do so, incorporating as well the context of communicative productions.

CHAPTER 2

TWO STUDIES OF THE RELATIONSHIP BETWEEN LINGUISTIC FORM AND CONCURRENT CONCEPTUALIZATION

The research presented here is an application of the concepts of gesture analysis outlined in the Introduction. It looks at the relationship between linguistic structure and on-line conceptual representations during language production, by comparing the speech and gesture of native speakers of typologically different languages from two unrelated language families, Mandarin Chinese and English. The speech and gesture of these two groups of speakers are compared in relation to two areas of linguistic form in which these languages have been said to differ, and that have been the focus of much linguistic analysis. These are verb aspect and the typological feature of topic- versus subject-prominence. In essence, both studies were exploratory in character. The tools of gesture analysis were used as a ‘naturalist’ would use them to find out, where languages differ structurally and in their expressive options, whether gesture differs systematically as well. The possibility that gesture would not differ observably was always open. Either way, the patterning of gesture in relation to structural differences among languages merits description in and of itself, and we may consider such patterning as giving clues concerning the process of thinking-for-speaking.

One motivation for the studies to be described here was the character of the chosen target linguistic differences themselves. In each case it seemed reasonable to hypothesize that speakers’ thinking may be influenced by variations in these particular

domains of expression, in ways that would be detectable by gesture analysis. Previous research with speech and gesture data from the same two speaker groups was another motivation. That work yielded the impression that systematic gestural differences related to the target linguistic differences would in fact be found between Mandarin and English speakers.

Study One examines the spoken expression of linguistic verb aspect and *Aktionsart* in the two languages, together with the gestures that occur in aspect- or *Aktionsart*-marked spoken contexts. Aspect and *Aktionsart* are both terms for linguistic mechanisms used to express the temporal contour of events. These terms, and how they are analytically distinct, will be further defined in Chapter Three. The goals of this study are two. The first is to determine whether there are features of gesture performance that covary with distinctions in verb aspect and *Aktionsart* in co-occurring speech, within-language. The second goal is to determine whether between-languages, given the differences in how Mandarin and English linguistically convey distinctions of aspect and *Aktionsart*, and in how verb aspect may combine with lexical *Aktionsart*, the gestures of the two speaker groups also differ, or whether they are the same. The potential significance of any similarities or differences found between the two speaker groups lies in what they can tell us about the nature of aspect and *Aktionsart* as conceptual categories and how these categories are related to the means of spoken expression.

Study One takes its cue from the fact that, according to Binnick (1991:446), Mandarin is a “classic tenseless language,” but marks a range of aspectual distinctions both grammatically and periphrastically. In lieu of tense marking, Mandarin locates events in time relative to the present by means of temporal adverbials and explicit temporal references. In contrast to Mandarin, English has been seen as relatively impoverished when it comes to marking verb aspect. English grammatically marks tense

and some aspect, but the two systems are confounded. The expression of the aspectual view of an event in spoken English depends more on lexically- and periphrastically-encoded *Aktionsart* than it does in Mandarin, which has a set of grammatical morphemes marking the major aspectual distinctions. Study One focuses on three spoken aspect distinctions that may be overtly grammatically marked and that occur with high frequency in Mandarin narrations. Each of these aspectual views may be signalled by a single-syllable morpheme: progressive ZAI, durative -ZHE, and perfective -LE. The gestures that co-occur with these markers in Mandarin and in comparable English speech contexts are examined and compared. These particular three verb aspects were chosen because they are sufficiently frequent in the narrations examined here to accumulate reasonably large data sets for analysis, not because they were the only aspectual distinctions expressed by the speakers.

Study Two is inspired by linguistic typological research concerning the feature of sentence or discourse structure known as ‘topic-prominence’ (Comrie, 1985; Li and Thompson, 1976, 1981; Tsao, 1990). According to this research, Mandarin, a topic-prominent language, differs from a subject-prominent language like English in the management of grammatical TOPIC. Mandarin, though regarded as being fundamentally a ‘SVO’ (Subject-Verb-Object) language, is described as having several high-frequency utterance structures that cause the ordering of reference and action in sentences to diverge from this ‘standard’ SVO structure. In most analyses, its topic-prominent nature is said to be evident in the relative typicality and frequency of utterances erected on a ‘TOPIC-COMMENT,’ as opposed to a SVO, sentence plan. This is an utterance structure that makes a TOPIC out of one of the constituents of a SV(O) utterance, or out of a constituent structurally unrelated to any one of them, and the utterance is completed with a

COMMENT to that TOPIC. Grammatically topicalized utterances are an option for speakers of English as well; however, such utterances are thought to be rather marked uses, perhaps identified with particular registers. The English speakers sampled for this study produced almost no grammatically topicalized utterances.

Reliably and significantly differing patterns of information flow in the surface structure of utterances may coordinate with quite different organizational requirements during thinking-for-speaking. Accordingly, the gestures of Mandarin and English speakers are examined for evidence of different thinking-organizational strategies related to different basic utterance structures.

General Methods

This section covers those aspects of the Method that are applicable to both Studies One and Two. Description of the details particular to either study will be found in Chapters Three and Four, respectively.

Participants

The participants were fourteen native Mandarin speakers (ten female and four male) and eleven native English speakers (six female and five male), ages 18 to 55. Not all of these participants were sampled equally for each of the two studies, as will be described in the Methods sections for the studies, and as is recorded in the Appendices. Of the Chinese participants, seven are from various regions of The People's Republic of China (Mainland China) and seven are from The Republic of China (Taiwan). The data for most of the seven Mainland participants were collected within a year of those participants' arrival in the United States. Six of these participants were University of Chicago graduate students, the other one a visiting scholar at the university. Four of the Taiwanese participants were young, middle-class mothers of small children filmed in

Taipei, Taiwan; the remainder, all students, were filmed in the U.S. and Canada. For the purposes of the present research, it was assumed that in regard to the grammatical features of aspect marking and lexical *Aktionsart*, and the typological feature of topic prominence, all of the Mandarin-speaking participants were linguistically comparable. Even though some of them were first-language speakers of some Chinese language other than Mandarin, all had spoken Mandarin from a very young age. All education from pre-school and up on both the Mainland and in Taiwan is carried out in that language, and it is also the language of business, government, and public life.

The participants were not informed that their gestures were an object of study.

Stimuli

Three different types of stimuli were used to elicit speech and gesture:

1) Cartoon: This is an American cartoon of a classic and well-known type. It features the cat and bird named “Sylvester” and “Tweety” and is about six minutes long. The cartoon’s plot is limited. It consists of eight action-filled episodes in which Sylvester the cat uses different methods to try to capture Tweety. Though produced in the U.S., there is little dialogue in the cartoon, so little English language. It has been shown to speakers of many different languages who have little or no knowledge of English. When these speakers re-tell the story of the cartoon, it is clear that their comprehension of the story line and the events portrayed is good.

2) Vignettes: These are a series of very short videotaped action sequences involving small plastic characters or inanimate objects. A set of 65 vignettes was used for this study. Each is around one and a half to two seconds long and depicts a character or object performing one or two actions, sometimes involving a second, stationary, object. The vignettes are part of a test battery designed to elicit morphological marking on verbs of motion in speakers of American Sign Language (Supalla, Newport,

Singleton, Supalla, Metlay & Coulter, 1993). The vignettes are designed to be viewed singly in a series, with time after each one for the participant to describe what she has just seen.

3) Movie: This is a very early feature film by Alfred Hitchcock; one hour and twenty-four minutes in length. Originally conceived as a silent film, it subsequently was modified to include sound and some short stretches of dialog. In its modified state, however, it retains an emphasis on visually-structured narrative, with the result that the story line and characterizations are accessible to viewers who speak little or no English. The story involves only three main characters and a very simple plot line, but is psychologically and morally layered and ambiguous.

Procedure

Elicitation

Participants viewed the stimuli on videotape, and were then themselves videotaped describing what they had just seen to a same-native-language listener. In the case of the cartoon and movie stimuli, speakers narrated to interlocutors who, they were told, had not previously seen these stimuli. The speakers were instructed to be as complete as possible in their descriptions, so that the interlocutor would then be able to re-tell the story to a third person. Interlocutors were encouraged to interrupt with questions if any part of the speaker's narration was unclear. For this reason, the communicative context was essentially conversational, even though one person did most of the talking.

Seventeen of the speakers, eleven Chinese and seven American, viewed and re-told the story of the cartoon. Six speakers, four Chinese and two Americans, viewed the vignettes. Two Chinese and two Americans viewed and re-told the story of the movie. Although most of the Mainland Chinese participants were less familiar with the

characters and conventions of the American animated cartoon and the Hitchcock movie than were the American participants, their narrations overall were quite comparable to those of the Americans. Cartoon and movie re-tellings were all to listeners who were, or who pretended to be, unfamiliar with the stimulus. All of the speakers who performed the Vignettes Task, Chinese and American alike, reported finding the vignettes stimuli bizarre, but generated a response to each vignette nonetheless.

Speech transcription and gesture coding

From the videotaped narrations and Vignettes responses, detailed transcriptions of all utterances and the gestures that accompanied them were created. The speech for most of the Mandarin cartoon and movie narrations was transcribed first by a native speaker and later checked by a fluent non-native speaker of the language. The rest of the data was transcribed first by a fluent non-native speaker, and then checked for accuracy by a native speaker. The step of checking by a non-native speaker is necessary to correct for the ‘normalizing’-type errors to which native-speaker transcribers are prone. For instance, it was found that native-speaker transcribers sometimes corrected grammatical ‘errors’ and smoothed out instances of speech dysfluency.

The speech transcriptions include all breaths, pauses (both filled and unfilled), and speech dysfluencies such as self-interruptions, self-corrections, and repetitions. The features of grammatical structure pertinent to Studies One and Two were flagged. The gestures accompanying these structures were examined frame-by-frame on the videotapes and were coded for form, type, semantic content, and function in relation to the speech. ‘Representational’ gestures were flagged. Finally, the timing of the gesture production relative to the speech was exactly coded, to within-syllable accuracy. The bulk of gesture coding on which these analyses rest has been verified by more than one trained gesture coder. Further, for both Studies One and Two, a separate and more systematic coding

reliability check was done by another experienced coder with knowledge of Mandarin Chinese.⁷

Types of Gesture

The studies to be described here are based on analyses of gestures with certain characteristics. These gestures occur together with speech, and their production is spontaneous and unreflective; that is, there is no evidence that they are planned prior to the moment of their production. They appear to be largely unconsciously produced. Certainly, speakers produce a variety of gestures of which they and their listeners are consciously aware. Such ‘gestures that are meant to be seen’⁸ are produced expressly to assist in conveying specific idea content, for example, visuo-spatial content that is not easy to convey in speech, or for which the desired term cannot be retrieved. Gestures that are meant to be seen are often, though they need not be, ‘deictically framed’ in the accompanying speech, as for example when a speaker holds both hands, palms facing, in mirror configuration some distance apart with thumb and index finger extended, while saying, “a box like this.”⁹

Speech-accompanying gestures having iconic, metaphoric, or deictic representational value, or some combination of these, were coded for the analyses presented here. Such gesture is typically densely-encoding, so usually more than one of

⁷Karl-Erik McCullough.

⁸I owe this description to S. Kita.

⁹Other varieties of consciously-produced and intentionally communicative gestures, in addition to those discussed in the various Kendon references given throughout this paper, include those of trained public speakers; for instance politicians, trial lawyers, and television announcers, whose gestures give the impression of being practiced or ‘canned’ and seem unnatural precisely because they are planned as opposed to spontaneously produced. Speakers like these receive training in how to gesture.

these representational dimensions is present in a given gesture.¹⁰

Following McNeill, the term ‘iconic’ is used for the dimension of gesture that represents features of concrete referents to which a speaker refers; for instance, two hands held out from the speaker’s body, some distance apart, and with their flat palms facing each other, may represent a wooden plank by showing its length. Gestures termed ‘metaphoric’ represent features of metaphorically-construed abstract concepts. For instance, two hands may be held together, cupped palms up, in the form of a bowl. In the appropriate speech context, a gesture of this sort is analyzed as the gestural representation of a ‘conduit’ metaphor (Lakoff and Johnson, 1980; Grace, 1987); a metaphor according to which abstract, unbounded quantities are conceived of as things that may be unitized and held in containers. The item depicted by a metaphoric gesture like this, the ‘vehicle’ that is a bowl-like thing, is of course something that exists in concrete experience; only the communicative context in which it is produced allows us to infer its particular representational properties as a gestural metaphor. A metaphoric gesture of the ‘conduit’ sort, or any other similar example, has an iconic base, and so is considered to be ‘representational’ in the same way.

Iconic and metaphoric gestures represent by depicting. Deictic gestures represent by indexing locations in gesture space that have come to stand for those meanings. In form, deictics are most typically points with the index finger to locations around the speaker. There are ‘concrete’ deictics whose referents are actual physical objects or locations in the speaker’s environment, either visible, or distant and non-visible. The studies here however are concerned only with ‘abstract’ deictics that index locations in

¹⁰I thank Karl-Erik McCullough for persistently drawing my attention to the fact that gestures always pattern within a multi-dimensional featural encoding space, such that only the rare, extreme gesture may be classified with reference solely to a single representational dimension, such as concrete iconicity, or abstract deixis, or with reference to a single level of linguistic analysis, for example the narrative or metanarrative level of discourse.

gesture space that are invested with certain meanings; either given by the speaker in previous discourse, or derived from a presupposed deictic frame (*up*, *down*, and the like). Deictics are common in the elicited narrations analyzed here, and do not always conform to the *index point* handshape. Speakers ‘point’ with different hand shapes; a flat hand,¹¹ or with different parts of their bodies — the chin or head, for instance, and this varies across speaker groups. Different vehicles for the expression of deixis are not a factor in the analyses presented here. The presence of the dimension of deixis alone is sufficient for inclusion in the representational gestures that are examined here.

As was noted above, gestures typically pattern on multiple representational dimensions simultaneously. Thus, any iconic or metaphoric gesture will typically reference configurational gesture space as well, and so can be said to have a deictic dimension. That is, speakers spatially ‘map’ the relations among referents by producing them at particular, marked, locations in gesture space. An example of a sort of gesture that combines two representational dimensions in this way is what is called a ‘localizer.’ This is essentially a deictic gesture that is infused with some iconic value, such that it does more than just *point* to a location some distance away in gesture space; but rather, is positioned in the referenced location, and may encode as well something of the physical nature of the referent; its object nature, for instance. A slightly cupped hand, palm up, moved first to one location in gesture space and then to another, for instance, can differentiate two concrete or metaphorically-construed objects in that space.

The features of gesture form and timing on which the analyses here are based are assumed to be dissociable from strict social mediation and so to be for the most part locally-created. These gesture features acquire their semiotic values from the whole communicative contexts in which they occur. In contrast to speech code, they

¹¹Adam Kendon (1995a) has explored some of the discourse-pragmatic implications of the choice of hand shape used to express deixis.

automatically encode few if any specific, socially-constituted meanings, yet they show a clear semantic and structural relationship to on-going speech in form, content and timing. This relationship is locally determined. Excluded from these analyses therefore are what have been called ‘emblems,’ or, ‘quotable gestures.’ Such gestures are those whose meanings are largely socially-constituted. Some gestures of this type can be produced as stand-alone statements, for example, the familiar “OK” sign of thumb and forefinger held together in the shape of an “O,” with the remaining fingers extended. So stable are the meanings of such gesture forms, that it is possible to study the history of their use across speakers and eras, sometimes across periods of centuries, and develop dictionary-like lists of them (Di Jorio (1832), cited in Kendon, 1995b). Such gestures and their meanings are recognized by others of a speaker’s own culture.

Locally-created, ‘representational’ gestures with meanings particular to a given speaker and a given narrative content at the moment of speaking, as described by McNeill (1985, 1987, 1992) are the objects of study here. When different speakers are constrained to narrate the same story content, such gestures will observably encode similar meanings and therefore have a high likelihood of sharing some features of form; nevertheless, they will differ enough from speaker to speaker to make it clear that standards of form of the kind that constrain, for example, an ‘OK’ sign to a certain configuration and location in gesture space, are not in force in the formation of every feature of gestural performance.

Excluded as well from these analyses are gestures, whatever their nature, that occur in the absence of speech. Such gestures may often be configured by the speaker to take the place of speech, and so may differ in important but as yet poorly understood ways from those gestures that participate *with* speech in the construction of meanings.

The Mechanics of Unreflective, Representational Gesture

Following Kendon (1972) and McNeill (1992), gestures are taken to consist of three primary phases: (i) a preparation phase, when the gesturing hand moves from a rest position, (ii) a stroke phase, the main representational movement or stationary phase, and (iii) a retraction phase, when the hand returns to its rest position. These phases taken together make up one gesture 'phrase.' All representational gesture phrases comprise at least a stroke phase. It is the crucial phase. Kinesically, the stroke is very often the movement focus of the gesture. Semantically, it is where the meaning of the gesture is apparent. For the analyses presented here it is important to note that, while the stroke is often a movement phase, this is not always the case. Frequently, the stroke phase of a gesture may be a stationary, meaningful configuration of the hand or hands; for instance, when a speaker iconically represents a small bird by holding her hand up with thumb and forefinger extended two or three inches apart, keeping it still for some moments. Further, a stroke 'phase' may also be *instantaneous*, as is the case when the preparation phase of a gesture is followed immediately by its retraction phase, and all there is between the two phases is a momentary, meaningful configuration of the hands. In such a gesture, the instantaneous stroke phase may suggest the shape or outline of some object, for instance.

Either of the other primary gesture phases, the preparation and retraction phases, may be overridden in production; that is, the preparation for one gesture may follow immediately upon termination of the stroke phase of the preceding gesture, thus eliminating the first gesture's retraction phase. Similarly, a gesture stroke phase may onset immediately from the point where the preceding gesture left off, such that there is no preparation phase for the second gesture. Thus, it is possible and it often happens that gestures are executed very rapidly in succession, with no return to rest position in between. This means that it is easily possible for a gesture, or even more than one gesture, to occur with each of multiple short phrases spoken in rapid succession in fluent

speech.

To illustrate the features of representational gestures just described, below is an excerpt from an English-speaking subject whose speech is accompanied by a gesture. The speech with which the gesture coincides is surrounded with square brackets and the gesture stroke phase is in bold face type.

(1) I dunno, [she **slugs** him] or throws him out the window

In the bracketed gesture, the preparation phase coincides with the first portion of the word “she.” The stroke phase consists of a single, rapid movement of the speaker’s right hand in the form of a fist, starting at the speaker’s right and moving slightly downward to the left — the movement of striking something. As shown by the bold face type, the stroke starts on the final portion of “she” and extends to cover roughly the first half of the verb “slugs.” By the end of the word “him” the hand has returned to its rest position in the speaker’s lap, thus concluding the retraction phase of the gesture.

In addition, gestural holds sometimes occur during the course of a gesture, when the gesturing hand pauses momentarily and then resumes motion. The hand may pause between the preparation and stroke phases or between the stroke and retraction phases of a gesture, or both, as the following example demonstrates.

(2) an[d he swings **smack** into the b][uilding

In the bracketed gesture shown here, the underlined portions indicate where the pre-stroke and post-stroke gesture holds occurred, relative to speech. The gesture, a flat right hand with the fingers together, pointing away from the speaker’s body, and with a left-facing palm, holds for an instant coincident with the word “swings.” This is referred to as a ‘pre-stroke hold’ (Kita, 1990). The speaker’s hand then arcs left and downward across the gesture space in front of the speaker, then back upward and to the left. There

is another gestural hold after the stroke phase, and this is called a ‘post-stroke hold.’

The Analytic Significance of Gesture Phases

Kita (1990) observed that pre-stroke holds are often a phenomenon of the dynamic, synchronized timing relationship between speech and gesture, wherein the stroke phase of a gesture that is expressive of a particular semantic content seeks, as the evidence would suggest, to time with a specific target linguistic unit in the accompanying speech stream. As well, most often, stroke phases time with the point of prosodic emphasis in the speech they accompany (Kendon, 1972, 1980). Thus, within a speech-gesture production envelope there is a point of emphasis where the dimensions of language production that pattern within the analog, holistic-synthetic framework of representation, that is, gesture and intonation (Bolinger, 1986), focus in on particular units of the speech code.

The stroke phase is defined as that portion of a gesture where it assumes a semantically interpretable form. Gestures encode meanings related to the speech they accompany; gesture and speech are co-expressive. The semantic relationship between gesture strokes and their co-occurring speech units may be ‘co-referring’ (essentially, redundant), or ‘complementary,’ or often, some of each. The former relationship is exemplified by the following speech-gesture production in which a two-handed gesture that depicts a spherical object co-occurs with the phrase “bowling ball” in speech.

(3) Tweety throws a [**bowling ball**] down the drainpipe

The speech-gesture semantic relationship of ‘complementarity’ is exemplified in a Mandarin language example in which the the two hands alternately rotate around one another, moving on wrist pivots, in the space in front of the speaker, thus expressing a MANNER of motion, *rolling*. Instead of co-occurring with a MANNER-expressive

component of the motion event in speech, “rolls around,” the gesture co-occurs with the expression of the motion event’s FIGURE component, “cat,” as follows:

- (4) [mao ne] zai da jie shang <ehn> gun-dong¹²
 貓呢在大街上 <ehn> 滾動
 cat TOPIC on big street-surface <ehn> roll-move
the cat <um> rolls around on the street

The analytic significance of the difference between speech-gesture ‘redundancy’ and ‘complementarity’ during stroke phases is as yet poorly understood and will not be a particular focus here. These examples are given in order to illustrate that there is more than one sense in which speech and gesture may be thought of a ‘co-expressive.’

Since speech production is linear and sequential, it is unavoidably extended in time. Gesture, in contrast, can be prepared and executed as a synthetic whole. As Kendon (1980) noted, gesture planning and production usually precede speech production by a small interval of time. Thus, the preparation phase of a gesture typically onsets sometime *before* the portion of the spoken utterance occurs with which the stroke is destined to time. In such cases, the gesture often halts in mid-motion and does not advance to its stroke phase until the speaker-gesturer has come to the portion of speech that bears the necessary semantic relationship with the content encoded in the gesture stroke. An instance of this is seen in example (2) above, “swings **smack** into,” where the underlining prior to the bold face stroke phase shows the presence of a pre-stroke hold. In these cases, as Kita (1990) described it, it appears that “gesture waits for speech to catch up.”

¹²The ‘pin-yin’ transcription system, developed by Mainland Chinese linguists with the help of Soviet linguists in the early 1950’s, and officially adopted in the PRC in 1958, is used to transcribe all of the Mandarin examples. This conforms with common practice in Chinese linguistics publication and language teaching.

Pre-stroke holds play an important analytic role in GP theory. Theoretically, the presence of such a hold in a gesture phrase constitutes the clearest indication of the specific content of the GP of an utterance. In an instance like the following:

(5) he goe[s up inside the drainpi]pe

the accompanying gesture stroke, timing with “side the drain,” is a path trace with the tip of the index finger *up* across gesture space in front of the speaker. The verb-satellite assembly “goes up” has semantic content in common with the path trace gesture, but is pointedly (judging by the pre-stroke hold) skipped by the gesture’s stroke phase. According to McNeill (1992) and McNeill and Duncan (in preparation), the analytic significance of pre-stroke holds is that they make particularly clear what is *not* part of a GP. In a case like example (5), the inference is that the semantically rather bleached main verb “go” and its PATH satellite “up” together are a part of the final utterance that is derived from the GP during the process of its unpacking into overt sequential speech forms, rather than being an explicit part of the initial organizing unit itself.

Within this theoretical framework, the analytic significance of a *post*-stroke hold, however, is quite different. A post-stroke hold is when a speaker’s hand remains motionless in mid-air for a few moments, at the position where a stroke phase terminates. This is analyzed as another consequence of the differences of execution in speech and gesture, a result of their essential organizational dissimilarity: speech being constrained to extend sequentially, while gesture is capable of more instantaneous execution. Thus, whereas a pre-stroke hold indicates what is *not* part of the GP of the utterance, a post-stroke hold prior to the gesture’s retraction phase can give a clear indication of what speech material *is* part of the GP, in addition to that which timed with the gesture stroke

phase itself. That is, that gestures often continue to hold for a short interval in a marked location in gesture space, and typically with a marked handshape as well, is also, as with the pre-stroke hold, a phenomenon of the dynamic timing and semantic relationships between speech and gesture. Unlike a pre-stroke hold however, the post-stroke hold makes explicit what is directly a part of the GP.

The presence of a post-stroke hold may only be accurately determined when there is as well an observable retraction phase to the gesture or a marked re-setting of the gesturing hand prior to execution of the next gesture-speech combination. There are styles of speech-gesture production manifested in some speakers where gesture strokes are separated by holds, one following the other, with no obvious retraction phases or re-sets to interrupt the sequence. The existence of this style of speaking-gesturing for stretches in the narrations of some individuals in no way disconfirms the general speech-gesture timing principles and their significance stated above. It is just that in such stretches of speech-gesture production, where the hand rarely returns to rest position or re-sets, it is not possible to say for certain where post-stroke holds leave off and pre-stroke holds begin. In production sequences like this, that particular analytic indicator of the nature of GPs isn't visible. On the basis of the data from the studies presented here, additional roles for gestural holds will be suggested that are specifically related to the target grammatical differences between Mandarin and English that motivated this research.

CHAPTER 3

STUDY ONE: GESTURAL INDICES OF VERB ASPECT AND *AKTIONSART*

Introduction

First, the distinctions of linguistic aspect and *Aktionsart* that are the focus of this study are described and after this, how the Mandarin Chinese and English languages differ with respect to how each signals a particular aspectual frame or viewpoint in spoken expression.

While verb tense indexes the location of events in time (past, present or future), verb aspect, according to Comrie (1981a), indexes a particular view of an event in time. He states, “aspects are different ways of viewing the internal temporal constituency of a situation.” Many linguists (Comrie, 1981a; Binnick, 1991) hold that the most fundamental of aspectual distinctions exhibited across all languages that do mark aspect is that between the perfective and imperfective. Perfective aspect is said to express an external viewpoint of events, while imperfective expresses an internal viewpoint. A speaker’s choice of aspectual view represents a decision either to expand in expression on the internal temporal constituency of an event, that is, open a view onto the internal workings of an event, or *not* to expand on these workings, and instead deal with the event as a collapsed whole. Distinct from, but related to aspectual view as it is used here, is the notion of *Aktionsart*. As opposed to the speaker-applied viewpoint expressed in choice of linguistic verb aspect, *Aktionsart* is a feature of event type itself, as encoded in the verb or periphrastically. According to Klein (1994) *Aktionsart* refers to the, “inherent

temporal features of the lexical content of verbs, and more complex constructions.”

Particular verbs of every language express event types with certain inherent temporal features (Vendler, 1967). For instance, the English verb ‘hit’ expresses an event structure that is inherently punctual; the verb ‘reach’ (one sense of it) inherently telic; the verb ‘deteriorate’ inherently durational; the verb-particle combination ‘fill up’ inherently completive. Aspect and *Aktionsart* are thus distinct analytic categories and in fact can both be explicitly distinguished in a single utterance, as the example from one of the Mandarin narrations below demonstrates.

- (6) ta ting-dao le
 他聽到了
 he listen-CMPL PRF
 he hears this

In this example the ‘dao’ that follows the verb “ting”/“hear/listen” explicitly signals completive *Aktionsart*, creating a so-called ‘resultative verb complement’ (RVC) structure. The same *Aktionsart* distinction is captured in English in the contrasting lexical items ‘hear’ versus ‘listen.’ In Mandarin the contrast is expressed via presence versus absence of the marker ‘dao.’ The speaker expresses an external aspectual view of this event by choosing perfective aspect.

Note, however, that, though they are distinct analytic categories, inherent *Aktionsart* can strongly prejudice choice of aspectual view. Leaving the perfective aspect marker off of an utterance like Example (6) can make it sound somewhat anomalous to Mandarin ears. Correspondingly, imposing an internal view of the event ‘ting-dao’ would result in quite an anomalous-sounding construction.¹³ In this way, verb aspect and *Aktionsart*, although not the same analytically, intersect as linguistic categories in the

¹³Hui-fang Hong (personal communication).

sense that verbs expressive of events with particular *Aktionsart* often coordinate more or less naturally with one linguistic verb aspect as opposed to another. For instance, ‘deteriorate’ is inherently durational, and so in combination with imperfective verb aspect its inherent temporal structure is expressed. In contrast, taking an internal view of an event expressed by a punctual verb such as ‘hit’ requires an effort of imagination, perhaps aided by picturing what a film of a single instance of hitting would look like, played in slow motion. Simply marking a punctual English verb for progressive imperfective aspect, if nothing else is added to the construction, generates a reading of iterativity (Comrie, 1981a), a different *Aktionsart* from punctual, indicating multiple occurrences of a punctual event in sequence, as in, “he is hitting him.” To focus linguistic expression on the internal constituency of a single iteration of a punctual event, it is necessary to elaborate on this in a very explicit expression. Perhaps something like, “as he is hitting him the third time,” captures such an event contour.

Thus, for the purposes of this study, aspectual view and *Aktionsart* are understood as categories that interact in expression in ways that create different understandings of the internal temporal contour of an event. Aspectual view is the target conceptual category here, however. Overt grammatical aspect marking is given precedence in inferring the aspectual view adopted by a speaker, including in those cases where it combines with *Aktionsarten* of different types. Absent such overt marking, aspectual view is inferred on the basis of *Aktionsart*.

The present study attempts to discover whether the distinction between perfective and imperfective aspects in speech is related to distinctions in the execution of gestures. A further distinction is made within the imperfective, or internal, view of events between the verb aspects progressive and durative. Although the latter terms for imperfective aspect are sometimes used interchangeably, the present analysis assumes a distinction

between them as characterized in Binnick (1991). Binnick adopts the framework in which progressive is described as momentary, and durative is described as having temporal extent. Binnick illustrates this contrast between the two imperfective aspects with the following English sentences:

- (7) a. the comet is coming (Progressive)
 b. the comet comes ever nearer (Durative)

Figure 1 is a timeline that illustrates the distinctions between the perfective and the two imperfective aspects. Progressive aspect ('be looking') situates the speaker at a dimensionless point somewhere in the course of an action in progress. Durative ('while looking') situates the speaker in the course of an action explicitly encoded as having temporal extent. Perfective aspect ('looked') situates the speaker outside the event, viewing it as a complete entity.

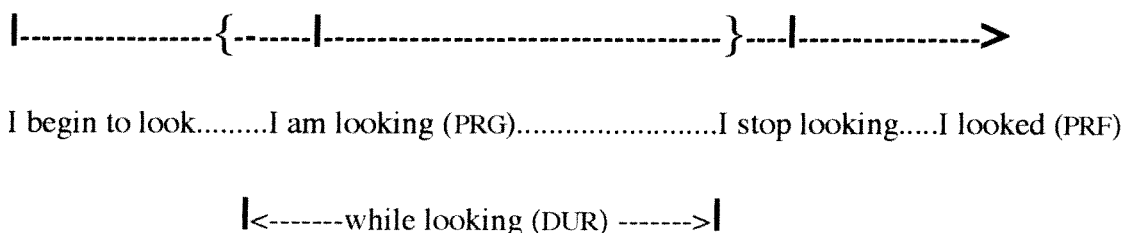


Figure 1. Comparison of aspectual viewpoints.

Linguistic Means of Conveying Aspectual View

As mentioned above, Mandarin makes available a set of three single-syllable morphemes for the unambiguous expression of these aspectual views. In Mandarin, progressive ZAI is always before the verb, as in the following example:

- (8) ta ZAI pao
 他 在 跑
 he PRG run
 he is running

Additional material may or may not intervene between the aspect marker ZAI and the verb that it governs. Durative -ZHE is a post-verbal morpheme. It always immediately follows the verb, as in this example:

- (9) ta bao-ZHE niao-long pao
 他 抱著 鳥籠 跑
 he carry-DUR birdcage run
 he runs carrying the bird cage

Sheng Ma (1985) and Li and Thompson (1981) point out that use of these two imperfective markers in Mandarin is restricted to certain classes of verbs, based on the lexical *Aktionsart* of those verbs. According to these authors, the use of either imperfective aspect particle with punctual or non-action verbs in Mandarin is ungrammatical.

Perfective -LE is also a post-verbal morpheme, however it may come either directly after the verb or appear at the end of the utterance after additional intervening material:

- (10) ta jiu pao-LE
 他 就 跑了
 he then runs-PRF
 then he runs (off)

In spoken English, the progressive verb aspect is grammaticized similarly to Mandarin, using the auxiliary 'to be' and verb final '-ing' marker:

(11) he is running

Sheng Ma notes, however, that the correspondence between Mandarin ZAI and the grammaticized progressive aspect in English is not complete. He says that progressive marking in English may extend to habitual or repeated activity in English, as in the example listed above where application of progressive aspect to the punctate verb ‘hit,’ gives rise to a reading of iterativity. Sheng Ma claims that in Mandarin such combinations are ungrammatical.

To say that aspect is not grammaticized in a given linguistic system is not to say that there are no means within that system to create aspectual contrasts linguistically. In sampling the English spoken data for instances of linguistic aspects other than the grammaticized progressive, the search was for periphrastic aspect and for lexical and periphrastic *Aktionsart* analogs to the other two target aspects. Instances were sampled in which speakers elected to use verbs or phrases that encoded inherent temporal features compatible with the target aspect distinctions. For example, the verb ‘meander’ in the following sample, taken from a speaker’s description of one of the Vignettes, was taken to express durational *Aktionsart*, and, by extension in its usage here, durative aspectual view:

(12) a baby doll meanders across a green background

In the English speech sampled for this study, phrases and lexical items such as the following were taken to express durativity:

(i) Forms with ‘as’ (or ‘while...,’ ‘when...,’ ‘in the process of...,’ ‘with [someone]...’ [doing something]). For example:

- a. as he’s coming up and the bowling ball is coming down ...
- b. with Granny and Tweety chasing him in the trolley car ...

(ii) VERB + VERB-ing forms such as:

- a. run screaming
- b. come climbing

(iii) 'keep' VERB-ing forms:

he keeps looking around the apartment

(iv) Lexical items with durational *Aktionsart*, for example:

slide, run, roll, fly

Perfective aspect is difficult to pin down in spoken English and often an 'external view of events' is not a feature of event representation that may be unambiguously determined on the basis of linguistic form. Aspect and tense are confounded in English; thus, simple past tense marking can be taken as expressing a perfective view of an event. For the purposes of this study, verbs judged to have an inherent punctual or completive *Aktionsart* were selected; for instance, verbs that in the absence of overt grammatical or phrasal indications of imperfective aspect, strongly resist an imperfective reading. Compare for example: 'hit' versus 'pummel,' 'hear' versus 'listen,' 'finish' versus 'draw to a close.' In addition, verb+particle constructions were included; those where the particle seems to add a perfective aspectual dimension to the linguistic representation of an event; for example: 'fly off' versus 'fly,' or 'fall off' versus 'fall.'

Summary

Study One is an exploratory study with two impetuses. (i) Verb aspect strongly suggests itself as a linguistic category that could relate to distinct ways of conceiving of events; that is, linguistic aspect distinctions may be expected to associate with differences in conceptual representation. If aspect is this kind of category then, within language, we

expect to find patterning in gesture related to aspectual distinctions. (ii) Mandarin and English have been described as differing significantly in how they mark aspect. Aspect marking in Mandarin is seen as less ambiguous and less confounded with systems marking other linguistic contrasts. May such differences in the linguistic signalling of aspect distinctions correspond to between-language differences in event conception during speaking? Analysis of the gestures of Mandarin and English speakers is used to explore this issue.

Methods for Study One

Participants

The response data from all 25 participants (fourteen Mandarin and eleven English speakers) and all elicitation tasks were sampled for this study, and all participants contributed utterance data to all coding categories, although not in equal proportion, since not all participants produced fluent, gesture-accompanied utterances in equal number in each of the categories. However, the participants were sampled as evenly as possible within the constraints imposed by the irregular, unpredictable nature of spontaneous speech and gesture data, and according to the sampling criteria outlined below. The table in Appendix B lists the contributions of each participant to the data set, by stimulus type.

Speech sampling

The sampling process began with an examination of the speech transcripts themselves. At this first step, no attention was paid to the gestures that accompanied the speech. The speech transcript for each participant was scanned for the target aspect-marked (in Mandarin), or aspect- and *Aktionsart*-expressive (in English) utterances. In the case of the English language transcripts, aside from sampling grammatically-marked progressive aspect, lexical and phrasal indications of the target internal temporal event

constituencies were the basis for the selection.

Gesture sampling

Of all the aspect- and *Aktionsart*-marked utterances identified by scanning the speech transcripts, those that were accompanied by representational gestures were extracted and further examined. These utterance-gesture pairs were examined to determine whether the gesture in each case bore any discernable relationship to the event (either an event of motion or of stasis) encoded by the *verb* in the utterance. Utterance-gesture pairs with such verb-related gestures were considered for inclusion in the data set for Study One and all others were excluded. This sampling criterion, based on a semantic correspondence between the verb in a utterance and a co-occurring gesture, excluded any gestures that were semantically related only to an ‘Element’ (a physical object, surface, location, or similar) being described in the utterance. If a gesture encoded nothing of the event of motion or of stasis expressed in the speech it accompanied, then it and its co-occurring speech were excluded from the sample.

The latter sampling criterion becomes clear if we consider the kinds of gestures that have been observed to co-occur, for example, with a spoken expression like, “and the bird flies around.” The accompanying gesture could reference only features of the bird, and omit features of the flying around. This would be the case if the speaker held up one hand with thumb and index finger extended in the shape of the letter ‘C’ to represent the bird by indicating its small size. This sort of ‘Element-only’ gesture was excluded from the analyses of Study One. In contrast, the speaker could make ‘wing flapping’ motions with the fingers of both hands, or could trace the bird’s flight path through the air with an index finger point. In the coding scheme used here, either of the latter two would count as an event-related representational gesture, and be included in the data set. Since the aim of Study One was a comparison among event contours in speech and gesture, which

depend for their realization in large part on how the event expressed by the verb is temporally construed, data selection was restricted to utterances accompanied by event-related gestures (events of motion or of stasis). The assumption was that, if there are distinctions in gesture form that are related to linguistic aspect and *Aktionsart*, then these distinctions would likely be most apparent in gestures related to verb content.

Third, the spoken portion of all utterance-gesture pairs was examined to determine if there were dysfluencies such as pauses, so-called ‘filled pauses’ (for instance, “um” and “uh”), self-interruptions, self-corrections, and repetitions. Completely fluent speech is a somewhat rare thing; at the same time, not all pauses are instances of dysfluency. A fair proportion of the speech elicited by the narration tasks used here was dysfluent in one of the ways just listed. Of particular concern, given the coding parameters of this study, are the sorts of speech dysfluencies identified by Butterworth and Beattie (1978) as associated with instances of lexical search. In the cases analyzed by these authors, speakers search unsuccessfully for a suitable word to express an intended meaning. Moments of lexical search such as these are frequently accompanied, not just by hesitancy and stuttering in speech, but also by small, beat-like gestures, executed rapidly in succession. Such ‘Butterworth beats,’ as they have come to be called, have a repeating, extendable character in common with certain gesture forms that accompany fluent, aspect-marked speech. The latter forms proved to be a significant factor in the analyses of fluent speech and accompanying gesture presented here, and more will be said about them presently. To avoid the possibility of confounding variables of gesture form with values having potentially different etiologies in the language production process, while mimicking one another on this dimension of patterning, dysfluent speech was excluded from the sample. The phenomenon of ‘Butterworth beats’ and what may motivate them will be taken up in the Discussion for Study One, particularly in regard to the possibility that some commonality may exist at a

different level of analysis, between these beats and the repeating, extendable gesture forms that appear in certain fluent, aspect-marked spoken contexts.

Sampling totals

Based on the above sampling criteria, from the Mandarin data 50 utterance-gesture pairs where speech expressed a perfective view of the event, and 50 utterance-gesture pairs where speech expresses an imperfective view of the event, were extracted. The set of perfective-marked utterance-gesture pairs consists of verbs all paired with the perfective aspect marker -LE. The set of imperfective-marked utterance-gesture pairs consists of 25 utterances marked with progressive ZAI and 25 marked with durative -ZHE. To this set of 100 Mandarin utterance-gesture pairs was added a similar set of 100 pairs from the English narrations. The two sets are analogous in terms of the aspectual views of event conveyed by the utterances of each.

Utterance-gesture pairs were selected for inclusion in the data set by starting at the beginning of each speaker's narration and moving sequentially through it, extracting those pairs that met the sampling criteria of the study. It was not possible to balance the contributions of all participants numerically, since not all subjects produced fluent, gesture-accompanied speech marked for each of the three target aspects in equal amounts. Further, as will be mentioned below, not all Mandarin speakers used the progressive marker ZAI, either because they did not frame any portions of their narrations in the progressive aspect, or they did so but utilized other, periphrastic means to express it. Similarly, there was one Mandarin speaker who produced no fluent, gesture-accompanied uses of speech marked with durative -ZHE. No more than five utterance-gesture pairs of one spoken aspect or *Aktionsart* context were extracted from the narration of any one speaker. These additional sampling criteria were necessary because

different speakers were found to have distinct speaking and gestural styles, that appear to involve some of the dimensions of gesture duration and form relevant to these analyses. A large number of samples of any one of the aspect types from a single speaker could therefore have skewed the results. Appendix B summarizes the contributions of each subject to the three target aspect categories.

Appendix A lists the utterance-gesture samples for each language. The speech-accompanying gestures are bracketed, stroke phases bolded and hold phases underlined. Examination of the English utterances listed in Appendix A will show that, in the case of the perfective and durative, the samples chosen typically reflect inherent event contour, the *Aktionsart*, as expressed in choice of verb or periphrastic construction, as opposed to, or in addition to, explicitly marked verb aspect.

Gesture coding

Several features of gesture production were coded. An exploratory pilot study had earlier revealed a possible relationship between verb aspect in speech, gesture stroke duration, and certain features of stroke form. These formal features of gesture strokes were systematically coded for Study One.

The duration of gesture strokes in 60ths of a second was recorded, based on the visible time codes burned onto the image of the videotapes themselves. These values were later converted to milliseconds for analysis. The presence and durations of all post-stroke holds were coded. A word is necessary here about gestural holds that were considered to be the stroke phases of the gestures in which they occurred. As was mentioned in the discussion of the mechanics of gesture production in the previous chapter, what allows us to identify a certain portion of a speech-accompanying gesture as its stroke phase is the appearance of semantic value in this phase of the gesture that is related in some way to the semantic value of the co-occurring speech. This means that the

stroke phase of a gesture need not be a movement phase. In a speech-gesture production where the focus is on a static, configurational ‘event,’ as opposed to a motion event, the stroke phase of a gesture is likely to be a held configuration of the hands. For example, the following utterance from one of the cartoon narrations is accompanied by a gestural hold, signified by the underlined, bold face type:

(13) [so he’s got the bird cage and he’s running away]

In example (13) the speaker’s arms are held out in a closed circle in front of her and maintained that way for the duration of the bolded, underlined portion of the speech. Thus, the speaker is depicting the static configuration of arms holding something, and the held character of the gesture is key to its semantic value. Such cases are called ‘hold-strokes.’

In addition, the presence of post-stroke holds was coded for this study. These holds that follow movement strokes are distinguished in coding from hold-strokes. Among post-stroke holds, a distinction was made between ‘syntactic’ versus ‘semantic’ holds. A syntactic post-stroke hold is here defined as one that itself, as a hold, appears to have no semantic content specifiable in relation to the portion of speech it accompanies. This can be a hold that spans pauses in speech, or one that terminates when the end of the syntactic, phrasal, speech unit in which the stroke occurs is reached. Kita’s (1990) analysis of *pre*-stroke holds as gestures that “wait for speech to catch up” is in keeping with what is being called a ‘syntactic’ hold here. The presence and extent of the pre-stroke holds analyzed in that work are tied to the linear syntax of an utterance; to the design detail of speech code that constrains it to express meanings in a sequentially-extended structure. Nothing about the pre-stroke hold itself, in combination with its co-occurring speech, suggests that its existence encodes narrative-level semantic content

related to the cartoon events being described.¹⁴ In Kita's analysis, further, there was no suggestion that the hold *as hold* was semantically interpretable in relation to the speech; rather, pre-stroke holds are cases where the organizational phase of gesture is completed before the process of speech production has brought gesture's target speech locus into being. Similarly, a syntactic post-stroke hold shows the scope of the gesture's stroke over sequentially-produced speech. This was described in the outline of GP theory given in the Introduction. Such holds make clear the exact portion of speech to which the semantic value of the stroke phase of the gesture is linked in thinking, without the movement stroke itself extending to cover the entire phrase.

What will be called here a 'semantic' post-stroke hold, in contrast, is one that bears some sort of evolving semantic relationship to the movement stroke that it follows as well as to the speech with which it, the hold itself, co-occurs. In such a case, the semantics of the hold *as hold* are clearly relevant for its interpretation. An example of this is when a speaker narrates the event of putting on a suit of clothes, accompanied by a gesture in which both hands in the form of loose fists move down simultaneously, in time with the verb "wears," from starting positions just in front of the shoulders, to a hold position at about chest level, as if hanging on to the lapels of an imaginary suit while saying,

(14) [he **wears the monkey's costume**]

A post-stroke hold of this kind is judged to have a semantically-motivated relationship to

¹⁴Holds other than 'hold-strokes' certainly are in and of themselves meaningful, but in a different sense and on a different level of linguistic analysis than what is intended here. The same may be said of persistence in maintaining, all in one area of gesture space (itself 'held'), an extended series of movement strokes interspersed with holds. Speakers routinely map out different areas of their gesture spaces to correspond to different discourse organizational units within their narrations. When gestural holds are considered in relation to this level of analysis, it is clear that how they index this organizational scheme is an element of their semiotic value.

the preceding movement stroke and, in addition, to the speech with which it co-occurs. Given that the onsets and offsets of stroke phases of speech-accompanying gestures are identified according to their semantic value interpreted relative to the accompanying speech, in those cases where the post-stroke hold was of the semantic type, it was coded as being part of the stroke, and contributed its duration to that of the movement stroke for the purposes of this analysis.

An alternative to analyzing such a movement stroke combined with the following hold as a single gesture for the purpose of calculating stroke durations, would be to consider them — movement and then hold — as two separate gestures. In the example listed above, the movement stroke could be ‘putting on’ the clothes, and the hold could be a separate gesture, a hold-stroke meaning ‘wearing the clothes.’ Arguing against this is the fact that the two gesture phases in sequence time with an extent of speech that is expressive of an event unity. Further, the possibility that two gesture phases arise from separate idea units, or are separate production ‘pulses’ is generally considered when there is some point of re-setting between the two, rather than a seamless join. Overall, the instances coded in the way described for example (14) as having a movement stroke and a semantic post-stroke hold, give the impression of a single speech-gesture ‘pulse.’

Another feature of gesture form appeared on the basis of the pilot study to be related to spoken aspect. This was a formal feature of stroke movement. Two categories were distinguished: ‘non-extendable’ and ‘extendable.’ Non-extendable strokes are those movements that appear to have inherent limitations on their extension. They are typically quick, uni-directional movements. The category name is motivated by the fact that many of these strokes are executed in such a way that they truly are physically non-extendable; that is, when a gesturing hand flaps quickly downward, for instance, pivoting on the wrist, it reaches a point where further movement on the wrist pivot is not possible. An extendable stroke in contrast is one that is executed either as a single-vector but sustained

movement, or with a multi-vector ‘agitated’ motion, as when a gesturing hand flaps repetitively up and down on a wrist pivot. Note that the term ‘agitated’ as a description of an extendable movement stroke is used here to distinguish multi-vector, typically repeating movements, from single-vector movements that terminate at the end of a sustained movement along one vector of movement. To be classified as ‘agitated’ a motion stroke need not be a jagged or uneven movement; it may as well be smoothly executed. For example, a repeating circular motion of the gesturing hand would be classified as ‘agitated’ according to the coding scheme used here, because of its multi-vector execution. In general, the category name ‘extendable’ indicates that these are movement strokes that have no physically-mandated instant of cessation. From the physical perspective, they have the potential to persevere for an extended period.

Several indices of gestural complexity were coded. Again, results of the exploratory pilot had revealed a possible relationship between verb aspect in speech and differing degrees of complexity of the accompanying gesture. Specifically, in the presence of imperfective aspect marking in speech, speakers’ gestures appear to encode more components of the stimulus event being described and of its context. Therefore, features of gesture form were coded such as whether the gesture is one- or two-handed, and if two-handed, whether the hands are ‘mirror images’ of each other or assume contrasting shapes and attitudes. When the hands mirror each other in a gesture, typically what is being represented is some unitary component of an event representation; when they contrast, it is often the case that multiple components of an event and their relationships to one another are simultaneously represented.

A related index of complexity was a kind of ‘density’ of semantic encoding. Varying degrees of this were noted by coding the presence or absence in gesture of the expression of the event components of PATH and MANNER (in cases of expression of motion events), as well as the presence or absence in gesture of ‘Element’- marking. The

latter refers to any specific features of the concrete physical features of objects or locations that are part of the event to which the speaker refers, for instance a bowling ball, a bird cage, or a surface on which a movement occurs. When such element features were integrated with the gestural expression of an event (whether an event of motion or of stasis) this was recorded. The gesture coding category of ‘Element’ can be understood as collapsing Talmy’s (1985) components of FIGURE/AGENT, GROUND, and LOCATION. To understand the category, consider a very commonly-observed kind of motion-expressive gesture that was *not* coded as Element-marked. This is when a speaker uses the tip of her index finger to trace an imaged PATH trajectory through the air. Such a gesture was regarded as encoding only the PATH of motion. A gesture would be coded as Element-marked as well if for instance instead of a point, wiggling fingers representing the running legs of an animate FIGURE, or a tensed, cupped hand representing a ball in motion, were to move along the represented PATH of motion, or if some element of the GROUND against which the PATH of the moving trajectory was instantiated were gesturally represented along with the motion.

Coding summary

The units of analysis that figured in the coding for Study One were: (i) in speech: the linguistic aspects of perfectivity and imperfectivity (progressive and durative), and (for English) their temporal structural allies in linguistically-encoded *Aktionsart*; (ii) in gesture: stroke duration or, where appropriate, duration of a stroke together with its post-stroke hold; a featural analysis of gesture form to distinguish non-extendable from extendable stroke phases as well as degrees of gestural complexity: (a) two-handed mirror versus contrast gestures, (b) the several components of motion events — MOTION,

PATH, MANNER, plus a generalized notion of Element encoded in gesture.

Table 1. Summary of coding categories in speech and gesture — Study One.

<u>Speech</u>	<u>Gesture</u>
<u>Aspect:</u>	<ul style="list-style-type: none"> • duration of stroke and post-stroke hold • where applicable: duration and category of post-stroke hold
• Perfective	• stroke category: non-extendable versus extendable
• Progressive	• where applicable: extendable strokes — sustained or agitated
• Durative	• stroke form: one-handed versus two-handed
	• where applicable: two-handed gestures — mirror or contrast
	• event components: motion (path and manner) and element

Analysis

Since the aspectual distinctions that are the focus of this analysis are unambiguously grammatically signalled in Mandarin, unlike in English, the data from the Chinese participants were analyzed first to determine whether any distinct gestural forms coincide with the use of these markers. The English-speaking participants' gestures from analogous speech contexts were then compared to those of the Chinese.

First examined were some outcomes of the speech sampling process that was based on the speech transcripts alone. Aspect marking in speech was examined to determine if participants tend to prefer one aspectual viewpoint over another, or if any stimuli differ in terms of the spoken aspect marking they tend to elicit.

Next, gesture stroke durations were calculated and averaged by aspect category and language. A two-way analysis of variance was performed to determine whether there is a relationship between imperfective versus perfective spoken aspect contexts and the durations of gesture stroke, and whether Mandarin and English speakers differ in this regard. For the within-imperfective aspect comparison, two-sample *t*-tests were run to determine if there is a difference in stroke duration in each language between progressive

and durative contexts.

The features of gestural form and complexity that seemed on the basis of the pilot study to covary with aspectual view and *Aktionsart* were tallied by aspect category and the relative proportions in which they occur determined by category and language. Chi-square analyses were performed to determine whether these features reliably vary by aspect category within each language.

Note that for these planned analyses, the unit of analysis was not the participant, but rather, the utterance-gesture pair. There were several issues that determined this feature of the design of the analyses. At root, the issues all derive from the nature of the elicitation itself, and the resulting naturalistic character of the spontaneous speech and gesture data on which the analyses are based.

First, the sampling procedure was designed to minimize the probability that the speech-gesture characteristics of any single participant would unduly influence the results. Recall that the attempt was to sample utterance-gesture pairs in every aspect category from every speaker that met the criteria of spoken fluency and accompaniment by a representational gesture, with no speaker being allowed to make more than five contributions. For the occasional speaker who produced criterion utterance-gesture pairs in excess in a given category, the excess were culled randomly. Second, the number of criterion pairs available for sampling from some of the speakers in some categories was quite small. The two probable primary causes for this are: (i) the requirement of sampling fluent, gesture-accompanied speech, and (ii) the fact that some speakers appeared to emphasize one or another of the aspectual views in their whole narrations, making utterances marked for the others infrequent in their elicited material.

Overall, as Appendices B and D show, the number of utterance-gesture pairs contributed by each participant in each category was usually rather small. If the analyses were to be organized around the participant as the unit of analysis, this would mean in

most cases taking the median value of a very small set of measurements; in some cases using a single number from one participant against the medians of measurements from others. Thus the sampling procedure itself was designed to minimize the participant as a factor in the results, and the design of the analyses is such that the origin of any particular utterance-gesture pair is ignored. Operationally, this amounts to an acknowledgment that an assessment of how gesture may vary in relation to the aspectual view of spoken utterances across individuals is beyond the reach of this sampling procedure and analysis. Yet, the sampling procedure and the design of the analysis are in effect mandated by the nature of the elicited language samples. What would be needed to assess individual variation in the relationship between spoken aspect and gesture is a somewhat different elicitation procedure, one that would reliably elicit speech marked for the different aspectual views in all participants, and in sufficient quantity to permit statistical comparisons both within and between subjects. It would have to be a procedure that would have as high a likelihood of eliciting accompanying spontaneous gesture as the descriptive narration tasks employed here, while not permitting participants to adopt a more or less exclusive tense-aspect 'frame.'

In addition to the planned analyses, some further phenomena in speech and gesture that appear to be aspect-linked were observed: (i) instances of apparent layering in gesture of features related to more than one aspectual view; (ii) features of metaphoric gestures in aspect-marked contexts that are similar to those of concrete iconics in these contexts; (iii) some instances of the occurrence of the gestural forms identified as related to spoken aspect, but that occur in non-aspect-marked speech contexts; (iv) evidence of the flexibility speakers have in choice of aspectual view.

Coding reliability

Ten percent of the data used in this study were randomly selected, half each from English and Mandarin, and coded separately by another experienced gesture coder for the details of form and content that are relevant to the analyses here. This coder is also knowledgeable of Mandarin. The reliability coding categories for Study One and percent agreement between the two coders are as follows:

1) discerning the presence of a gesture stroke	96%
2) identifying a gesture as representational	93%
3) identifying post-stroke hold type	86%
4) form of gesture stroke: non-extendable, extendable, agitated	95%
5) identifying stroke as one or two-handed, mirror or contrast	93%
6) element-marking	96%
7) identifying stroke onset and offset (± 0.10 sec.)	84%
8) identifying post-stroke hold offset (± 0.10 sec.)	92%

Results of Study One

For the sake of simplicity, the term ‘aspect’ will be used as the label for all of the comparisons of linguistically-encoded temporal constituency distinctions discussed in this section. Analysis of the gestures in the set of 200 Mandarin and English utterance-gesture pairs shows that there are differences in the timing and form of gestures related to aspectual view as expressed in speech.

Outcomes of the Speech Sampling Process

Across the Mandarin speakers, use of the -LE perfective aspect marker was more than twice as common as the two imperfective markers, ZAI and -ZHE, combined.

A preference for framing responses largely in the perfective aspect, or at least avoidance of overt indicators of imperfective aspect, seemed to be true of both speaker groups. Only three of the Mandarin and three of the English speakers framed the larger portions of their responses in an imperfective aspectual view. This finding is in keeping with Slobin and Berman's (1994) observation of a preponderance, cross-language, of the perfective view of events in the narrative context of re-telling a story or witnessed event. These authors (1994:27) note that cross-language, narrators seem to prefer perfective aspect for these kinds of re-tellings, and have a tendency to select an overall anchor tense and aspectual view, to which they then adhere for most of the narrative. In these cases, use of the opposing tenses and/or aspects constitute marked usages. These assertions appear generally borne out in the present set of narration data. Certainly, extracting sufficient samples of perfective-marked speech from these narrations for inclusion in the analysis was not a problem, as every elicited speech sample had them in abundance. As for imperfective aspect, among Mandarin speakers the durative -ZHE marker was more common than the progressive ZAI marker. In regard to the latter, extracting a sufficient number of utterances containing progressive ZAI was in fact a challenge, as use of this grammatical particle was comparatively infrequent in these data. Several Mandarin narrations did not have a single use of ZAI in fluent, gesture-accompanied speech. This was less true of progressive marking in the English sample; perhaps a reflection of broader applicability of progressive aspect across classes of verbs in English that was mentioned in the Introduction.

It should also be noted that the Chinese participants frequently chose not to use one of the aspect-marking grammatical particles, but chose instead from among Mandarin's periphrastic means of expressing aspect. For example, it is possible in Mandarin to express an imperfective aspectual view in one of the following ways: (i)

verb reduplication, (ii) the VERB-*come*-VERB-*go* form, or (iii) the VERB-*one*-VERB form. Such usages are sometimes described as expressing a progressive aspectual view. If this is their function, then the relatively infrequent use of ZAI may in part be accounted for by preference on the part of some speakers for these periphrastic usages. These were frequent in the data. Below are some illustrations of imperfective aspect conveyed by these periphrastic means. It is not clear from such examples that these forms should properly be considered specifically progressive as opposed to durative. It could rather be that they are globally imperfective and do not distinguish between the senses of ‘dimensionless slice’ versus ‘having temporal extent’ that are said to distinguish those two aspectual views.

- (15) a. nei-zhi mao zai jie-shang gun gun gun
 那隻貓 在街上 滾滾滾
 that-CL cat on street-top roll roll roll
 the cat rolls / is rolling on the street
- b. ta zou-lai-zou-qu
 他走來走去
 he walk-come-walk-go
 he paces / is pacing back and forth
- c. mao zai dian-xian-shang zou-yi-zou
 貓 在電線上 走一走
 cat on electric-wire-top walk-one-walk
 the cat walks / is walking on the electric wires

Which mode of expression to use, grammatical or periphrastic, for establishing an imperfective view of events in speech, seems to be up to speaker discretion. For example, comparing two of the Taiwanese speakers who participated in the Vignettes Task, both of whom framed a large proportion of their responses in imperfective aspects: one chose the grammatical-particle option of ZAI for expressing progressive aspect more

than half the time, the other chose to use ZAI only once.

Some mention must be made of the relative difficulty of judging the aspectual view of many spoken English utterances, since, with the exception of the imperfective progressive, aspect is not unambiguously marked using grammatical means, as was outlined in the Introduction to this chapter. Verb and periphrastic *Aktionsart* guided the extraction of the English perfective samples and to some extent the durative as well. For the durative, key words and phrases that denote durativity (for example, ‘while,’ ‘as,’ and so on, as listed above) were cues to the presence of this aspectual view. Nevertheless, from utterance to utterance it was frequently difficult to judge the speaker’s choice of aspectual view on the basis of linguistic form alone in English. The contributions of each speaker to the data set for Study One are listed in Appendix B.

The choice of aspect marking in speech for both groups of speakers appeared to be largely independent of the particular stimulus event being described. This was particularly clear in the data from the Vignettes Task, where a given vignette might elicit responses from different speakers framed in terms of differing aspects. More will be said on this point below.

Gesture Stroke Durations

Table 2 shows that the durations of gesture strokes that accompany perfective- and imperfective-marked speech contexts differ. The strokes of gestures that occur in perfective contexts are much shorter on average than those that occur in imperfective contexts. Table 2 shows the mean stroke durations of gestures accompanying spoken perfective and imperfective aspects in Mandarin and English. Within each language the difference in stroke durations of gestures accompanying speech marked for these aspectual distinctions is highly statistically significant. An additive model, two-way

Table 2. Mean durations in seconds, and standard deviations, of gesture strokes accompanying utterances with perfective versus imperfective aspect; Mandarin versus English.

	Mandarin	English
Perfective:	mean = 0.274 std.dev. = 0.186 <i>n</i> = 50	mean = 0.249 std.dev.= 0.114 <i>n</i> = 50
Imperfective:	mean = 1.018 std.dev. = 0.934 <i>n</i> = 50	mean = 1.158 std.dev. = 0.623 <i>n</i> = 50

Table 2a. ANOVA table for gesture stroke duration; Mandarin versus English, perfective versus imperfective aspects.

Source	DF	SS	MS	F	P
Language	1	0.164	0.164	0.50	0.480
Aspect	1	34.114	34.114	104.31	0.000
Error	197	64.430	0.327		
Total	199	98.707			

analysis of variance comparing these stroke durations between the two languages yielded a significant effect for aspect, $F(1) = 104.31$, $p < .000$, and no effect for language, $F(1) = 0.50$, $p = .480$. In addition, a separate analysis of variance was run using a multiplicative model and the language-by-aspect interaction was not significant ($p > .38$).

Table 3 shows the mean gesture stroke durations for the gestures accompanying spoken progressive and durative, the two imperfective aspects. The results of two-tailed *t*-tests for differences between independent means showed that, among the Mandarin speakers, this difference in stroke durations of gestures accompanying progressive versus durative spoken contexts is not statistically significant, $t(41) = 0.29$, $p = 0.77$, but it is

Table 3. Mean gesture stroke durations in seconds and standard deviations: progressive versus durative aspect; Mandarin versus English.

	Mandarin	English
Progressive:	mean = 0.979 std.dev. = 0.739	mean = 0.952 std.dev. = 0.420
Durative:	mean = 1.057 std.dev. = 1.110	mean = 1.363 std.dev. = 0.726

significant in the English speakers' speech-gesture productions, $t(38) = 2.45$, $p = 0.019$. Thus English, although similar to Mandarin in that gesture stroke phases are longer in imperfective spoken contexts than in perfective ones, nevertheless appears to differ from Mandarin within the category of imperfective spoken contexts. In English durative-marked contexts, stroke durations are longer on average than those in progressive-aspect contexts. However, the results of a two-way analysis of variance presented in Table 3a indicate that the main effects due to language and aspect are not significant; neither is the language-by-aspect interaction.

Table 3a. ANOVA table for gesture stroke duration; Mandarin versus English, the imperfective aspects: progressive versus durative.

Source	DF	SS	MS	F	p
Imperfective Aspect	1	1.4908	1.4908	2.40	0.124
Language	1	0.4858	0.4858	0.78	0.379
Imp*Lang	1	0.6906	0.6906	1.11	0.294
Error	96	59.5832	0.6207		
Total	99	62.2505			

Table 3b shows the results of a one-way analysis of variance that was run to determine if the mean stroke durations of gestures accompanying utterances of each

imperfective aspect category are both significantly longer than stroke durations in perfective-marked spoken contexts. There are significant differences in stroke durations accompanying the three categories of spoken aspect, $F(2) = 55.58$, $p < .000$. Figure 2 indicates that the differences lie between the perfective and each of the two imperfective aspects.

Table 3b. ANOVA table for gesture stroke duration; Mandarin and English together: Imperfective Progressive versus Imperfective Durative aspects.

Source	DF	SS	MS	F	p
Aspect	2	35.605	17.802	55.58	0.000
Error	1	97	63.102	0.320	
Total	1	99	98.707		

Aspect	<i>n</i>	Mean	Std.Dev.	
Progressive	50	0.9658	0.5953	-----+-----+-----+-----+
Durative	50	1.2100	0.9410	(----*---)
Perfective	100	0.2619	0.1538	(--*---)
Pooled Std. Dev. =		0.5660		-----+-----+-----+-----+
				0.35 0.70 1.05 1.40

Figure 2. Individual 95 Percent Confidence Intervals for Mean, Based on Pooled Standard Deviation.

The Complexity of Gesture Forms

The differences in stroke duration in different spoken aspect contexts are in keeping with more general divergences in gesture production across spoken aspect contexts. Results in the coding categories that had to do with gestural complexity show tendencies toward reduced versus elaborated gesture forms that vary with spoken aspect.

Table 4 summarizes these features of gesture forms, as proportions of total gestures occurring in different spoken aspect contexts. This table incorporates the differences in stroke durations just outlined. The findings summarized in Table 4 make it possible to profile the kinds of gestural performance that are most typical of the target spoken aspect contexts, and to identify the specific gestural features that best characterize each of these contexts. The gestural profiles of the perfective, progressive and durative spoken aspectual views each are presented in turn in the following paragraphs.

A few details must be mentioned about the proportions listed in Table 4. The table compares gestures that occur in the three aspect contexts in terms of the following features:

(i) Stroke durations in milliseconds. Following this feature, the rest of the comparisons are in terms of proportions of total gestures by aspect category and language. These totals are listed at the bottom of Table 4.

(ii) Non-extendable strokes versus extendable strokes, calculated as a proportion of the total speech-gesture sample within each aspect category and language; that is, 50 perfective-, 25 progressive-, and 25 durative-marked utterances accompanied by gesture sampled from across the speakers of each language. The proportion of agitated motion strokes within the category of extendable strokes is listed in square brackets following the overall proportion that includes both smooth/sustained motion strokes, as well as agitated ones. Note that this latter, square-bracketed proportion is itself also a proportion of the same total number of gestures within aspect category and language; that is, 50, 25 and 25, as above. For example, the proportions listed for Mandarin progressive-marked contexts are '.92 [.84].' This means that 92% of the 25 gestures in Mandarin progressive contexts are extendable stroke gestures, and a subset of these, 84% of the 25 total, are agitated-motion strokes. The difference in incidence of extendable versus non-extendable strokes across spoken aspect contexts is significant in Mandarin, chi-square (2) = 69.890,

Table 4. Gesture durations and complexity across different spoken aspect contexts, Mandarin versus English.

		PERFECTIVE	PROGRESSIVE	DURATIVE
(i) <u>Gesture Stroke duration (means):</u>				
MA		0.274 secs.	0.979 secs.	1.057 secs.
EN		0.249 secs.	0.952 secs.	1.363 secs.
(ii) <u>Stroke form — proportion of gestures within aspect category:</u>				
MA	non-extendable	.98	.08	.24
	extendable [agitated stroke]	.02 [.02]	.92 [.84]	.76 [.32]
EN	non-extendable	1.00	.04	.08
	extendable [agitated stroke]	---	.96 [.88]	.92 [.52]
(iii) <u>Gesture Holds — proportions of stroke followed by holds, and their types:</u>				
MA	(a) no post-stroke hold	.80	.88	.52
	post-stroke hold	.20	.12	.48
	(b) syntactic	.80	.67	.17
	semantic	.20	.33	.83
	(Subset of Total Gestures : 10		3	12)
EN	(a) no post-stroke hold	.84	.92	.60
	post-stroke hold	.16	.08	.40
	(b) syntactic	1.0	1.0	.00
	semantic	.00	.00	1.0
	(Subset of Total Gestures: 8		4	10)
(iv) <u>No. of hand(s) and the relationship of the hands in 2-handed gestures:</u>				
MA	(a) 1-handed	.78	.68	.28
	2-handed	.22	.32	.72
	(b) 2-handed / mirror	.64	1.00	.72
	2-handed / contrast	.36	---	.28
	(Subset of Total Gestures: 11		8	18)
EN	(a) 1-handed	.68	.56	.20
	2-handed	.32	.44	.80
	(b) 2-handed / mirror	.81	.82	.50
	2-handed / contrast	.19	.18	.50
	(Subset of Total Gestures: 16		11	20)
(v) <u>Proportion of all non-pantomimic gestures that encoded element features:</u>				
MA		.14	.08	.28
	(Subset of Total Gestures: 7		2	7)
EN		.10	.12	.24
	(Subset of Total Gestures: 5		3	6)
<hr/>				
Total Number of Gestures:	MA	50	25	25
	EN	50	25	25

$p < .000$, and also in English, chi-square (2) = 88.760, $p < .000$. The difference in incidence of stroke agitation across aspect contexts is significant as well in both languages; Mandarin, chi-square (2) = 53.429, $p < .000$ and English, chi-square (2) = 60.967, $p < .000$.

(iii) The data on presence or absence of post-stroke holds are presented in two sections, (iii) *a* and (iii) *b*. Note that the proportions listed in these two sections are calculated on the basis of different totals. The sections (iii) *a* compare the proportions of the total gestures within each aspect category and language (50, 25 and 25, as above); those that have no post-stroke hold versus those that do. For example, we see that only 10 of the 50 gesture phrases accompanying perfective-marked utterances in Mandarin, or 20%, include a post-stroke hold of any kind, whereas 48% of the 25 gesture phrases in durative contexts in that language include a hold of some kind. Overall in Mandarin, the incidence of post-stroke holds varies significantly by spoken aspect category, chi-square (2) = 9.973, $p < .01$. The same is true for English, chi-square (2) = 9.000, $p < .02$. The sections (iii) *b* in Table 4 compare, within the subset of gestures having post-stroke holds, the proportions of each of the two kinds of post-stroke holds that were coded for this analysis, ‘syntactic’ and ‘semantic.’ Note that the proportions in sections (iii) *b* are calculated on the basis of the subset total only. The numbers of gestures in each such subset are listed underneath the proportions in the (iii) *b* sections. As was just mentioned, in Mandarin, 12 of 25 gestures in durative-marked utterance contexts, or 48%, have post-stroke holds of some kind. Of this subset, 10, or approximately 83% of the 48%, are ‘semantic’ post-stroke holds. In this comparison, the expected frequencies in several cells of the chi-square table are too small to test for statistical significance, however a pattern is evident in the data. In English for example, the number of post-stroke holds in spoken perfective and progressive contexts is eight and four, respectively and all these

holds are of the 'syntactic' type; in spoken durative contexts the number of post-stroke holds is ten, all of the 'semantic' type.

(iv) As was true in section (iii) of the table, the data on the number of hands used to gesture, one versus two, are presented in two sections, (iv) *a* and (iv) *b*. Again the proportions listed are calculated on different bases in the two sections. In the sections (iv) *a* the gestures that are one-handed versus two-handed are listed as proportions of total gestures within each aspect category and language (50, 25 and 25, as above). For example, 78% of the 50 gestures that occur in Mandarin spoken perfective contexts are one-handed and 22% are two-handed. Overall, the Mandarin speakers produced 63 one-handed and 37 two-handed gestures; English speakers produced 53 one-handed and 47 two-handed gestures. Between languages and collapsing the spoken aspect categories, the difference in incidence of one-handed versus two-handed gestures is not significant, chi-square (1) = 2.053, $p = .15$.

In sections (iv) *b* the proportions are based on subsets of the 50, 25 and 25 gesture totals for each aspect category. For example, 80% of all of the 25 gestures that occur in English durative contexts are two-handed — a subset comprising 20 gestures. Of this 20-gesture subset of the total, 10, or 50%, are two-hand contrast gestures. The numbers of gestures in each such subset again are listed underneath the proportions in sections (iv) *b*. Between spoken aspect contexts in Mandarin the incidence of 2-handed mirror versus contrast configuration gestures was not significantly different, chi-square (2) = 3.554, $p > .15$. The difference in incidence by aspect category in English, however, was marginally significant, chi-square (2) = 5.241, $.10 < p < .05$.

(v) The final section of Table 4 lists the proportions of Element-marked gestures by aspect category and language and is also based on subsets of the total gestures. The set of data considered for these comparisons is a rather small subset of the whole. All pantomimic gestures were excluded from this comparison, for the reason that

Element-marking in pantomimic performances is in a sense less ‘meaningful’ than it is in non-pantomimic gestures. This is because the Element of FIGURE/AGENT is automatically a part of such performances, as is something like INSTRUMENT in the case of a pantomimic enactment of hitting a cat with an umbrella, for example, where the hand will have a form suggestive of holding this INSTRUMENT. Most of the non-pantomimic gestures in this sample from both languages are not Element-marking. This is 54% of the total Mandarin sample and 61% of the total English sample.

Spoken Perfective Aspect

With but one exception, all of the 100 gestures in the two languages that accompany perfective-marked speech are single-vector, non-extendable gestures. The exception is a quick, bi-directional iconic gesture that depicts a ‘snatching’ action.

In regard to the complexity index of one- versus two-handed gestures, for both speaker groups, the proportions of two-handed gestures are lowest in perfective-marked aspect contexts. Table 4 displays a cline that holds for both languages, relating the proportions of one-handed to two-handed gestures across the three spoken aspect contexts. To better illustrate this cline, the subset of the data shown in Table 4 that is a comparison of one-handed and two-handed gestures is reproduced in Table 4a, reconfigured to highlight it. In speakers of both languages, the proportion of the speech-accompanying gestures that are one-handed declines from spoken perfective-, to progressive-, to durative-marked contexts, and the proportions of two-handed gestures across these aspect contexts of course increases correspondingly.

Table 4a. Hand(s) used to gesture across different spoken aspect contexts; Mandarin versus English.

		PERFECTIVE	PROGRESSIVE	DURATIVE
MA	1-handed	.78	.68	.28
EN	1-handed	.68	.56	.20
MA	2-handed	.22	.32	.72
EN	2-handed	.32	.44	.80
<i>Number of Gestures:</i>				
	MA	50	25	25
	EN	50	25	25

Referring now back to Table 4, for those cases where spoken perfective aspect is accompanied by a two-handed gesture, in the majority of those cases, 64% in Mandarin and 68% in English, the hands are mirror images of one another; for example, as in two hands together appearing to knock down a ball. Table 4 also shows that gestural encoding of features of the physical objects involved in an event, or other object features of the scene as a whole, here referred to as Element-marking, is proportionally less in spoken perfective contexts than in durative contexts. Of the non-pantomimic gestures in perfective contexts, 14% of the Mandarin speakers' and 10% of the English speakers' were element-marking, compared to 28% and 24%, respectively, in durative aspect contexts.

Overall, post-stroke holds were infrequent in spoken perfective contexts. In the sample of 50 gestures accompanying such contexts in Mandarin, 16% had post-stroke holds of the 'syntactic' type; 4% of the 'semantic' type. Of the 50 such English gestures, again 16% were followed by syntactic post-stroke holds and no gestures accompanying spoken perfective in English had semantic post-stroke holds.

Another subset of the data from Table 4 is presented in Table 4b. These are proportions based only on those utterance-gesture productions that are expressions of

motion events. Motion event expressions represent 62% of the Mandarin sample and 69% of the English sample. These subset-based comparisons are motivated by a pattern uncovered in the exploratory pilot of gestural encoding of motion event components in perfective-marked contexts. There it was observed that (i) gestures in spoken perfective contexts are often limited to the expression of PATH and (ii) there appears to be less MANNER- and Element-marking (FIGURE, GROUND, and so on) in perfective than in progressive and durative spoken contexts. Comparisons involving the event components of PATH and MANNER make sense only in relation to speech-gesture productions expressive of motion events. The remainder of the sample, those productions dealing with stasis events or metaphoric expressions were therefore excluded from the summary in Table 4b. When considering the proportions listed in Table 4b, note that all instances of multi-vectored motion in gestural expressions of motion events were assumed to encode MANNER of motion. Under the ‘MANNER’ heading in the table, therefore, are all those gestures that have agitated motion strokes. Note further that the data are divided into non-pantomimic and pantomimic modes of representation because, as was mentioned above, the import of the feature Element-marking is dependent upon this parameter. Pantomimic gestures make up 16% of the Mandarin subset of motion event-expressive productions represented in Table 4b, and 20% of the English subset.

Of the motion event-expressive gestures in spoken perfective contexts, 64% in Mandarin and 60% in English encode only the PATH trajectory of some moving FIGURE. This is very often an abbreviated trajectory, given the short duration typical of gesture strokes in perfective contexts. These gestures often seem only to point or suggest a PATH rather than to trace its complete trajectory. Thus in perfective contexts it is common to reduce all event dimensions to PATH, and then frequently to reduce even that, resulting in

Table 4b. Expressions of motion events only — gestural encoding of event components across different spoken aspect contexts; Mandarin versus English.

		PERFECTIVE		PROGRESSIVE		DURATIVE	
		(Pantom.)		(Pantom.)		(Pantom.)	
MA	PATH only	.64	n/a	.06	n/a	.33	n/a
	MANNER only	0	n/a	.25	n/a	0	n/a
	PATH+MANNER	0	n/a	.44	n/a	.40	n/a
	PATH+ELEMENT	.14	.22	0	0	.20	0
	PATH+MANNER+ELEM.	0	0	.13	0	0	0
	MANNER+ELEM.	0	0	.06	.06	0	.10
	<i>Total Gestures: 62</i>	36		16		10	
<i>Counts: PATH only</i>		23		1		3	
<i>Counts: All other compositions</i>		13		15		7	
EN	PATH only	.60	n/a	0	n/a	.22	n/a
	MANNER only	0	n/a	.09	n/a	0	n/a
	PATH+MANNER	0	n/a	.64	n/a	.28	n/a
	PATH+ELEMENT	.15	.20	.09	0	.17	0
	PATH+MANNER+ELEM.	.03	.03	.09	0	.11	0
	MANNER+ELEM.	0	0	0	.09	0	.22
	<i>Total Gestures: 69</i>	40		11		18	
<i>Counts: PATH only</i>		24		0		4	
<i>Counts: All other compositions</i>		16		11		14	

quite a minimal gesture form. For both speaker groups, gestures in spoken perfective contexts had the fewest agitated motion strokes, or, strokes expressive of MANNER, of all gestural expressions of motion events. Chi-square tests were run on the count data presented in Table 4b that compare PATH-only gestures to gestures of all other compositions. For both languages the test values were statistically significant, indicating that

the tendency to produce the simple PATH-only gesture forms is not independent of aspect category in the accompanying speech; Mandarin, chi-square (2) = 15.860, $p < .000$ and English, chi-square (2) = 16.284, $p < .000$.

Spoken Progressive Aspect

Table 4 shows that in contrast to the single-vector, non-extendable stroke gestures typical of perfective-marked speech contexts, 84% of gestures accompanying spoken progressive aspect in Mandarin and 88% in English were agitated-stroke ‘extendable’ gestures. In the case of the expression of motion events, the data summarized in Table 4b, agitated-motion strokes may often be interpreted as the gestural expression of the MANNER component. Of motion event-expressive productions summarized in Table 4b, gestures in progressive-marked contexts show the most MANNER-marking of the three aspect categories. Calculating on the basis of the figures in Table 4b, we see that in Mandarin, a total of 94% of motion event-expressive gestures have agitated-motion strokes; 91% of the same subset in English have such strokes. The corresponding proportions of agitated-motion strokes in perfective- and durative-marked contexts was zero and 50%, respectively, for Mandarin; 6% and 61% for English.

As is the case in perfective-marked speech contexts, among two-handed gestures in progressive-marked contexts, the mirror-image configuration predominates. Table 4a shows that the hands are mirror images of one another in 100% of the two-handed gestures in progressive-marked contexts among Mandarin speakers. This is true in 82% of such cases among the English speakers.

Again, there is proportionally less Element-marking in progressive as opposed to durative spoken contexts, based on the subset of non-pantomimic gestures summarized at the bottom of Table 4. Of these, 8% of the Mandarin speakers’ and 12% of the English

speakers' gestures in progressive aspect contexts are Element-marking, compared to 20% and 24%, respectively, in durative aspect contexts.

Also, as is the case in perfective contexts, the overall incidence of post-stroke holds is low in progressive contexts. Of the 25 gestures accompanying such contexts in Mandarin that were examined, 8% have post-stroke holds of the 'syntactic' type; 4% of the 'semantic' type. Of the 25 such English gestures, again 8% of strokes in progressive contexts are followed by syntactic post-stroke holds and none in English have semantic post-stroke holds.

The example speech-gesture productions below illustrate some features common to gestures in progressive-marked speech contexts. For comparison, a contrasting example is given of a gesture that accompanied perfective-marked speech. Example (16) was a response elicited in the vignettes task:

(16) a construction m[an is **rolling across** the screen /]

The progressive-marked speech here is accompanied by a rather elaborate gesture in which the speaker's forearm, representing the FIGURE, moves in an orientation perpendicular to his torso, repeatedly looping while moving across the speaker's gesture space in a PATH and MANNER similar to that of the FIGURE's rolling motion in the eliciting stimulus. This is typical of the many cases in which an extendable, agitated-motion gesture stroke iconically represents the MANNER component of a motion event.

Note however, that the gestural feature of agitation is found not only to accompany the spoken expression of motion events as an expression of the MANNER component. It emerges elsewhere as well, in gestures accompanying progressive-marked verbs of *stasis* in speech; for instance, 'sit,' 'listen,' or 'stand,' or mental state verbs such as 'worry.' MANNER of motion cannot be part of the lexical semantics of such

verbs. That the gestural agitation feature emerges in such contexts means that it is not, as might be inferred from the speech-gestural expressions of motion events, simply doing the work of iconically expressing MANNER. It appears instead, on the basis of such evidence, to be in part a reflection of the progressive aspectual view itself.

This is another example from a Vignettes Task participant who says,

(17) [a ruler **was standing up and then**] [fell over]

The underlining with bold face type indicates a kind of hold-stroke, however in this case, the ‘hold’ is not completely motionless. The speaker holds her forearm and flat hand vertically in front of her body and *shakes* it slightly while saying, “was standing up.” In the next bracketed gesture phrase, the hand and forearm drop down in an arc trajectory to a horizontal position, pivoting on the stationary elbow joint.

Similar examples come from speakers narrating a scene in the cartoon where Sylvester the cat, partially concealed in a hiding place, listens for information about the bird. Those who mention this scene provide evidence both for how flexible speaker choice of aspectual view is, and also for the degree of abstraction of its gestural expression. In the eliciting stimulus, the act of listening is depicted by the cat perking up an ear and holding it steady while listening for the information he wants. There is no movement associated with the listening in the stimulus. As in real life also, the listening pose is one of complete stillness. Yet some speakers re-tell this cartoon event by saying, “he’s listening,” while at the same time slightly shaking a cupped hand, palm forward, next to one of their ears.

In contrast, one of the Mandarin speakers chose a different view of this event. Rather than refer to the period of time during which Sylvester holds still and listens, this speaker focuses on the very instant that the cat hears the information he has been listening for. She says,

- (18) ta ting-dao le
 他聽到了
 he listen-CMPL PRF
he hears this

This example was used earlier to illustrate the use of the completive *Aktionsart* marker ‘dao’ to create a RVC structure. Such a usage emphasizes that the focus is on the very instant that the cat’s ear catches the information. The speaker expresses an external view of this event by choosing perfective aspect. Synchronized with the RVC structure, the speaker’s hand makes a very rapid, single-stroke ‘snatching’ motion next to her own ear. Her open hand snaps shut as though grabbing hold of something, a gestural metaphor for ‘catching’ a sound that, in its single-stroke execution, is very different from the gestures produced by speakers who opt for an imperfective-progressive view of the same situation. Such speech-gesture productions illustrate how features of gesture execution that reliably occur in spoken aspect contexts influence the nature of the iconicity of gestural representation. In such cases, motion emerges as part of the gestural expression of an event that itself has no inherent motion.

A second way that aspect-related features of gesture form influence the iconicity of gesture is illustrated in cases where a speaker produces a repeating form of agitated motion stroke that is in explicit *contrast* to the MANNER of motion actually depicted in the eliciting stimulus. These are cases where the event referred to does have inherent motion, yet the gestural expression of its MANNER is a poor match for what the speaker saw when watching the event. One Mandarin speaker for example, when describing a vignette that depicted a ruler sliding smoothly and uni-directionally across a surface, chooses a progressive aspect frame as follows:

- (19) you yi-ge chi shi zai zhuo-zi -mian / zai dong
 有一個尺 是在桌子面 / 在動
 have one-CL ruler be on table surface / PRG move
there's a ruler (and it's) moving on the surface of a table

The speaker accompanies the progressive-marked portion of this utterance with an agitated motion stroke that could have been interpreted as depicting an object moving back and forth. In progressive spoken contexts, there are many instances where the accompanying gesture bears this feature of agitation, even when there is no corresponding motion in the stimulus.

Spoken Durative Aspect

As for the gestures that speakers produce to time with durative-marked speech, what is most noticeable in both languages is a sustained quality of the gestures, expressed either as an extendable, sustained motion stroke, or a long hold phase that may extend over more than one spoken clause. Extendable strokes comprise 76% of the 25-gesture Mandarin sample; 92% of the same English sample. A proportion of the extendable-stroke gestures that accompany durative-marked speech, eight of nineteen in Mandarin and twelve of twenty-three in English, display the 'agitation feature' described in the previous section as well, close to half of the total set. Collapsing across the two languages, 37% of all the extendable stroke gestures that occur in durative-marked spoken contexts have the agitation feature. A 95% confidence interval for this proportion was formed, $.241 < p < .499$. The lower bound is greater than zero, so the proportion is non-negligible. About these agitated-motion strokes that occur in durative-marked speech contexts, it is interesting to note that roughly half of them co-occur with speech that embeds some kind of overt progressive marking within a durative-framed utterance. Example (20) illustrates this often-encountered pattern.

(20) [but while **the chase is going** o][n Alice is sitting there worrying

‘While’ was taken to signal durative aspect. The bolded portion of the bracketed utterance, “ile the chase is goi,” shows the extent of an agitated motion stroke.

In both Mandarin and in English, 40% of the gestures accompanying durative-marked speech are followed by post-stroke holds of the ‘semantic’ kind described above. These holds are defined as sharing an interpretable semantic relationship with the accompanying speech. Such holds often extend across marked syntactic boundaries and as such do not appear to share the kind of timing relationship with utterance structure that Kita (1990) described. In contrast, only 4% of the gestures in perfective or progressive contexts in the Mandarin sample have semantic post-stroke holds, and none of the gestures sampled from the English perfective and progressive contexts do.

As was mentioned earlier in the section on spoken perfective aspect and gesture, two-handed gestures occur in higher proportion in durative-marked speech contexts than in the other two aspect contexts. Table 4a shows that, in the Mandarin sample, 72% of the gestures accompanying durative speech contexts are two-handed. In English the proportion is 80%. Results of the exploratory pilot study had led to the additional expectation that the majority of two-handed gestures in spoken durative contexts in both languages would be of the ‘contrast’ type. Two-handed contrast gestures are those where one hand represents one entity or action, the other hand another, and the two stand in relation to each other somehow. Contrary to this expectation, contrast gestures were the minority of two-handed gestures, in the Mandarin durative sample (29%, or five of a total of eighteen productions) and comprised only half of the English sample (ten of a total of twenty). Also note here that each of the two two-handed contrast gestures observed in English progressive contexts (see Table 4), accompanies a progressive-marked utterance that was embedded in a durative-marked framing utterance.

Thus overall there appears to be a greater tendency among English speakers than among Chinese to produce two-handed contrast gestures in durative-marked speech contexts. This is one factor where the data reveal a possible difference between the gesture patterns of the two speaker groups. Table 4 shows that, whereas for English speakers the proportion of two-handed contrast gestures is larger in durative, compared to perfective and progressive spoken contexts, in Mandarin this is not the case. In fact, comparing the Mandarin samples of two-handed gestures in durative and perfective contexts summarized in Table 4, we see that the proportion of contrast gestures accompanying perfective-marked speech is actually slightly higher than that accompanying durative-marked speech: 36% versus 29%.

A scan of the samples of Mandarin durative-marked speech, however, suggests a possible reason for this difference, one that may be simply happenstance. In the Mandarin sample a relatively higher proportion of the durative-marked verbs netted by the sampling procedures described in the Methods section are those having to do with holding symmetrical objects than was true in the English sample, for example, holding binoculars, the bird cage, or the bowling ball. Gestural depictions of such actions as these could be expected to elicit two-handed mirror iconic gestures. Another source for this apparent difference in gesture production between the two speaker groups could be the differing proportions of motion event expressions in the Mandarin and English durative samples. Table 4a shows that half as many of the sampled Mandarin durative utterances (ten of a total of twenty-five) are expressions of motion events as in the English sample (twenty of twenty-five). These two sources of variation taken together mean that a larger proportion of the Mandarin durative sample is expressive of events of stasis that involve holding symmetrical objects. This sampling outcome would yield fewer opportunities for the Mandarin than for the English speakers to demonstrate gestural encoding of two contrasting entities or events.

Element-marking is more common in gestures that accompany durative-marked than perfective- or progressive-marked speech, as was outlined in the previous two sections. Further, there is reason to think that the proportions of element marking observed in the extracted sample of durative-marked speech examined for Study One are underestimates; an artifact of sampling only verb-related, single gestures. The example below is an instance of a kind of element-marking that appears to be quite common in durative-marked spoken contexts, but that would not have been reflected in the Study One samples.

(21) [/ you know [**the stre**]et / that he has to swing across on this rope]

Superimposed on the preparation phase of an iconic gesture that depicts swinging across a certain distance is an additional gesture that briefly indicates the endpoints of the path, the locations of two buildings in a scene in the cartoon. Example (21) illustrates a phenomenon whose incidence cannot be estimated because no specific measurements were taken of it for this study. The data give the impression that gestures accompanying durative-marked speech make more extensive and cohesive use of gesture space than those found in the other two spoken aspect contexts. That is, within one speech-gesture production envelope whose scope appears to be tied to a particular durative-marked expression, there are often additional superimposed, or co-produced, gestures expressive of additional events and/or entities. Not only is more sometimes encoded in terms of elements and actions at the level of the individual gesture, but also such gestures appear more often to reference the spatial metaphoric relationships established in previous gestures, and in general to extend and maintain more complex spatial arrangements.

Summary of the Gesture Profiles

Distinctions in gesture form and execution appear to be associated with the three aspectual distinctions examined here. In general, the gestures of Mandarin speakers and English speakers manifested a similar range of distinctions. The majority of gestures sampled from spoken perfective contexts in both speaker groups are short and simple in execution. Often, in the case of motion event expressions, they express only a reduced PATH trajectory. The gestures sampled from spoken imperfective contexts are typically longer and more complex. A large number of the gestures in progressive contexts have agitated motion strokes that often combine with trajectories in the expression of motion events to create elaborate PATH-plus-MANNER productions. Typical of spoken durative contexts are gestural holds of the ‘semantic’ kind, as well as somewhat greater gestural complexity. The category of durative-marked speech and accompanying gesture is where a difference is observed between the two speaker groups, but this may be an artifact of somewhat different classes of event expressions having been sampled from the groups.

Layering of Aspectual Views

A durative aspect frame permits foregrounded actions to be distinguished in speech from co-occurring backgrounded actions. As can be seen from the samples listed in Appendix A, in speech there is in fact a tendency in the presence of durative marking to subsume more event information within what appear to be single speech expository units or ‘production envelopes.’ A given durative-marked event expression may have a structuring role in relation to one or more additional event expressions, and this is often mirrored in the gesture that accompanies these expressions, where we can see layering of the features of gesture execution that the foregoing analyses have shown may be related to spoken aspect. Consider the following example from a Mandarin narration, in which a

gesture hold-stroke, typical of durative-marked contexts, is further overlaid with some ‘agitation’ movements of the sort that are typical in progressive aspect-marked contexts:

- (22) [ta yong hen da-de wang-yuan-**jingding-zhe ta kan** /][ran-hou
 他用 很大的望遠鏡盯著他看 / 然後
 he use very large-MOD telescope stare-DUR him look / next
he uses a very big telescope to stare at him, and

jiu ^kan-zhe ta kan-^zhe ta ra^an-hou /]
 就看著他看著他然後
 just watch-DUR him watch-DUR him next /
then he’s watching and watching him, then ...

Here the speaker sets up a framing event in speech that is the cat staring at the bird through binoculars. The verb in speech, ‘ding’/‘stare’ has the durative marker -ZHE attached. Superimposed on the accompanying hold-stroke is a small, repeating bi-directional movement away from, and back toward, her face. This small gestural agitation may represent the ‘eidola’ (‘rays of sight’) imagined to be emanating from the cat’s eyes (McNeill, 1992:156). The verb reduplication in the second statement of “kan”/“watch” is an instance of the periphrastic expression of progressive (as it is often analyzed) aspect described earlier. This is an example in which aspectual views layered in speech are similarly layered in gesture. Many such instances were observed in durative-marked speech contexts of layered gestural expression of events and their associated aspectual views.

Aspect and Metaphor

The gestural indices of aspect outlined here are not confined to iconic gestures depicting concrete events and actions. Metaphoric gestures also display the same range of features, and there are several instances of this in the data examined here. For

example, in a movie narration, a speaker describes the protagonist's troubled state of mind by saying that she's, "sitting there worrying." This is accompanied by a gesture in which the speaker's two hands rotate around one another alternately, an agitated-stroke gesture expressive of some kind of processing metaphor.

Other examples of aspect-marked metaphoric gestures are those found at episode boundaries and at the conclusions of narrations. At such narrative junctures, speakers often make some metanarrative statement to the effect of, "that's it for that part," "that's how it ends," or a simple paranarrative/interactive, "OK?" with an inquiring look at the listener. In Mandarin, such statements usually carry the perfective aspect marker. In speakers of both languages, these metanarrative statements are frequently accompanied by small gestures that appear quickly to sweep something away at the edge of gesture space, or by conduit metaphoric gestures that open up in a quick, single stroke, appearing to release their contents.

Aspect-linked gesture features that occur in non-aspect-marked contexts

Some of the formal features of gesture outlined above that appear to covary with spoken aspect context are observed also to occur in aspect-neutral contexts, or contexts at least where aspect is not overtly marked locally. Recall that it is not obligatory to mark aspect in either spoken Mandarin or spoken English. Of course, no instances of this kind of 'aspect-linked' gesture that is observed in the absence of aspect-marked speech are included in this study's data set. In order for any utterance-gesture pair to be considered for inclusion in the data set for Study One, aspectual view had to be overtly expressed in speech.

An example of gesturing of this kind that occurs with some frequency in the cartoon narrations of both the Mandarin and English speakers is when a speaker, introducing the scene in the cartoon that opens with an organ grinder playing his

instrument, says, “there’s an organ-grinder down in the street” accompanied by an extendable, agitated-stroke iconic gesture representing the act of turning the instrument’s crank, in the style of gesture typical of progressive-marked contexts. A Mandarin speaker, getting a detail wrong, produced the same kind of speech-gesture combination, saying in Mandarin, “the old man there plays violin.” The verb in this utterance is not marked for aspect. One of the metaphoric gestures mentioned in the previous section is another example. This is the speaker who concluded narrating one cartoon event by saying to his listener, “OK?” and making a small sweeping motion away from his body at the edge of his gesture space.

On the basis of speech in such unmarked contexts, the only way to infer the narrator’s aspectual viewpoint would be with reference to some general ‘tense-aspect frame’ as outlined by Slobin and Berman (1994). The question of whether such features of gesture may index aspectual view in the absence of overt marking in speech is taken up in the Discussion.

Speaker Choice of Aspectual View

Finally, evidence of the degree of flexibility that a speaker has in whether to take the expanded, internal view of an event expressed by imperfective aspect, or to collapse the event expression with perfective aspect, was everywhere apparent in the narrations that were examined for this study. Across speakers, different choices are made in identical event contexts as to which aspectual view to apply. It is the exception rather than the rule for particular individual events in the eliciting stimuli to preferentially trigger one aspectual view over another.

An illustration of this flexibility on the part of the speaker comes from a set of data that was not actually used in the present study, but it is especially compelling for a

particular reason. The example is drawn from data collected for another study¹⁵ in which participants were asked to narrate portions of the cartoon story twice, to two different audiences. With such an elicitation, there is the opportunity to observe the flexibility of aspectual view within-subject, across two tellings of the same event content. In one case, an English speaker narrates the scene where the cat rolls down a hill, propelled by a bowling ball that he has swallowed. In the first re-telling, the speaker describes this event using a progressive-marked verb saying, “the cat is rolling down this hill.” The co-occurring gesture has the extendable, agitated-motion stroke phase encoding both MANNER and PATH that is typical of such spoken aspect contexts. In the second re-telling however, the speaker chooses to collapse this event saying, “it sent him down the hill.” The accompanying gesture in this case is a quick, single-stroke path trace gesture of the sort typically seen in perfective speech contexts.

Discussion

According to Klein (1994), “all languages have devised means for the expression of this fundamental domain of human reasoning, temporality.” Both Mandarin and English have the means to encode distinctions of aspect and *Aktionsart*. The gestural evidence examined for Study One allows us to consider whether the differences between the two languages in their systems of encoding these temporal distinctions have an impact on thinking-for-speaking. The results presented here suggest that there are consistencies in gestural patterning related to aspectual view that seem independent of specific linguistic expressive options. Conceptual representation of the internal temporal constituencies of events during real-time language production appears to be largely

¹⁵These data are from a study by Asli Ozyurek, designed to determine the effect of narrating to one listener versus two on the way a speaker organizes her gesture space. I thank Asli for bringing this example to my attention.

unaffected by the surface differences between Mandarin and English in the linguistic means used to encode aspect and *Aktionsart*.

In what follows, some points concerning particular findings of Study One are taken up. Then some more general issues concerning aspectual view as represented in speech and gesture, and concerning the nature of linguistic aspect itself, are discussed.

Stroke durations

It should be stated that the differences observed in gesture stroke durations across aspect categories cannot be related to any differences in the time it takes to utter the portions of speech with which the strokes are co-expressive. First, although perfective -LE and durative -ZHE are both post-verbal morphemes, durative -ZHE is constrained to follow immediately the verb it marks, while perfective -LE may occur some distance away at the end of an utterance. Gestures most typically time with speech with which they are closely related in semantic content, and the verb-related gestures sampled for this study indeed often accompany the verbs of their utterances, as the bold face portions of speech in Appendix A show. To the extent that this is true, gestures accompanying -ZHE-marked speech could, even if they were of short duration, in every case extend to co-occur with both the verb and its aspect-marking particle. However, this would be the case in perfective -LE-marked speech only some of the time. The rest of the time, in order for stroke phases to cover the entire -LE-marked utterance, they would have to be of longer duration. Thus, the facts of association of verbs with these different aspect particles, if stroke durations were tied to the duration of utterances, would predict *longer* stroke phases in perfective -LE-marked spoken contexts than for durative -ZHE-marked contexts. The results reported above show that the reverse is true.

Further, although no analysis of the specific synchrony between verb-related gesture strokes and particular constituents in the accompanying speech was performed for Study One, a scan of the speech-gesture pairs listed in Appendix A shows that stroke phases can turn up in a variety of places relative to the spoken utterance. That is, they are not necessarily roped to the verbs and their associated aspect particles. Depending on the nature of the GP in any one instance, any of a number of alternative timing relationships may emerge. This makes the issue of the ‘time it takes to utter the associated speech’ specious, since the time it takes will always be exactly the duration of the gesture stroke. Since strokes have flexibility in terms of what they accompany in the utterance, and although they often co-occur with verbs and their aspect particles but this is not always so, there appears to be no way to meaningfully relate the Study One’s observations concerning gesture stroke durations with the durations of stretches of spoken expression.

The durative difference

A difference was observed between the Mandarin and English samples in the stroke durations in progressive and durative aspect contexts, and in the proportions of two-handed contrast gestures in durative contexts. A single factor in the composition of the samples from the two languages may account for both of these observed differences.

In regard to the latter, the exploratory pilot study had led to the expectation that both speaker groups would have a high proportion of the complex two-handed gestures in the context of durative-marked speech. That gestures in durative contexts might sometimes depict more than one event is predicted by the semantics of durativity, as discussed by Slobin and Berman (1994:5) and Sheng Ma (1985). A duratively-construed event can form the background for another, foregrounded event. As was mentioned, although the narration data were not systematically examined in this regard, they give the impression that gestures accompanying durative-marked speech make more extensive

and cohesive use of gesture space than those found in the other two spoken aspect contexts. That is, within one speech-gesture production envelope whose scope appears to be tied to a particular durative-marked expression, there are often additional superimposed, or co-produced, gestures expressive of additional events and/or entities. Instances of the production of two-handed contrast gestures in support of this density of encoding are attested in the samples of both speaker groups. Their proportion was lower in the Mandarin sample, however.

In the Mandarin sample of durative-marked speech, a higher proportion of the verbs were those having to do with holding symmetrical objects, for example the bird cage or binoculars, than was true in the English sample. Gestural depictions of such actions as these could be expected to elicit two-handed mirror iconic gestures. This higher proportion could be happenstance; a consequence of a sampling fluke that netted a narrower range of event expressions in durative aspect contexts in the Mandarin sample than in the English. Alternatively, it could also be a function of the restrictions on the classes of verbs to which the durative marker may attach. Recall that Mandarin restricts application of the durative aspect particle to verbs of particular *Aktionsarten*. The English sample of durational verbs and durative aspect periphrastic constructions would not have been subject to such a restriction.

The differing composition of the Mandarin and English durative aspect samples likely accounts as well, at least in part, for the fact that in the imperfective aspect sample there was a statistically significant difference in mean stroke durations in English progressive spoken contexts versus durative ones, but none in the Mandarin sample. Given that the Mandarin durative sample is skewed toward expressions of unitary events accompanied by gestures that correspondingly show a lesser degree of complexity, it makes sense that stroke phases would on average be executed more quickly.

Further research is needed to sort out such issues. One way the procedure needs to be altered is in terms of the criteria for speech sampling. The sample from the Mandarin data was limited to grammatically-marked aspect signalled by the presence of one of the three aspect-marking particles, while the English sample leaned more toward lexical and periphrastic expressions that prejudice particular aspectual views. This means that the two samples are not strictly comparable, in potentially significant ways. A way to address this would be to broaden the Mandarin sample to include a set of utterance-gesture pairs more analogous to the types sampled from the English. The Mandarin data should present no barrier to such a procedure. As was outlined, Mandarin has other means, in addition to its set of aspect-marking particles, to express the target distinctions of temporal contour, and a variety of these means is observed in the narration data. These could be sampled and subjected to the analyses described above.

One change in procedure that has been suggested, but which the data would not support, is an exhaustive, clause-by-clause analysis of aspectual view, with examination of the accompanying gestures. However, because aspect marking is not obligatory in either Mandarin or English, it is often not possible to infer with any certainty the view of events that a speaker takes within every single utterance frame. Therefore, the analyses pursued here are only possible on the basis of speech where the relevant distinctions are in some way overtly signalled.

The nature of iconicity in gesture

The findings of Study One concerning aspect-related features of gesture form make it necessary to think about the nature of iconicity in gesture, and ask how features of gesture may relate to a more abstract level of conceptual or linguistic organization. A witnessed motion event, such as a cat rolling down a hill with his legs flailing, may be collapsed to an 'iconic' gesture that encodes nothing but a reduced PATH, or may instead

be more elaborately encoded in a gesture that preserves many more of the features of the event. At a minimum, a phenomenon such as the ‘agitation feature’ calls into question claims that gesture comes from visual imagery via a “direct route” (Hadar and Butterworth, 1992); that it is, “the motor manifestation of imagistic activation,” as these authors claim. The findings suggest that the tendency for gestures with specific features to co-occur with certain spoken aspects is very strong, because this co-occurrence is possible even when it violates the iconicity of the gestured depictions of events. Gestures in spoken perfective contexts typically collapse event expression by encoding a minimum of detail, while those that occur in spoken progressive contexts sometimes add a dimension of movement, agitation, that is not present in the stimulus. Such instances of gestural diminishment versus augmentation of the ‘image’ are evidence that features of production such as rapid movement strokes and the agitation feature are not always or simply iconically expressive. What shows up as MANNER-marking in many motion gestures may in fact be an overlay on a more fundamental, abstract, non-iconic feature that derives from a kinesthetic representation of a state of being, or an action in progress, as something that is somehow ‘agitated’ or ‘enlivened.’ In choosing a progressive aspectual view, a speaker manifests her sense of a point in an ongoing process. The data show that a gesture that accompanies such speech will likely incorporate the agitation feature. Where the referent is something that may be expressed with a MANNER-encoding verb of motion, the agitation feature may achieve expression in iconic MANNER marking. Where the verb is one of stasis, however, the feature may show up as a completely non-iconic shake or wiggle. The existence of such patterning, and of its opposite— the unelaborated, non-extendable gestures that occur in perfective-marked contexts — are evidence that aspectual view is a powerful determinant in the content and structure of expression, gestural and otherwise. Aspect itself appears to set a significant

part of the framework for iconicity.

The possibility of indexing view-of-event in gesture alone

It has been pointed out that there are many instances in the narration data where gestures with the same formal characteristics as those found in aspect-marked speech contexts occur in the absence of spoken aspect marking. For instance, they may occur in the presence of English simple present or in instances where gesture expresses some action being carried out by a character, even though the action is not expressed in speech. This is the case in those instances where a speaker says, “There’s an organ grinder in the street,” and accompanies this with an agitated-stroke gesture iconic for the motion of cranking an organ grinder’s musical instrument. Such cases make it reasonable to consider whether, for languages such as Mandarin and English, in which aspect marking is not obligatory, gesture alone may sometimes convey the aspectual content of a conceptual representation. It may not be possible to judge definitively, but the features of gesture execution that appear on the basis of the results reported here to be related to aspectual view, may in some instances signal the presence of a distinct aspectual view of the event in conceptual representation, one that is not encoded in the accompanying speech. One approach to exploring this possibility would be to coordinate such gestural observations with analysis of their narrative contexts in terms of ‘tense-aspect frame’ as discussed by Slobin and Berman (1994).

The description in the preceding chapter of the different ways in which speech and gesture may be ‘co-expressive’ shows why it is reasonable to consider this possibility. A high degree of redundancy across the two modalities is common; for example, “he goes [**up**],” where the word “up” is accompanied by an index finger point trace of an upward path. However, at least some degree of speech-gesture complementarity is typical of many productions as well. An example of this is when a

Mandarin speaker says (according to the syntax of that language), “th[is cat] TOPIC MARKER on the big street rolls around,” where the gesture that times with “cat” in speech is a motion iconic expressing ‘roll around.’ Two quite different components of the event representation are thus manifested in a conjunction of speech (FIGURE) and gesture (MOTION). In this instance, the feature of motion expressed in gesture is later unpacked explicitly in the speech, “rolls around”; however, often such further unpacking into speech does not occur. In such cases, the gestural encoding of a given event component that occurs in complementary conjunction with an element of speech, remains the only expression of that component within the utterance. Although aspectual view, as a component of event representation, is different in kind from FIGURE, PATH, or MANNER, there seems to be no reason to exclude the possibility that its gestural manifestations could combine with speech not linguistically marked for aspect, in the same way that other components of events with complementary meanings are conjoined in speech and gesture.

One application of this approach suggests itself in the search for an explanation for what were referred to above as ‘Butterworth beats.’ Recall that these are agitated, repetitive movements of the hands that sometimes accompany pauses and speech dysfluencies, when these result from lexical retrieval difficulties. Recall also that all speech of this kind was systematically excluded from the present study, as it was suspected that the ‘agitation feature’ in such cases might have some different source than the gestural agitation that can accompany fluent, aspect-marked, spoken expression of narrative-level events. The intention in excluding dysfluent speech was to avoid confounding two potentially distinct sources of similar patterning in one dimension of gesture.

As an exercise in taking a more comprehensive view of the distribution of such gestural patterning, one can ask whether these Butterworth beats may be a manifestation of something at another level of discourse structure that is abstract but that shares some conceptual domain with progressive verb aspect. Since it is clear that gesture patterns in association with more than just the narrative level of analysis (McNeill, 1992; Levy and McNeill, 1992), it may be reasonable to consider Butterworth beats a kind of metapragmatic expression of progressive aspect, related to conversational turn-management, similar on some level to the movements of the same sort that appear in progressive-marked narrative contexts. On this view, lexical search beats are the external metaphor of some kind of repetitive processing activity associated with lexical search, and make manifest the fact that the current disruption of language production is only temporary; that the speaker will resume speaking.

Linguistic aspect as a heterogenous category

Study One has proceeded with something like an assumption that verb aspect is a unitary category; one linguistic feature with multiple settings, three of which were examined here. Yet the results presented here, as well as surveys of the wide range of phenomena that have been brought together under the term 'aspect' (cited in Binnick, 1991), lead to the conclusion that, as an analytic entity, aspect has very broad reach; a reach not limited to considerations of the internal temporal constituencies of particular events labelled by verbs.

The earlier section on the nature of iconicity discussed the possibility that aspectual view is implicated in the structure of 'imagistic' representations. It seems clear as well that the influence on event conception of this conceptual-linguistic category is broader than the limited domain suggested by the label '*verb*' aspect. Choice of aspectual view appears to have consequences for the structuring of whole units of

discourse, determining how event descriptions will be situated within a discourse framework. A notion that aspectual view is something tied to the scope of a single verb captures only a portion of the patterning that appears related to this conceptual-linguistic distinction. Aspect appears to be a more comprehensive zone of patterning than the analyses that have focused on overt verb markers of it in different languages would suggest.

The data examined for Study One contain many speech-gesture productions that illustrate how choice of aspectual view is manifested in gesture's system of spatial and temporal coordinates. Example (23), below, is such a case. The speaker here described a rather complex scene involving two moving FIGURES and a GROUND component that plays a significant role in how the two FIGURES move in relation to one another. The underlining in the first portion represents the perseveration of a spatial configuration set up in the preceding phrase that is relevant to the motion event to be expressed next.

- (23) ran-hou ne zhei dian-che nei-ge* nei-ge gan ne
 然後呢這電車那個* 那個 gan 呢
 next NE this train that-CL* that-CL pentagraph NE
then the train and that that pentagraph*

jiu zhe-ma gen-zhe gen-zhe ta jiu shi-jin pao ...
 就這麼 跟著 跟著
 他就使勁跑
 then like.this with-DUR with-DUR him just furious run
then follows along with him as he runs furiously

The features of this complex gestural display are those typical of durative-marked contexts: a gestural hold reflecting a cohesive use of gesture space, a two-handed contrast use of the hands, and a sustained motion stroke. What is interesting is that the durative aspect marker itself does not emerge in speech until a point at which the speaker has already for some time had the gestural configuration with which it is co-expressive

activated. Thus, there is a sense in which the overt grammatical durative aspect marker may convey only a fraction of the temporal extent that the event being expressed has in conception, as evidenced by the very extended gestural accompaniment. Such are instances where the patterning of gesture in relation to speech provides an ‘enhanced window’ on the conceptual representations that participate in language production. We see that aspect participates in defining a broader domain of reference than may be inferable only on the basis of its interactions with verbs in spoken utterances.

In regard to the durative aspect-marking particle -ZHE in Mandarin, and taking a broad view of aspect, it is interesting to reflect as well on the range of lexical items to which it can attach. Consider the syllable to which -ZHE attaches in example (23), above. Sometimes classified as a co-verb, sometimes as a preposition, ‘gen’/‘with’ is basically a locative or spatial relational particle, one of a class in Mandarin that includes such items as ‘xiang’/‘facing,’ ‘yan’/‘along,’ ‘bei’/‘behind,’ ‘dui’/‘toward,’ and ‘ai’/‘adjacent to.’ The durative aspect particle may attach to any of these and does so in clauses where there is usually another lexical item that functions as the main verb of the clause. The following two examples from the Mandarin narrations illustrate this:

- (24) ta bei-zhe shou zou-lai-zou-qu
 他 背者 手 走來走去
 he behind-DUR hand walk-come-walk-go
he is pacing with his hands behind his back
- (25) ta yan-zhe nei-ge pai-shui-guan pa-shang-qu
 他 沿者 那個排水管 爬上去
 he along-DUR that-CL drain-water-pipe climb-up-go
he climbs up along the drainpipe

In a general statement meant to apply to the entire class of coverbs in Mandarin, of which the spatial relational particles are a subset, Li and Thompson (1981:360) state,

“That some coverbs can occur with verbal aspect markers is explained by the fact that they used to be verbs: a coverb that can take an aspect marker has progressed less far along the historical route from verb to preposition than one that cannot.”

Given an *a priori* assumption that anything to which a durative aspect particle may affix must be of the grammatical form class ‘verb,’ the above statement seems straightforward. Yet there seems to be no reason not to consider the possibility that such affixation in the case of spatial relational particles is instead evidence that the domain of linguistic aspect is not limited to interactions with verbs; rather that some of its functions link in fundamental ways to spatial cognition. Many authors note, for instance, that what serves as the progressive aspect morpheme in many languages is a grammaticized spatial preposition (see Li, 1991, for a discussion of this in Mandarin). The term ‘aspect’ itself encodes a visuo-spatial metaphor. It and the spatializing gestural reflexes of spoken aspect distinctions invite a broader interpretation of this category.

Because the range of phenomena associated with linguistic aspect is wide, and because it has proved problematic to port analyses of aspectual distinctions among languages — analyses of the distinction between perfective and imperfective, for instance — some linguists recommend that such analyses, and the notion ‘aspect’ itself should be reserved for the temporality distinctions that are marked on verbs in Slavic languages, analysis of which gave rise to the term in the first place. Bache (1989, cited in Klein, 1994) advocates this very narrow treatment of the category of aspect. At the other extreme are theorists who would see aspect as part of a broad discourse pragmatic framework (see discussion in Binnick, 1991:446 ff).

The view through the gesture-enhanced window that highlights, for example, the evident discourse scope of aspect choice, its determining role in the nature of iconicity,

and the possible consistency of patterning across varying linguistic contents and contexts, recommends those theoretical treatments of aspect that broaden rather than restrict the framework in which we consider its functions. As Binnick (1991:446) states,

“Clearly, tenses and aspects have distinctive [discourse] pragmatic functions ... there is some point to the specification of temporal relations *beyond the mere indication of temporal relations themselves*. ... They serve to structure discourses by foregrounding and backgrounding information, by providing the logical structure of narrative, or indeed by indicating the very nature of the discourse as narrative or something else. They serve to glue events together into sequences of events or to indicate their independence from one another. ... They interrelate crucially with the systems of marking indicating the information flow through the discourse or text. They invite inferences and provide overtones.”

CHAPTER 4

STUDY TWO: GESTURAL INDICES OF UTTERANCE DYNAMICS

Introduction

Within the spoken utterance, features of different languages such as standard word order, as well as mechanisms for distinguishing arguments of the verb and for the management of discourse topic, affect the flow and organization of information in spoken expression. In most linguistic analyses, Mandarin Chinese, like English, is said to follow a basic SVO sentence plan. Li and Thompson (1976, 1981) claim that Mandarin and English differ, however, on the linguistic typological dimension of topic prominence (see also Tsao, 1990; Comrie, 1985, 1994). These authors and others have argued that Mandarin is a ‘topic-prominent’ language, as opposed to English, that is ‘subject-prominent.’ In existing analyses, topic versus subject prominence is determined by the preponderance of certain types of utterance structures in ordinary discourse. Topic-prominent languages are said to have a preponderance of sentences that begin with a statement of a TOPIC that is followed by a COMMENT; the latter being speech that expands on, or in some way follows from, the content of the TOPIC. In general, the SUBJECT-PREDICATE relationship central to utterance formation processes and to discourse cohesion in a subject-prominent language is thought to be more peripheral in the grammar of a topic-prominent language.

To get a better sense of the essential contrast between a TOPIC-COMMENT utterance and one that requires a specified SUBJECT-PREDICATE relationship, consider the following Mandarin example:

- (26) nei-ge xue-sheng zhong-guo hua shuo-de bu-cuo
 那個學生 中國話說得不錯
 that-CL student china language speak-MOD not bad
 TOPIC COMMENT
 (As for) that student, (her) Chinese (is) spoken pretty well

Even though a couple of different English sentences capture the sense of example (26), it is problematic to devise an ‘accurate’ translation equivalent of it. The reason for this is that the Mandarin sentence specifies nothing that may be interpreted as a SUBJECT-PREDICATE relationship. In addition to the English gloss given above, other reasonable equivalents would include, ‘that student’s Chinese is pretty good,’ ‘that student’s Chinese is spoken pretty well,’ or ‘that student speaks Chinese pretty well.’ There may be others as well.

Chafe (1976) and Xu and Langendoen (1985) discuss two types of TOPIC-COMMENT constructions that occur in Mandarin, the ‘English style’ and the ‘Chinese style.’ According to these authors, the English style of topicalization is one in which the topicalized constituent is identical to one that occupies a role in an underlying SV(O) sentence. The occupant of this role may be overtly expressed again in the COMMENT that follows the TOPIC, as in example (27) *a*, or may be ellided as in (27) *b*. In examples (27), the SUBJECT of what is interpretable as a SV sentence is topicalized. In (27) *a* this constituent is repeated in pronominal form in the COMMENT:

framing the utterance in this way does not make sense in the narrative context.

It is hoped that the above comparison makes clear the difference between Mandarin and English utterances that analyses of Mandarin as a topic-prominent language have attempted to capture. This is that specifying a SUBJECT-PREDICATE relationship seems central to utterance formation in English, but not so central in Mandarin. Evidence of this is that it is not possible to create a translation equivalent of example (28) without making this relationship linguistically explicit in the English version; that there are multiple possible equivalents is evidence that the relationship is underspecified in the Mandarin original. An issue that will be taken up later in this treatment is whether this ‘underspecification’ in many utterances of Mandarin necessitates reconsidering what may be the essential nature of those utterances in which the SUBJECT-PREDICATE relationships appear to be more fully specified. It becomes necessary to ask if such Mandarin utterances are in fact organized similarly to the utterances of a subject-prominent language like English.

In many analyses of Mandarin, the topic statement is assumed most typically to convey an element of given information. The relevance of the distinction between given versus new information status for analyses of topicalization, however, is currently being debated (for example, Huang, 1994) and more will be said about this below. In the ‘standard analysis,’ following the statement of TOPIC in a Mandarin sentence, the remaining portion of the SV(O) sentence, the COMMENT, is then assumed to add new information to the ongoing discourse. The syntactic, sentential TOPIC, according to this analysis, may or may not correspond to the syntactic subject of the SV(O) construction that links to it or ‘underlies’ it.

As in Levelt (1989:99), it will be important in this description of Study Two to distinguish two types, or levels, of TOPIC in both Mandarin and English. The linguistic analyses mentioned above are concerned with TOPIC as a syntactic constituent, most typically in a sentential kind of construction. A goal of this kind of linguistic analysis is to distinguish TOPIC from other sentential units through its syntactic behavior, as will be outlined below. On another level however, there is something like Reinhart's (1982) notion of TOPIC as 'about-ness'; that is, what a sentence, or any unit of discourse, is about. In this latter sense of TOPIC, Mandarin, English and all other languages, are alike in that speakers' utterances tend to be *about* something. They all have TOPICS in this sense of the term.

Study Two presents another test of the potential of using gesture analysis as a tool to enlarge the window on on-line conceptual representation during speech production. This study concentrates on the features of Mandarin that have been the target of the above-mentioned linguistic analyses; those that support syntactic topic prominence, and that have been analyzed as operating at the level of sentential syntax. The ordering of reference and action in the sentential structures of spoken event representations can differ significantly in Mandarin and English. This is equivalent to saying that the distribution of the components of meaning across sequentially-ordered utterances and what causes them to cohere in the two languages differs. Since this is the case, we seek to know whether the non-spoken representations that are active during language production differ between the two languages as well; whether thinking-for-speaking differs. The assumption guiding the present research is that analysis of the gestures that accompany these utterances, considered together with the units of speech that they accompany, permits precise inferences about thinking-for-speaking. If gesture patterning differs in

ways that relate to the different segmentation and organization of information as expressed in the typical sequences of speech code in the two languages, we may infer different thinking-for-speaking strategies in the production of utterances. Alternatively, if gesture patterning in two speaker groups does not differ, in spite of differences in the way surface utterances are ordered, then we would say that there is no evidence for differences in thinking-for-speaking. Should this be the case, then a different view of the relationship between thinking and speaking would be supported; one along the lines that sentential propositions have a fundamental existence separate from their means of expression.

The target topicalizing structures of Mandarin sampled for this study were of the sort that typically single out one constituent of a spoken event representation, a nominal or a verbal constituent. This constituent is isolated in utterance-initial position, or made part of an utterance-initial phrase. Such linguistic marking of utterance constituents as is found in TOPIC-COMMENT and ‘object-fronting’ utterance structures (to be described below) may focus the attention of Mandarin speakers on the components of events about which they are communicating in a way that is different from the focus promoted by SUBJECT-PREDICATE oriented English. For example, a topicalized PATIENT or AGENT of an action could be more salient for a Mandarin speaker and have a realization in conception that is more separate from ACTION itself than may be the case for an English speaker who produces utterances where the AGENT-as-SUBJECT is syntactically linked to the PREDICATE of a sentential construct, and the PATIENT of an action is likewise linked with the verb expressive of that ACTION in a cohesive PREDICATE.

Points from Current Analyses of Mandarin Syntax

English has a basic SVO utterance structure (Comrie, 1981b) and Mandarin is usually also categorized as a SVO language as well. Li and Thompson (1981:19), however, have suggested that Mandarin may better be viewed as unclassifiable relative to this linguistic typological distinction. These authors claim that, “the notion of subject is not a structurally well-defined one in the grammar of Mandarin,” and further,

“... the order in which basic words and phrases occur is governed to a large extent by considerations of meaning rather than of grammatical functions. This means that sentences with verbs at the beginning, in the middle, and at the end can be found in Mandarin. Languages that are relatively easy to characterize in Greenberg’s [that is, VSO, SVO, or SOV] terms are always those in which word order is determined primarily on strictly grammatical grounds ...”

As an illustration of Mandarin’s word order flexibility, Comrie (1994) lists these three grammatical utterances:

(29) a. wo mai shu le
 我 買 書 了
 I buy book PRF
I buy/bought the book

b. shu wo mai le
 書 我 買 了
 book I buy PRF
I buy/bought the book

c. wo shu mai le
 我 書 買 了
 I book buy PRF
I buy/bought the book

The SUBJECT-PREDICATE relations and general word order patterns on which the SVO typological categorization is based in other languages may be tenuous factors in the

formation of utterances in Mandarin and as such largely irrelevant to syntactic analyses of the language.

There seems to be a preference on the part of linguists in the American and European traditions to put some construal of the nature of SUBJECT-PREDICATE relations at the center of their analyses of every language (Tsao, 1990). Historically, this has meant that linguistic analyses of Mandarin stumbled over the issue of what to do with the quantity of utterances in fluent, grammatical speech whose cohesion seemed not to be explainable on the basis of sentential syntax of the familiar type involving the SUBJECT-PREDICATE relationship. This situation gave rise to analyses based on such notions as ‘major’ and ‘minor’ sentences in Mandarin (Chao, 1968). Major sentences were those seen as having the ‘standard’ components of SV(O) sentences that form the basis of analysis in Indo-European linguistic research practice. Minor sentences were everything else. Unfortunately for such analyses, the minor sentences in speech samples could outnumber the major sentences, and there was no way to account for the internal workings of these non-SV(O) utterances, nor for their relationships with neighboring utterances (Erbaugh, 1982). Although it has been held that Mandarin lacks a grammar altogether, as Erbaugh (1991) points out, this is clearly not the case since it is possible to speak, as she puts it, “wildly ungrammatical” Mandarin.

An alternative approach to accounting for the structure of Mandarin utterances has been to claim that the surface patterning of units is somehow largely ‘semantically’ driven, or that the ordering of actions in Mandarin sentences mirrors their ordering in the real world (Tai, 1985). In line with this notion of some semantic framework having precedence over the sentential-syntactic, are the results of sentence comprehension studies such as Miao’s (1981; replicated by Tzeng & Hung, 1984). Miao found that

native Chinese speakers relied more heavily on semantic cues than on syntactic ones (defined as word order) in sentence interpretation. On the basis of such evidence she inferred that word order, thus, sentential syntax, is not a particularly important source of information in Chinese sentence processing. Other researchers as well (Liu, Bates and Li, 1992) have claimed that Chinese subjects make more use of semantic cues than word order cues in interpreting simple sentences.

It is a fact that Mandarin has a ‘telegraphic’ sound to native English speaker ears, and there is a strong reflex in the English-speaking listener to infer ellided elements, or insert grammatical functors, when interpreting Mandarin utterances. Within the generative-transformational school of linguistics, Huang (1984) develops a grammatical analysis that posits ‘zero’ elements that underlyingly structure utterances in Mandarin. Tsao (1990) finds that this generative-transformational approach fails to yield a coherent explanation for certain key topicalized structures in the language.

Comrie (1994) notes:

“It is therefore not surprising that we find extreme discrepancies in the attitudes of different linguists towards the analysis of Chinese, ranging from those on the one hand who argue that the various syntactic distinctions that are overt in English or Latin are present in Chinese, only in covert form, to those who argue that the syntax of Chinese must be organized on completely different principles—or even that Chinese has no syntax, its apparent syntactic properties being in fact pragmatic.”

Whether Mandarin turns out to be the fundamentally SVO language that many linguists believe it to be, or this distinction in fact does not properly apply, as suggested by Li and Thompson, the language does permit the formation of utterances that, in terms of the overt ordering of surface constituents, pattern similarly to those common in a SVO language like English. Likewise, much of what is found in spontaneous Mandarin narrations does not map neatly onto such a SV(O) pattern, but patterns instead like the

utterances that have prompted linguists to propose alternative syntactic analyses for the language. These facts about the utterances of the language make possible an examination of the effects of different surface-level orderings of constituents on the organization of conceptual representations during speaking.

The Nature of Mandarin Topicalized Utterances

Mandarin possesses several commonly-used utterance structures in which the ordering of reference and action differs from a SV(O) pattern. These include TOPIC-COMMENT structures and so-called ‘object-fronting’ BA (把) constructions. The TOPIC-COMMENT pattern is of course an acceptable expressive option in certain registers of English as well; for example, sentences of the form, ‘As for *p*, *q*,’ or, ‘Now *x*, *y*,’ where *p* and *x* are TOPICS; *q* and *y*, COMMENTS to these TOPICS. However, such constructions do not occur with the same frequency in English as they do in Mandarin. An indication of this is that the American English narration data used for Study Two contained only one instance of linguistically-marked topicalization. The speaker said, “[TOPIC] now this Granny, [COMMENT] she { ... }.” In the narration of the remaining English speaker, who is British, there were two or three utterances that could be interpreted as topicalized. This overall scarcity is in contrast to the data from the Mandarin speakers, in which non-SV(O) utterance structures predominated. It was in fact difficult to extract samples of the ‘standard’ SV(O) utterances from the Mandarin narrations sufficient to balance the data sets used in Study Two. More will be said about this below.

In their reference grammar, Li and Thompson (1981:85) state that the TOPIC of an utterance, “ ... sets a spatial, temporal, or individual framework within which the main

predication holds. In addition, the TOPIC always refers either to something that the hearer already knows about — that is, it is definite — or to a class of entities — that is, it is generic.” Li and Thompson note, as do other treatments, that any part of an utterance may be topicalized; nominals as well as verb phrases. The rest of the utterance then fills out the remaining semantic specification.

Although topicalization in Mandarin often causes the surface order of constituents in complete utterances to vary from a SV(O) pattern, this is not always the case. There are utterances where a sentence-initial TOPIC and syntactic SUBJECT are one and the same constituent. As will be seen below, analyses such as Li and Thompson’s assume that, in the absence of an explicitly marked TOPIC, set apart from the rest of the utterance, the element filling the ‘S’ slot of a SV(O) utterance is to be considered simultaneously both syntactic TOPIC and SUBJECT.

As it is typically characterized in such treatments, TOPIC must always link in prescribed ways to the preceding discourse. Thus topic prominence would seem to be inherently a phenomenon of discourse level, as opposed to sentential level, structure. The Li and Thompson quotation above, for example, shows how some treatments endorse the notion that the information encoded by the TOPIC must be given or presupposed, that it must be something the speaker and listener already know about. This reveals the scope that TOPICS may be supposed to have in such analyses. Nevertheless, Li and Thompson’s and others’ analyses of topicalization are framed almost entirely in terms of suppositions about sentence-internal syntax. So for instance, Li and Thompson state, “... a topic always occurs in *sentence-initial* position ... ” and, “a topic can be separated from the rest of the *sentence* (called ‘comment’) by a pause or by one of the ‘pause particles’ — ‘a’ (or

its phonetic variant ‘ya’), ‘me,’ ‘ne,’ or ‘ba’¹⁶ — although the use of the pause or the pause particle is optional” (1981:86; italics mine).

In the Results section, several examples of another utterance pattern that is often discussed in relation to topicalization in Mandarin will be given as well. The latter are extended sequences; examples of the so-called TOPIC ‘chain.’ TOPIC chains, discussed briefly by Li and Thompson (1981:659) and in more detail by Huang (1994), Shi (1989), and Tsao (1990), consist of a series of phrases, often clause-like, where each phrase is a COMMENT-to-TOPIC pivot point between what precedes it in the discourse and what follows it. In relation to what precedes it, the phrase stands as a COMMENT; in relation to what follows, it stands as TOPIC. Shi (1989) gives the following example of such a construction:

- | | |
|---|--|
| <p>(30) wo kan-shang le zhei-ge gu-niang /
 我看上了這個姑娘 /
 I fall.in.love PRF this-CL girl /
 <i>I fell in love with the girl /</i></p> | <p> ta ye kan-shang le
 他也看上了
 he also fall.in.love PRF
 <i>he also fell in love (with her)</i></p> |
| <p> zui-hou bei ta qiang-zou le
 最後被他搶走了
 finally PAS he rob-away PRF
 <i>in the end (she) was taken away by him</i></p> | |

TOPIC chains are the most frequent pattern seen in the data used for this study. Their length and propositional scope, however, made them unsuitable within the methodological framework of Study Two, given the need to sample utterances from both of the languages that were comparable in propositional content. Such utterances are not

¹⁶Note that the ‘ba’ pause particle is a different morpheme and written character from the object-fronting BA grammatical particle, and that their grammatical functions are unrelated.

‘sentential’ in the usual sense. Grammatical analyses of Mandarin, though, do typically assume that chains and TOPIC-COMMENT sentences are both driven by the same topicalization dynamic of the language. The more complex chained sequences are given in the Results section to illustrate how the linguistic typological feature of topic prominence is a larger factor in the structuring of Chinese narrative than can be captured by the methodology employed in this study. Study Two focuses on the restricted class of shorter, sentence-like utterances and their associated gestures. The longer sequences of chained TOPICS, or TOPICS followed by chains of COMMENTS, therefore had to be excluded from the analyses reported here, but will be taken up in the Discussion.

The standard analysis of topicalization in Mandarin as outlined above was adopted for the methodological framework of Study Two. This analysis attempts to deal systematically with those classes of utterances in Mandarin that have proved problematic for traditional analyses in the style of Indo-European linguistics. As such it is sensitive to utterances that pattern differently from those of English in important ways. The strategy of using the standard analysis has the practical advantage as well of being the most obvious way to align the sets of Mandarin and English narration data in terms of expressive content; that is, on the basis of comparable ‘sentence’-propositional content.

In the Results and Discussion for Study Two, more will be said about the theoretical issues that are raised by examination of topicalized structures actually found in Mandarin narrations, and by analyses of the ways that topicalization processes function at the discourse level. The extra-sentential framework from which the utterances examined in Study Two were excised is equally important to understanding the utterance structuring dynamics of Mandarin and English, and to applying the findings of this study toward further development of GP theory. The preceding comments and the examples of Mandarin topicalized structures presented in the following section are meant simply to

frame the more restricted analysis that is presented here, and make reasonably explicit the sorts of utterance structures with which this comparison of Mandarin and English narrations is concerned.

TOPIC-COMMENT Constructions

The listing in this section gives several examples of topicalized utterances of the type found in grammars of Mandarin. The examples that follow are taken from three different sources: Li and Thompson's (1981) *Functional Reference Grammar*, a Mandarin Chinese instructional textbook called *Practical Chinese Dialogue II* (Li, 1988), and from the Mandarin narration data used for this study. These examples will help to give a sense of what the most-researched topicalized utterance structures are like, and what it means to create topic statements out of different utterance constituents; for instance, noun phrases versus verb phrases. Examples such as the ones in this listing were used as templates for sampling utterances from the Mandarin narrations. Some of the examples serve as well to illustrate the difficulties involved in distinguishing topicalized from non-topicalized utterances, and the difficulty of determining relationships among syntactic TOPIC, SUBJECT, and the rest of the material making up an utterance. In actual practice, not the least of these difficulties lies in determining, in free-flowing narration, exactly where 'utterance-initial' position is. This particular criterion of the standard theory was at times difficult to apply in putting together the data set used in Study Two. Some of the implications of this difficulty as it relates to evaluating the Study Two data set, as well as to current theories of topic versus subject prominence in languages are taken up in the Discussion.

According to Li and Thompson, example (31) is a sentence with both syntactic TOPIC and SUBJECT constituents:

- (31) *nei-zhi gou* *wo* *yi-jing kan-guo le*
 那隻狗 我 已經看過了
 that-CL dog I already see-EXP PRF
 TOPIC & SUBJECT
 OBJECT COMMENT — — —>
 that dog, I have already seen

Li and Thompson explain, “... we can identify the topic as *nei-zhi gou* ‘that dog’, since it tells what the sentence is about, it is definite, it occurs in sentence-initial position, and it could be followed by a pause particle. It is not the subject, however, since it has no ‘doing’ or ‘being’ relationship with the verb *kan* ‘see’.” Instead, it is the subject *wo* ‘I’ that has a ‘doing or being’ relationship with the verb in this sentence. Therefore, the syntactic relationships with which the standard analysis is concerned are identified with reference to the main verb of the utterance.

According to Li and Thompson, a single utterance-initial element of a sentence may be simultaneously both its sentential SUBJECT and its TOPIC, as in example (32). It is an utterance that looks very much like an English sentence with “wo”/“I” as its syntactic SUBJECT.

- (32) *wo* *xi-huan chi ping-guo*
 我 喜歡吃蘋果
 I like eat apple
 TOPIC & COMMENT
 SUBJECT
 I like (to) eat apples

Li and Thompson’s focus on the syntactic TOPIC status of the constituent “wo”/“I” reflects these authors’ theoretical view that TOPIC is a primary and obligatory syntactic constituent in Mandarin, while SUBJECT and the SUBJECT-PREDICATE relationship are secondary. Signalling the latter grammatically is optional, according to these authors. To

native speakers of a subject-prominent SVO language such as English, such a phrase constituent appears to be nothing more than a syntactic SUBJECT; however, according to standard topicalization theory, in Mandarin it must be considered to be the syntactic TOPIC as well by virtue of its sentence-initial position. To get a sense of what this analysis may mean from a Mandarin speaker's perspective, realize that a pause or one of the pause particles may be inserted after any utterance-initial spoken constituent, such as the "I" in example (32). The equivalent sense in an English translation of the sentence would perhaps be a very marked form such as, "As for me, I like to eat apples." Further, and perhaps more importantly, following on a pause particle in Mandarin, multiple different constituent organizations are allowed. That (32) has no pause, and that the COMMENT that follows on the utterance-initial "I" is overtly structurally similar to an English (complementized) verb-object PREDICATE, on the Li and Thompson view, are not a reflection of such an utterance's particular structural cohesion. As will be seen in several of the examples below, a Mandarin utterance may lack an overt syntactic SUBJECT, and still both cohere and be a grammatical sentence.

Examples (33) *a* through *c* are further examples of topicalized utterances with nominal TOPICS, taken from the Mandarin grammar textbook.

- (33) a. xing-li dou na-dao le
 行李 都拿到了
 suitcase all pick up-CMPL PRF
 TOPIC & COMMENT
 OBJECT
 (As for the) suitcases, (they have) all (been) picked up

- b. nei-ge xue-sheng zhong-guo hua shuo-de bu-cuo
 那個學生 中國話說得不錯
 that-CL student china language speak-MOD not bad
 TOPIC COMMENT
 (As for) that student, (her) Chinese (is) spoken pretty well

- c. che yi-jing ting zai nar le
 車 已經停在那兒了
 car already stop at there PRF
 TOPIC & COMMENT
 SUBJECT
 (As for the) car, (it is) already stopped/parked there

The next example is a topicalized structure of a sort that has received much attention from linguists (Li and Thompson, 1976; Huang, 1984; McCawley, 1989; Tsao, 1990; others). Many of these treatments follow standard terminology in calling it a ‘double-subject’ TOPIC-COMMENT structure.

- (34) da-xiang bi-zi hen chang
 大象 鼻子很長
 elephant nose very long
 TOPIC COMMENT
 (as for) elephants, (their) noses (are) very long

As McCawley (1989) points out, this standard usage is a misnomer, for if syntactic SUBJECTs must stand in a ‘doing or being’ relationship with their predicates, then clearly “elephant” cannot be considered to be nominative case; rather, only “nose” fulfills that criterion.

The following are examples of the variety of utterance constituents that may serve as TOPICS, taken from the textbook or the narrations, as indicated. From the text, here is an example of a ‘stative verb’ as TOPIC:

- (35) pian-yi bu yi-ding hao
 便宜 不一定好
 inexpensive not certain good
 TOPIC COMMENT
(just because a thing is) inexpensive, (does) not (mean it is) necessarily good

From the narration data: verb phrase as TOPIC:

- (36) zou-yi-zou ne hou-mian lai le yi-liang dian-che
 走一走 呢 後面來了一輛電車
 walk-one-walk TOP behind come PRF one-CL train
 TOPIC COMMENT
(as he is) walking, a train comes from behind

From the text: verb-object as TOPIC:

- (37) ban jia fei-chang ma-fan
 搬家 非常麻煩
 move house extremely trouble
 TOPIC COMMENT
(as for) moving, (it is) extremely inconvenient

From the text: 'co-verb'-noun-VP as TOPIC:

- (38) gen ta shuo-hua dei shuo man yi-dianr
 跟他說話 得說慢一點兒
 with him speak must speak slow one-bit
 TOPIC COMMENT
(in) speak(ing) with him (one) must speak (a) bit (more) slowly

From the narration data: 'clause' as TOPIC:

- (39) ta xia-de zai li-mian luan pao
 他嚇得 在裡面亂跑
 he frighten-MOD at inside disorder run
 TOPIC COMMENT
he (is) frightened (and) runs around crazily inside (the cage)

Taken together, these utterances illustrate the fact that syntactic SUBJECT is an optional element in Mandarin utterances. The topicalized structures listed above are the sort that were sought in the Mandarin narrations, and the selected samples all match one or another of these templates.

The BA Construction

The other utterance structure in Mandarin that was sampled in Study Two is the BA-construction. The BA (把) grammatical particle is often referred to as an ‘object-fronting’ particle. Its function within an utterance is typically analyzed as marking an argument positioned in front of the verb as being that verb’s direct object. This is analyzed by some as a topicalizing movement transformation (Tsao, 1987).

Historically, BA was a content word in Mandarin, a verb meaning ‘to grasp.’ Over the past several centuries however, BA has undergone grammaticization to assume its current particle function (Chappell, 1991). The prototypic account of BA is that it moves the direct object of a transitive verb forward from its ‘standard’ location following the verb, on the view that Mandarin is fundamentally a SVO language. Thus the arrangement of constituents in a sentence like example (40) *a*, below, changes when the BA particle is introduced to become a sentence like the one shown in example (40) *b*.

- (40) a. lao tai-tai fang niao long-zi zai chuang-tai-shang
 老太太 放 鳥籠子 在窗台上
 old lady put bird cage on window-sill-surface
 the old lady puts the bird cage on the window sill

- b. lao tai-tai BA niao long-zi fang zai chuang-tai-shang
 老太太 把 鳥籠子 放 在窗台上
 the old lady OBJ bird cage put on window-sill-surface
 the old lady puts the bird cage on the window sill

or perhaps:

- c. *the old lady, as for (that) bird cage, (she) puts (it) on the window sill*

The latter, alternative, English gloss of the BA-construction attempts to capture what may be a more appropriate reading if the BA-construction is properly to be considered a topicalized utterance structure. On such an account, BA is topicalizing because it moves a noun phrase in the direction of utterance-initial position, the place for TOPICS. Such accounts do not mention that this movement is only *toward*, rather than always *to* initial position, a peculiar omission given the insistence of many treatments of topicalization in Mandarin that utterance-initial position is key. However, the narration data used for this study show that the BA-marked element is almost always an utterance constituent that has already been referred to in the discourse context prior to the point where it appears with the BA particle. The moved NP is thus always marked for definite reference, so is a given element of discourse. This fact fits with many treatments of topicalization.

It is interesting to reflect on the motivation of the term ‘object-fronting particle,’ as applied to BA. Implicit in this term is the assumption that there is an underlying sequential structure to utterances like the BA-construction above that is SVO instead of SOV. Yet native Mandarin speakers,¹⁷ when asked which of the two sentences, (40) *a* or *b*, sounds more natural, select the object-fronted one. The SVO version, while not

¹⁷Hui-fang Hong (personal communication).

grammatically incorrect, still sounds somewhat anomalous or marked to Mandarin ears. Thus, behind a term like ‘object-fronting particle’ lies a theoretical assumption about what we should take to be the fundamental word order pattern of the language. The relative frequency of SOV as compared to SVO constructions in the spontaneous narration data used here, as well as native speaker preferences, leave such an assumption open to question.

The prototypic account of BA specifies as well that the kinds of object NPs that may be ‘fronted’ in this way are those that actually receive the action of the verb, and that there is a preference for this to be manipulatory action, or an action with observable effects upon the object; that is, the object should suffer the effect of the verb. This was the case in example (40). Such features of the particle's use are attributed to the particulars of its previous semantics as a content word, ‘grasp.’ Expanded and non-prototypic accounts such as that of Chappell (1991), however, discuss other versions of the BA-construction, including uses involving intransitive verbs. The following grammatically acceptable sentence for instance,

- (41) wo BA wu-ge ping-guo chi le liang-ge
 我 把五個蘋果 吃了兩個
 I OBJ five-CL apples eat PRF two-CL
 of the five apples I ate two

is one example that shows how the requirement that the BA-marked element in an utterance be the transitive object of a manipulatory verb does not hold up.¹⁸ We see that the apples actually eaten in example (41) are only a sub-set of the set identified in the BA-marked NP. In this usage, a function of the BA-construction appears to be

¹⁸I thank Lin Fu-wen for bringing examples like this to my attention.

establishing a domain of reference (a TOPIC), or a FOCUS.

Thus we see that BA has wider application than prototypic treatments of it typically highlight. As the examples above show, the BA-construction, whether specifically topicalizing or not, is another utterance structure in Mandarin that patterns differently from utterances bearing similar propositional content in English. BA-constructions are extremely frequent in the Mandarin narrations used for Study Two, and are comparable to English SV(O) sentences in terms of propositional content. For these reasons, such utterances comprise a comparison sample that met the requirements of the study.

Summary

As is true of linguistic verb aspect, the focus of Study One, the notions of discourse TOPIC, syntactic TOPIC, and the strategies different languages use to manage them, are contentious areas of linguistic research. Most current analyses locate explanations for the differences in the patterning of reference in topicalized as opposed to SV(O) utterance structures on the level of sentential syntactic analysis, but there are many unresolved issues. In his summation of the research by various schools of linguistics on the nature of sentences in English and Mandarin, Tsao (1990) comments that, while for English there is a degree of unanimity among analyses as to the two critical component categories, subject and finite verb, that are central to the sentential unit in English, he says, “... Chinese linguists, on the other hand, do not seem to have had the luck.” He cites Chao’s (1968) failure to provide a strictly ‘grammatical’ analysis of Mandarin utterance structure, as an example of how linguists in that tradition have been unable to provide a coherent account at the level of sentential syntax of the roles of SUBJECT and

TOPIC in Mandarin utterances. He cites the generative-transformational grammarians, for example Huang (1984), as failing to account for the so-called ‘double-nominative’ sentences. Tsao writes,

“In the course of discussion, we have seen that both topic and subject, rather than subject alone as in the case of English, seem to play a part in the grammatical organization of Chinese. ... these two important categories, due to the fact that such a distinction does not play an important role in the description of English and other Indo-European languages, are often mixed up [in linguistic analyses of Chinese] ...”

That Mandarin and English present different patterns of utterance structure is assumed for the purposes of this study. Examination of the narration data used for Study Two bears this out. In the Discussion, the potential of the results of Study Two to point a direction for further linguistic research into utterance structure in topic-prominent Mandarin versus subject-prominent English is taken up, as is their potential to contribute to an understanding of the underlying differences in utterance production in the two languages.

Methods for Study Two

Participants

The response data of eleven of the Mandarin- and seven of the English-speaking participants were sampled for Study Two. These were the participants who contributed cartoon narrations. Not all participants produced fully fluent, gesture-accompanied utterances of the target types in equal number; therefore, the samples of target utterance types could not be completely balanced across participants. With one exception, however, all participants contributed at least two samples to each target utterance category. The exception was a Mandarin speaker who, in an eight-minute narration, produced no fluent, gesture-accompanied TOPIC-COMMENT utterances of the sentence-

like variety required for inclusion in this study, although she did produce fluent utterances of the other three types, accompanied by gestures. The table in Appendix D summarizes the contributions of each participant to the data set for Study Two.

Elicitation

There are two reasons that only the cartoon narration data were sampled for Study Two. First, one of the virtues of the cartoon elicitation is that it gives somewhat tighter control than does the longer and more layered Hitchcock film, over the input to the conceptual representations that speakers form, while at the same time providing sufficient story content for a genuine narrative. The restricted plot and simple characterizations of the cartoon insure that few speakers' narrations diverge too far from a certain set of story events. This makes the cartoon narrations a better source than the movie narrations for utterances with comparable propositional content across speakers. Second, data from the Vignettes Task were excluded because close examination of the responses the vignettes elicited revealed a heightened attention across all of the participants, but especially the Chinese, to the peculiar object features of the moving character or object depicted in each vignette. The Chinese commented far more often on such features and in general on how odd were the characters and actions depicted. Since many Chinese participants appeared to regard the Vignettes Task as one of characterizing the objects themselves in some detail, it was felt that this could skew the analyses presented here. These analyses were intended to pick up on how depiction of elements and events ordinarily is handled in naturalistic, spontaneous expression. For these reasons, only the cartoon narrations were examined for Study Two.

Speech sampling

In the same way as described for Study One, the sampling process began with an examination of the speech transcripts themselves. The gestures that accompanied speech were considered only in later steps.

The speech transcript of each participant's cartoon narration was searched for speech sequences matching the target utterance structures. From the Mandarin narrations, four utterance types were sampled: the two topicalized varieties described in the preceding section (TOPIC-COMMENT and BA-constructions), and two varieties that bore no overt linguistic indicators (disregarding the sentence-initial default TOPIC position) of topicalization (SV and SVO). The latter, in other words, were utterances that in terms of the spoken order of their constituents, patterned similarly to SV and SVO sentences of subject-prominent English.

The English cartoon narrations were scanned for utterances having SV or SVO sentence structures, controlling for equivalent or closely equivalent informational content to those selected from the Mandarin narrations.

The utterance types sampled from the two speaker groups are as follows:

(i) TOPIC-COMMENT: From the Mandarin narrations, the TOPIC-COMMENT utterances sampled were restricted to those of the 'textbook,' sentence-like variety, where the TOPIC and COMMENT components of the utterance are adjacent to one another and the TOPIC is either clearly set off from the rest of the utterance through linguistic marking and thus distinguishable as a TOPIC element, or is overtly the same entity as the syntactic SUBJECT but identified as syntactic TOPIC by the presence of a following pause particle such as 'ne'/呢, 'ba'/吧, 'ma'/嘛, or the like.

(ii) Object-fronted: These are Mandarin utterances that include the object-marking particle BA/把. Each has a sentential SUBJECT, a verb, and a DIRECT OBJECT (PATIENT) of that verb. In each case the DIRECT OBJECT appears in front of the verb, marked by the BA particle.

(iii) SVO: These are utterances in Mandarin and English with sentential SUBJECTs and verbs followed by DIRECT OBJECTs.

(iv) SV: These are utterances in Mandarin and English with sentential SUBJECTs and verbs, but no DIRECT OBJECTs of the verb.

Additional utterance elements, such as locative phrases, constituents expressive of the GROUND components of the events described, along with various descriptive elements, were unavoidably part of many of the utterances sampled as well, in both languages. In Mandarin, such elements usually occur before the verb and so are lodged within the internal utterance structure.

As was the case in Study One, dysfluent speech was not sampled. In Study Two again the particular reason for not sampling broken, hesitant, or repetitious speech was the fact that such speech often occurs when a speaker is attempting to recall a word or phrase, but experiencing a momentary failure to do so. In addition to the repetitious movements or 'Butterworth beats' discussed in the context of Study One, dysfluencies of this type have also been observed to be accompanied by iconic gesturing of a particular kind. Speakers appear sometimes in these situations appear to use gesture to support the lexical search process (Krauss, et al., 1991), and/or to gesturally convey a particular meaning to the listener, in the absence of the target spoken lexical item. When such gestural substitutes for speech co-occur with pauses in narration that are associated with dysfluency, we are no longer dealing with strictly 'speech-accompanying' gesture. Not

enough is known about the possible differences between this kind of highly communicatively-loaded gesture and the less (explicitly) communicatively-loaded, unwitting gestures that accompany fluent speech. The motivating factors behind gesture form and execution may be different in gestures that occur in the absence of speech versus those that occur with speech (the speaker's degree of conscious intention to communicate may differ, for example); therefore, the analyses here were restricted to the latter variety.¹⁹ Thus, only fluent speech was sampled from the narrations.

Gesture sampling

All representational gestures that co-occur with the sampled utterances were included in the analysis. In contrast to Study One, where only gestures related to the action expressed by the verb of an utterance were included for analysis, in Study Two, gestures related to the physical entities in play were included as well. In Study Two, the only gestures that were excluded were those analyzed as non-representational at the narrative level of analysis; beats, for example. These are relatively few in number.

Sampling totals

The data set for Study Two comprised a total of 160 utterance-gesture pairs sampled from the Mandarin and English speakers according to the above sampling criteria. The 80 Mandarin utterances consisted of two subsets as follows:

(i) 40 topicalized utterance structures (20 TOPIC-COMMENT and 20 object-fronted utterances) and their accompanying representational gestures.

(ii) 40 SV(O) utterance structures (20 SV and 20 SVO utterances) and

¹⁹The theoretically very important issue of when and whether gestures are communicatively loaded, are produced with intention to communicate, or do function to communicate propositional content, is beyond the scope of this treatment.

accompanying representational gestures.

The 80 English utterances consisted of two subsets as follows:

- (i) 40 SV utterance structures, and accompanying representational gestures.
- (ii) 40 SVO utterance structures, and accompanying representational gestures.

As was the case in Study One, utterance-gesture pairs were selected for inclusion in the data set by starting at the beginning of each speaker's narration and moving sequentially through it, extracting those pairs that met the sampling criteria of the study. An attempt was made to balance the contributions of all participants numerically, although this was not always possible, since not all subjects produced fluent, gesture-accompanied target speech in equal amounts. Again, no more than five utterance-gesture pairs in one utterance category were extracted from the narration of any one speaker. The number of utterances contributed by each speaker is listed in Appendix D. The complete listing of utterance-gesture pairs from each language is presented in Appendix C.

Gesture coding

The results of the exploratory pilot study had suggested that Mandarin speakers gesture at a higher rate than English speakers and that they tend more than the English speakers to produce gestures expressive of the physical objects or entities that figure in their re-tellings. These preliminary findings pointed to possible differences in thinking-for-speaking between the two speaker groups, and it is hypothesized that these are related to organizational differences among the target utterance structures in Mandarin and English. In a procedure similar to that described for Study One, the gestures accompanying the utterances sampled from the narrations were examined for evidence of the encoding of event components such as ACTION, AGENT/FIGURE, PATIENT, GROUND and so on. No representational gestures were excluded in Study Two. This means that,

unlike in Study One, many of the utterance-gesture ‘pairs’ contained more than one gesture. All of the representational gestures accompanying each of the Mandarin and English utterances in the samples were classified as belonging to one of three categories, according to the manner in which components of the cartoon event are encoded gesturally. The gestures were categorized as follows:

(i) ‘Isolating’ versus ‘Incorporating’: As was discussed in the previous chapter in regard to Study One, the gestural realization of an event may be pared down to a single event component, PATH for example, or may conflate multiple event components into a more elaborate production, one that incorporates FIGURE, PATH, and MANNER, for instance. Categories *a* through *c*, below, captured this distinction.

a. Element Only: These are Isolating gestures that encode only features of some item in the narrative re-telling — some physical object or a cartoon character; a FIGURE, for instance. Examples of such Element-encoding include gestural expression of the shape of a bird cage, the extent of a wire, the small size of the bird, or the position and/or some feature(s) of the entity filling the AGENT or the PATIENT role in a proposition. As well, gesturally expressed features of the GROUND against which an event was enacted were categorized as Element-encoding.

If a gesture expressed any feature of an Action involving the Element, it was not coded as an Element Only gesture.

b. Action Only: This is the other category of Isolating gestures. Action Only gestures encode only features of some action in an event of motion, or of some event of stasis. In the case of motion events, this is typically the PATH or MANNER of motion.

If the gesture encoded any features of an Element conflated with the Action, for instance a PATH-related movement with an Element-marked handshape, the gesture was

not coded as Action Only.

c. Incorporating: Such gestures simultaneously encode features of both Elements and Actions in the narrative context. An example of an Incorporating gesture that encodes the motion event components of FIGURE, PATH, and MANNER is one in which a speaker moves her hand horizontally across her gesture space, with wiggling fingers oriented downward to represent a cartoon character tightrope-walking on a wire. A subset of the gestures in the Incorporating category was further distinguished as ‘enactments.’ This judgment was based on the viewpoint that a speaker manifests in gesture: character or observer. The category of viewpoint is explained below in (iii). Enactments are gestures in which the speaker represents a cartoon action by pantomiming it. Such gesture performances are considered in the framework here to be maximally Incorporating, in that they encode multiple features of the components of an event in a highly synthetic fashion.

(ii) Event Component: Every non-pantomimic gestural expression of an event component — ACTION, AGENT/FIGURE, PATIENT, and so on — was classified according to what role its semantic content occupies in relation to the verb of the accompanying utterance. An explicit gestural reference to the action encoded by a verb of motion, for instance, was coded as an instance of gestural depiction of ACTION. Element-marking was assigned to semantic content categories according to which argument of the verb is indexed or depicted. This was essentially an assignment of case, functionally described here as, for instance, the AGENT or PATIENT. Gestures were categorized according to which of the following event components they encode:

a. AGENT / FIGURE — These two categories were collapsed into one for this analysis. The utterance category itself directs which way this content should be

construed in any particular instance.

b. ACTION

c. PATIENT of the action encoded by the verb

d. ‘Other’ — GROUND, LOCATION, INSTRUMENT, RECIPIENT, and other references — these being additional event components that appear in the speech and gestures of both speaker groups.

The event component category targets the issue of whether the different categories of utterances examined here differentially associate with explicit gestural references to actions or to particular arguments of the verb in spoken utterances. A question that motivated this study is whether such structures as Mandarin non-SV(O) utterances, in which there is special linguistic marking of particular sentence constituents, are associated with particular salience of their referents in conception. Particular salience would correspond with a tendency of some components to be singled out for specific and/or separate realization in the gestures that accompany speech. The object-fronting BA-construction, for instance, might be expected to associate with gestures that depict features of the fronted object, the PATIENT of the action encoded in the verb. Such features may be expressed singly in Isolating gestures, separated from additional gestures in multi-gesture productions, or may be expressed in Incorporating gestures as Element-marking conflated with Action features. We may expect as well that topicalized NPs of any sort could be singled out in co-expressive gestures having closely related semantic content, perhaps conflated with additional reference to the action expressed by the verb.

Note that, as was true in Study One, the depiction of Element and Action in pantomimic enactments is considered a qualitatively different kind of gestural realization than is specific featural Element- or Action-marking in non-pantomimic gestures. This is

because the pantomimic mode of representation brings all event components ‘along for the ride,’ so to speak. All of the components are present in such a representation by default without any one of them or a subset of them being specifically singled out.

(iii) Viewpoint: In order to sort pantomimic enactments from other gestures, a judgment was made of the speaker-gesturer’s perspective, in gesture, on the re-told cartoon events. Every gesture was coded as manifesting one of three viewpoint distinctions according to the considerations outlined below.

a. Observer: A speaker gesturally represents Actions or Elements in a way that does not involve assuming the perspective of the thing depicted. The gesture described above in the explanation of Incorporating gestures that represented the tightrope-walking cartoon character is a good example of an observer viewpoint gesture.

b. Character: A speaker uses her own hand *as a hand*, for instance, to represent a cartoon character hitting something with a stick. In such a case, the hand is configured so that it looks like it is holding the instrument of the action. Note that the pantomimic enactment gestures just described in the section on Incorporating gestures are always by default character viewpoint gestures.

c. Dual: A gesture that simultaneously has elements of both observer and character viewpoints is coded as displaying a dual viewpoint. For example, there is a scene in the cartoon in which the cat catapults himself up to the window where the bird is sitting, grabs the bird in his hand, and then plummets back down to the ground while still holding the bird. A dual viewpoint gesture that some speakers perform when narrating this scene is one in which the speaker’s hand depicts the cat’s hand holding onto the bird. This *holding* configuration is maintained while the speaker moves her hand from upper gesture space to lower and then across, to show the path taken by the cat and bird as they fall together and move away. The key feature of such a performance that signals a dual viewpoint is that, although pantomiming the act of holding the bird, the way that the

speaker's hand-plus-bird then moves to depict PATH cannot be construed as part of this pantomime; rather, in that movement, the configured hand is simultaneously representing both the hand of the cat as well as the cat's and bird's bodies as they fall down together and run away. Dual viewpoint gestures in the data examined here are often combinations of a pantomimic enactment together with an additional, explicit gestural reference to Element or Action.

Coding summary

The units of analysis that figured in the coding for Study Two are summarized in Table 5.

Table 5. Summary of coding categories in speech and gesture — Study Two.

<u>Speech</u>	<u>Gesture</u>
<u>Utterance structure:</u>	• Isolating versus Incorporating:
• SV	— Element Only
• SVO	— Action Only
• TOPIC-COMMENT	— Element and Action Incorporating
• BA-construction	• Event component and case role marking:
	— AGENT / FIGURE
	— ACTION
	— PATIENT
	— 'Other': GROUND, LOCATION, INSTRUMENT, and so on
	• Viewpoint:
	— observer
	— character
	— dual

Analysis

The speech sample. The outcomes of the speech sampling process were examined first. The narrations in the two languages were compared in terms of the

relative ease or difficulty of extracting ‘clean’ exemplars of the target utterance structures, and also in terms of what utterance structures were found to predominate in each.

Comparison of gesture-to-utterance ratios. Since the Mandarin topicalization processes that are a focus of Study Two were described above as potentially having the effect of singling out individual components of events for separate focus, analysis targeted the possibility that Mandarin speakers, if they do gesturally individuate event components, could show a higher ratio of gestures to utterances than English speakers. The total numbers of representational gestures by language, and within language by target utterance category, were tallied, and the ratios of gestures to utterances compared.

Comparisons of gesture semantic content. To determine whether differences in gesture-speech ratios are a function of a tendency selectively to isolate event components versus conflate them in gestures, the extent to which narrators in each language gesturally single out specific components in Isolating gestures, or alternatively, produce Incorporating gestures that simultaneously encode multiple components was analyzed. The two types of Isolating gestures — Element Only and Action Only — were tallied, as were Incorporating gestures. Within the category of Incorporating gestures, separate tallies were made of pantomimic enactments. These tallies were compared by target utterance structure and language.

To the extent that event components are realized in Isolating or in specifically marked Incorporating gestures, and these gestures occur either singly, or together with other gestures to accompany single sentential utterances, it is potentially interesting to know whether, in the context of individual speech-gesture productions, the gesture production is configured to encode event content in specific ways. For example, in which forms do gesturally-realized event components come packaged together with different target utterance structures? With which arguments of the verbs of utterances are gestures

co-expressive? Is this encoding different in systematic ways when gestures occur singly versus together with other gestures as accompaniments to single spoken utterances?

Thus the gestural encoding of the components of events — AGENT/FIGURE, ACTION, PATIENT, and ‘Other’ — was considered in relation to the following factors: (i) language; (ii) utterance structure: SV, SVO, TOPIC-COMMENT, and BA; (iii) type of gesture: Isolating (and within Isolating gestures, Element versus Action) versus Incorporating (and within Incorporating gestures, non-enactment versus enactment); (iv) single-gesture accompaniment versus multiple-gesture accompaniment of utterances. In the case of the final factor, instances where single utterances are accompanied by single gestures are compared to instances where single utterances are accompanied by two or more gestures in sequence, in order to determine whether the priority given in gestural encoding to particular components of events varies with this factor, and whether this interacts with target utterance structure.

In regard to event component encoding in gesture, the analysis undertaken here related to the four factors just described is limited. It is reasonable to suppose that event component encoding is influenced by each of the factors listed above; influenced in fact by multiples of them in interaction with one another. With this number of factors and levels within factors, and considering the complexities of naturalistic spontaneous language data, a larger data set than the 160 spoken utterances and their 222 accompanying gestures sampled here would be required to sort out the various effects of the four factors singly and in combination, and to exclude the influence of additional factors. The tables that summarize the Study Two data according to the four factors demonstrate how thinly spread the data quickly become when they are considered in relation to such an array of categories. Nevertheless, certain islands of patterning emerge. Across the two languages there appear to be some differences in the gestural

encoding of event components related to target utterance structure and possibly to whether gestures occur singly or multiply as accompaniments to utterances. These are presented as tentative findings and discussed in relationship to utterance structuring dynamics in the two languages.

Additional observations. The speech-gesture phenomena with which the analyses described above are concerned all have to do with the different types of gestures the two speaker groups produce and which parts of propositional or event content the gestures may differentially highlight in each of the target utterance contexts. Details concerning the exact timing of gesture relative to co-expressive speech are not featured in such an analysis. Two additional phenomena were observed in the course of sampling from the narration data and analyzing individual productions. These have to do with the issue of speech-gesture synchronization and with how speech-gesture productions coordinate with one another in sequence. The phenomena appear to be related to the differences in utterance structure that motivated this research and so will be described in the Results following the planned analyses.

The first phenomenon concerns where gestures expressive of particular semantic content time with respect to their co-referential units of speech within individual production envelopes. Speech-gesture synchronization in the Mandarin non-SV(O) utterances extracted for Study Two is described. This is followed by a description of what appears to be a general timing phenomenon in Mandarin. Apparently contrasting tendencies between the two speaker groups were observed in how gestures typically synchronize with particular constituents of the spoken utterances they accompany. These synchronization tendencies are not manifested in every utterance but are prevalent nonetheless. The potential for different patterns of speech-gesture synchrony to provide clues to the utterance-structuring dynamics of Mandarin and English is discussed.

The second phenomenon emerges from consideration of the connections between speech-gesture productions in sequential discourse. TOPICs appear to be established and maintained in gesture as well as in speech. Gesture repetition plays a role in this. Observations of this kind are external to the planned analyses of Study Two, because the design of the latter ignored the connections of each extracted speech-gesture production to its surrounding context. Such additional observations arose in the process of examining connected sequences of discourse in order to extract samples for Study Two's data sets. These findings are not yet systematic. They are described in order to suggest some directions for further research.

Coding reliability

Ten percent of the utterance data used in this study, divided evenly between English and Mandarin, was coded separately by another experienced gesture coder for the details of form and content that are relevant for this study. The other coder is also knowledgeable in Mandarin. The reliability coding categories for Study Two, and per cent agreement between the two coders are as follows:

1) discerning the presence of a gesture stroke	96%
2) identifying a gesture as representational or not	92%
3) identifying gestures as Element, Action, or Incorporating	95%
4) identifying the event component(s) depicted or indexed	88%
5) identifying viewpoint	90%

Results of Study Two

Outcomes of the Speech Sampling Process

There were several noteworthy outcomes of the speech sampling process. First, in spite of the attention and analysis that sentential TOPIC-COMMENT structures in Mandarin have received in the literature on Chinese linguistics, there are relatively few ‘textbook’-quality examples of these structures in the Mandarin narration data examined for Study Two. That is, it was not always a straightforward process to extract samples of fluent speech-gesture productions with appropriately limited propositional content matching any of the templates given previously in the section on TOPIC-COMMENT constructions. At least, not every one of the eleven narrations examined had them in quantity. Further, the incidence of fluent SV(O)-like utterances suitable for sampling in the Mandarin narrations is also quite low.

These outcomes are odd, given the standard analysis of Mandarin as a topic-prominent, SV(O) language. The expectation would be that sentences with TOPIC-COMMENT utterance structure, or those with overt SUBJECTs and PREDICATES should be easy to find but such was not the case. By contrast, in the English narrations fluent, gesture-accompanied SV(O) utterances are abundant and easy to identify.

What the Mandarin narrations have in abundance are TOPIC chains of the sort described in the Introduction to this chapter. Material from these chains could not be sampled for Study Two since typically there seems to be no reasonable way to excise one sentence-like portion of the chain and leave the rest. To give a sense of the kinds of such chained utterances the Mandarin speaking participants in this study produced, below are some examples drawn from the narrations. In the following transcriptions, ‘#’ and ‘/’ indicate breath and unfilled pauses, respectively; ‘ne’ and ‘a’ are TOPIC-marking pause

particles.

- (42) nei-zhi xiao niao # ke-shi nei-zhi xiao niao ne / dang-ran you zhu-ren a /
 那隻小鳥 # 可是那隻小鳥 呢 / 當然有主人 呵 /
 that-CL little bird # but that-CL little bird TOP / of course have owner TOP /
 TOPIC TOPIC COMMENT-TOPIC
 now, the little bird # but as for the little bird, of course he has an owner /

zhu-ren shi yi-ge lao-tai-po / gui-fu ren ne nei-yang-zi #
 主人是一個老太婆 / 貴婦人 呢 那樣子 #
 owner be one-CL old-lady / honorable person TOP that-way #
 COMMENT-TOPIC COMMENT-TOPIC COMMENT
the owner is an old lady, a respectable type, like that

ran-hou nei- gui-fu ren ne jiu ba ta gei gan-chu-lai le
 然後那貴婦人 呢 就把他給趕出來了
 next that- honorable person TOP just OBJ him give hurry-exit-come PRF
 TOPIC COMMENT
as for the lady, then (she) hurries him away

The following is a chained sequence from a different participant. It has fewer overt TOPIC-marking ‘pause particles.’

- (43) hou-lai # xiao niao zhuan -ge tou # kan-dao mao
 後來 # 小鳥轉個頭 # 看到貓
 after # little bird turn -CL head # see-CMPL cat
 COMMENT-TOPIC COMMENT-TOPIC
 afterward, (the) little bird turns (his) head (and) sees (the) cat

HOOH / xia-de / <eehh> zai li-mian luan pao
 HOOH / 嚇得 / <eehh> 在裡面亂跑
 ONOM / fright-MOD / <eehh> in inside disorder run
 COMMENT-TOPIC COMMENT-TOPIC COMMENT
yikes! (he) takes fright <eehh> (and) runs around inside (his cage)

The most common type of chaining encountered in these Mandarin narrations used narrative events in sequence as pivot points in TOPIC chains. The preponderance of this particular kind of temporal sequencing in these narrations is likely an artifact of the cartoon story structure — a linear sequence of events that proceed to a single goal. The following example is an instance of this kind of temporal sequencing TOPIC chain:

- (44) ta jiu pa / pa-shang-lai # pa-shang-lai yi-hou
 他就爬 / 爬上來 # 爬上來以後
 he just climb / climb-up-come # climb-up come after
 COMMENT-TOPIC COMMENT-TOPIC TOPIC
then he [the cat] climbs / climbs up # (and) after (he has) climbed up

nei-ge xiao niao ne / tao-zou le # tao-zou le yi-hou ne
 那個小鳥 呢 / 逃走了 # 逃走了以後 呢
 that-CL small bird TOP / flee-go PRF # flee-go after TOP
 TOPIC COMMENT TOPIC
as for the little bird / (he) flees # and after fleeing

ta di-yi ci mei-you cheng-gong mei-you cheng-gong yi-hou ne ...
 他第一次沒有成工 沒有成工以後 呢
 he first time not-have succeed not-have succeed after TOP
 COMMENT TOPIC
he didn't succeed the first time (and) after not succeeding ...

This sequence is a portion of a chain of such phrases, some of which are clauses, some not, in which the same formula is maintained throughout: '*p*; after *p*, *q*; after *q*, *r*; after *r*, *s*; ... ,' and so on. Li and Thompson claim that TOPIC chains such as the three examples listed here can extend for many utterances with no statements of syntactic SUBJECT, nor any overt indexing of SUBJECT-PREDICATE relationships. Such claims find support in the data used for Study Two.

No material from such TOPIC chains was sampled for Study Two. The reason for this becomes clear when we look, for instance, at the last portion of example (42), “as for the lady, then (she) hurries him away.” On the face of it, the very last TOPIC-COMMENT unit of that chain may have been suitable for inclusion in the data set. However it is questionable when we consider what has preceded the Mandarin utterance in its chain, how comparable such an utterance really is, in terms of the TOPIC constituent’s information value, to an English sentence with the same propositional content. By the point in narration where the last portion of example (42) is uttered, there have already been four phrases referring to the old lady. This may cause the information value of this sentence constituent to differ significantly from that of the SUBJECT of a propositionally similar English SV(O) sentence, in ways that could influence the content of gesture, or even the tendency for gesture to occur at all.

Finally, as to the propositional content of the speech sampled for Study Two, the utterances sampled for the Mandarin utterance category TOPIC-COMMENT are somewhat ‘messier’ than those in the other three categories of utterance (BA, SVO, and SV). This is in the sense of having, in a few cases, more propositional content than any one utterance of the other three categories would typically have. The scope of TOPIC-COMMENT utterances, even when sampling is restricted to sentence-like exemplars, can be more encompassing than the other types sampled. Within the TOPIC-COMMENT sample were four utterances, for instance, that actually contained some portions of two utterance structures, SV or SVO, meshed together.

Gesture-to-utterance ratio

Study Two was based on a sample of 160 utterances. Recall that an effort was made in the speech sampling process to constrain the propositional content of each utterance as much as possible to what is typical of a SV(O) sentential utterance in either language. This sampling criterion allows us to consider whether the two speaker groups differ in the number of gestures they produce relative to roughly equivalent propositional content. The two speaker groups together produced a total of 222 representational gestures to accompany the 160 sampled spoken utterances; thus, overall the ratio of gestures to sentential utterance was higher than one-to-one. On average, across the two speaker groups, the mean number of gestures per utterance was approximately 1.4.

Table 6. Mean number of gestures per utterance (and standard deviation), all utterance structures, Mandarin versus English.

	Mandarin	English	Totals
Gestures per utterance	1.63 (0.7)	1.14 (0.3)	1.38 (0.6)
<i>Number of gestures</i>	<i>130</i>	<i>92</i>	<i>222</i>

Table 6 shows the ratio of gestures-to-utterance by language. We see that the Mandarin speakers' mean ratio is higher than that of the English speakers when all the utterance types are considered together. A two-tailed *t*-test for difference between independent means showed that the difference is statistically significant, $t(113) = 5.47$, $p < .0000$. Table 6a further breaks down the variable of gestures-to-utterance ratio by target utterance type. The table shows that the Mandarin speakers' ratio is higher than that of the English speakers in all utterance structure categories, with TOPIC-COMMENT

utterances accompanied by the highest number of all, an average of more than two gestures per utterance. A within-language one-way analysis of variance was run to

Table 6a. Mean number of gestures per utterance (and standard deviation), by utterance structure, Mandarin versus English.

	Mandarin	English
SV	1.4 (0.4)	1.2 (0.3)
<i>Number of gestures</i>	27	47
SVO	1.4 (0.4)	1.1 (0.3)
<i>Number of gestures</i>	27	45
T-C	2.1 (0.8)	
<i>Number of gestures</i>	42	
BA	1.7 (0.6)	
<i>Number of gestures</i>	34	
Overall	1.6 (0.7)	1.1 (0.3)
<i>Total Gestures</i>	130	92

compare the four Mandarin utterance structures. Table 6b shows that the differences between the mean gesture-to-utterance ratios are significant, $F(3) = 5.85$, $p < .001$. Figure 3 indicates that the TOPIC-COMMENT utterances and their accompanying gestures are the biggest contributor to this difference, while the two SV(O) structures in Mandarin appear to have identical distributions.

Within English, a two-tailed t -test for independent means showed that the very small difference in mean number of gestures accompanying each of the two SV(O) utterance structures is not significant, $t(77) = -0.32$, $p = 0.75$. Collapsing the SV(O)

Table 6b. ANOVA table for gestures per utterance by utterance structure in Mandarin.

Source	DF	SS	MS	F	p
Structure	3	7.650	2.550	5.85	0.001
Error	76	33.100	0.436		
Total	79	40.750			

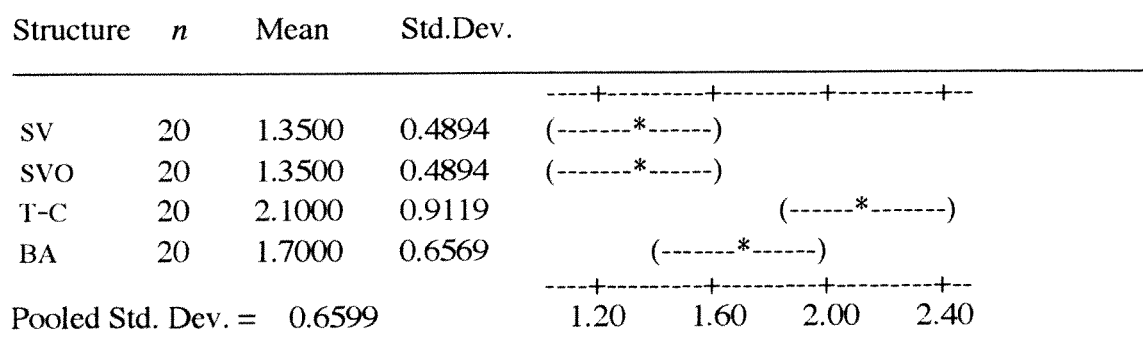


Figure 3. Individual 95 Percent Confidence Intervals for Mean, Based on Pooled Standard Deviation.

categories within each language, however, and comparing the means of gestures per utterance between the two languages, a two-tailed *t*-test for independent means showed the mean numbers to be significantly different, $t(59) = -2.48$, $p = 0.016$. Thus, despite their surface similarity, utterances in the two languages associate with different amounts of gesture.

That the overall ratio of gestures-to-utterance is higher in Mandarin speakers makes it likely that Mandarin speakers more often produce utterances accompanied by more than one gesture than do the English speakers. Table 7 shows this as proportions of total utterances. About half of all the Mandarin utterances sampled are accompanied by multiple gestures, while only 15% of the English utterances are accompanied by more

Table 7. Proportion of utterances accompanied by one gesture versus multiple gestures: Mandarin versus English.

	Mandarin	English
One Gesture	.49	.85
<i>Number of utterances</i>	39	68
Multiple Gestures	.51	.15
<i>Number of utterances</i>	41	12
<i>Total Utterances</i>	80	80

than one. This difference in utterance counts between the two groups is statistically significant, chi-square (1) = 23.728, $p < .000$. Further, of the Mandarin utterances accompanied by multiple-gestures, in 9% of the cases there are more than two gestures. In the Mandarin sample, for example, two of the twenty TOPIC-COMMENT utterances are accompanied by four gestures each. No English utterance in the sample is accompanied by more than two gestures. Analysis of the content of gestures, in terms of the encoding of particular components of events when utterances are accompanied by one gesture or by more than one, will be presented following the more general presentation on gestural encoding of event components, next.

Gestural Encoding of Event Components

To determine whether the larger amount of gesture in the Mandarin speakers is related to whether and how they break out the components of events in gesture, the semantic content of the gestures of the two groups was compared, according to the Isolating versus Incorporating category distinction. Recall that Isolating gestures are those expressive of a single event component — Action or Element. Incorporating

gestures are those that are simultaneously expressive of multiple event components.

Table 8 shows the number of Isolating and Incorporating gestures produced by both the English and Mandarin speakers. Initially, the gestures in these two categories were tallied without regard to the distinctions of Element Only versus Action Only or pantomimic enactments versus non-pantomimics, and without regard to the utterance structure in which they appeared.

Table 8. Isolating versus Incorporating gestures as a proportion of total gestures by language: Mandarin versus English, all utterance structures.

	Mandarin	English	Totals
Isolating	.67	.34	.54
<i>Number of Gestures</i>	87	31	119
Incorporating	.33	.66	.46
<i>Number of Gestures</i>	43	61	103
Totals	.59	.41	1.00
	130	92	222

We see that the general character of event component expression in gesture differs between the two speaker groups. Mandarin and English speakers show almost inverse proportions of such gestural encoding as represented by these two categories of gesture content. Two-thirds of the Mandarin speakers' gestures express an event component in isolation; for example, a reference to a FIGURE (an Element Only gesture) or to the PATH of its movement (an Action Only gesture). In only one-third of the gestures of Mandarin speakers are two or more event components conflated. In contrast, two-thirds of the English speakers' gestures incorporate more than one component. For example, a speaker could represent a FIGURE via her handshake while moving this hand along a

PATH. The differences among the proportions of responses in each category is statistically significant, chi-square (1) = 23.223, $p < .000$.

Table 8a. Element Only, Action Only and Incorporating gestures as a proportion of gestures by language: Mandarin versus English, all utterance structures.

	Mandarin	English	Totals
Element Only	.38	.11	.27
<i>Number of Gestures</i>	50	10	60
Action Only	.29	.23	.26
<i>Number of Gestures</i>	37	21	58
Totals	130	92	222

The findings of the exploratory pilot study had suggested that Mandarin speakers may produce Element Only Isolating gestures in higher proportion than English speakers. To see whether the two speaker groups differ in the content, Element versus Action, expressed in their Isolating gestures, the numbers of these gestures produced by both the English and Mandarin speakers were tallied separately, again initially without regard to the utterance structure the gestures of each type accompanied. The tallies in these two categories, and the proportions these represent of overall gesture production, are presented in Table 8a. The differences in proportions in each of the two categories of Isolating gesture categories is statistically significant (chi-square = 5.202, 1 df, $p < .05$). The tallies in the category of Incorporating gestures from the previous table are included again as well, for comparison. Proportionally more of the Mandarin speakers' total gesture production consists of Element Only gestures than is the case for the English speakers. Of the Mandarin speakers' gestures, 38% are Element Only, compared to 11% of the gestures of the English speakers.

We might expect that the gestures accompanying Mandarin non-SV(O) utterances would contribute in greater proportion than SV(O) utterances in that language to Mandarin's higher proportion of Element Only gestures. This is because, for instance, the BA-construction may be construed as putting a focus on fronted object NPs. Likewise, those TOPIC-COMMENT constructions that have NP TOPICS (the majority in this data set: 85%) may correspond to salience in conception of a character or a physical object in the narrated content that could be gesturally represented via an Element Only gesture.

On the issue of the gestural accompaniment of TOPIC statements, however, recall from the Introduction that McNeill (1992) hypothesizes that spontaneous gestures arise from the GPs of utterances, and that these GPs are where new information is differentiated from presupposed information in discourse. Many theories of topicalization in Mandarin bestow an association between TOPICS and given information. In light of these two theoretical positions, we might expect that the TOPIC statements themselves of TOPIC-COMMENT utterances would not typically be accompanied by gestures at all, or that any gestures that do co-occur with them may have special characteristics.

In order to determine whether target utterance structure is a factor in the different gestures-to-utterance ratios displayed in Table 8a, Table 8b further breaks out the proportions of Incorporating versus Isolating gestures on the basis of utterance category. In this table the category of Isolating gestures is further subdivided between the categories of Element Only and Action Only gestures. The pantomimic enactment sub-category of Incorporating gestures is presented separately as well. Note that in Table 8b the proportions in the Element, Action, and Enactment columns of both the

Table 8b. Isolating (Element Only, Action Only) and Incorporating (distinguishing pantomimic enactment) gestures as a proportion of total gestures within utterance structure category, Mandarin versus English.

	Mandarin					English				
	Isol.	Elem.	Act.	Inc.	Enact.	Isol.	Elem.	Act.	Inc.	Enact.
combined										
SV(O)	.69	.33	.35	.31	.20	.34	.11	.22	.66	.38
(n)	37	18	19	17	11	31	10	21	61	35
SV	.78	.22	.56	.22	.04	.43	.13	.30	.57	.17
(n)	21	6	15	6	1	20	6	14	27	8
SVO	.59	.44	.15	.41	.37	.24	.09	.15	.76	.60
(n)	16	12	4	11	10	11	4	7	34	27
combined non-										
SV(O)	.67	.42	.25	.33	.27					
(n)	50	32	19	26	21					
T-C	.76	.52	.24	.24	.17					
(n)	32	22	10	10	7					
BA	.53	.29	.26	.47	.38					
(n)	18	10	9	16	13					
Total Gestures					130					92

Mandarin and the English data sets are calculated in the same way as are the bold face columns of Isolating versus Incorporating gestures; that is, as proportions of the total number of gestures within each utterance category. Thus for instance, Table 8b shows that 76% of the forty-two gestures that accompany Mandarin TOPIC-COMMENT utterances are Isolating; 52% of those same forty-two total gestures are Element Only encoding; 17% are pantomimic enactments.

Table 8b shows again that for Mandarin speakers across all target utterance categories, the proportions of Isolating gestures are higher than those of Incorporating gestures. In each utterance category of English, the proportion of Incorporating gestures is the higher. Examining the Mandarin data further and beginning with the non-SV(O)

categories, we see that TOPIC-COMMENT and BA utterances differ in terms of the proportions and in terms of the content of the Isolating gestures with which they associate. The difference in proportions of Isolating versus Incorporating gestures that accompany TOPIC-COMMENT utterances, 76% versus 24%, is much greater than the difference in the corresponding proportions of gestures that accompany BA-constructions, 56% Isolating versus 44% Incorporating; further, the larger portion of this difference is accounted for by a higher proportion of Element Only gestures accompanying TOPIC-COMMENT utterances. Two-thirds of the Isolating gestures that accompany TOPIC-COMMENT utterances are Element Only. In contrast, Element Only and Action Only each make up about half of the Isolating gestures that accompany BA-constructions. More will be said below concerning the specific event components typically indexed or depicted by the Isolating and Incorporating gestures that accompany BA-constructions and TOPIC-COMMENT utterances.

Examining now the Mandarin SV(O) utterances, recall that it seemed reasonable to suppose that SV and SVO utterances might *not* elicit high proportions of Element Only gestures. This a theory-motivated expectation. These utterance structures are, unlike BA-constructions and TOPIC-COMMENT utterances, not generally theorized to be the result of anything like a topicalizing or focusing movement transformation. The construction resulting from such a movement transformation could be assumed to single out one constituent, usually a nominal, from the rest. The manifestation of this in the spoken sequence is overt linguistic marking. If the theoretical analysis of these utterances is correct, we might expect gestural highlighting of something related to the marked linguistic constituent. Since theoretically, SV and SVO utterance structures reflect

the base, unmarked word order of the Mandarin, we therefore might not expect these to associate with any particular gestural marking of any of the case roles. Further, given that the SV(O) structure is basic to English, the expectation based on theory is that the gestural performances accompanying Mandarin SV and SVO utterances would be similar to those accompanying such utterances in English.

When the SV and SVO utterance structures are considered in combination in Table 8b, the supposition that they do not associate with a high proportion of Element Only gestures is supported. Of all the Isolating gestures that accompany the combined Mandarin SV(O) utterances, Element Only and Action Only each account for about half. However a difference emerges when we examine the SV utterances separately from the SVOs. There we see roughly inverse proportions of Element Only and Action Only gestures. Of the Isolating gestures that accompany SV utterances, slightly fewer than one-third are Element Only. Element Only gestural accompaniments to SV utterances would typically be representations of FIGURE or, less frequently, GROUND. In contrast, of the Isolating gestures that accompany SVO utterances, three-quarters are Element Only. Element Only gestural accompaniments to SVO utterances would be representations of AGENT/FIGURE, PATIENT, or GROUND. Thus, similarly to the TOPIC-COMMENT utterances, and contrary to expectation, the SVO utterances in this Mandarin sample associate with a relatively high proportion of Element Only gestures; BA-constructions a relatively low proportion. The SV context stands out as associating predominantly with Action Only Isolating gestures.

Table 8c. The utterance structure categories of Mandarin, in descending order, the proportions of Element Only gestures within the category of Isolating gestures.

	<u>Proportions of</u> <u>Isolating gestures that are:</u>		<u>All gestures:</u>		<i>Total Gestures</i>	<i>Total Utterances</i>
	Element Only	Action Only	Isolating	Incorp.		
SVO (n)	.75 12	.25 4	.59 16	.41 11	27	20
T-C (n)	.69 22	.31 10	.76 32	.24 10	42	20
BA (n)	.53 10	.47 9	.56 19	.44 15	34	20
SV (n)	.29 6	.71 15	.78 21	.22 6	27	20
					130	80

Table 8c displays just the Mandarin subset of the data of Table 8b, rearranged so that the proportions of Element Only gestures in the four Mandarin utterance structure categories appear in descending order down the left side of the table. Note that these proportions in this table are calculated differently from the related values in Table 8b. In Table 8c, the proportions 75%, 69%, 53%, and 29% Element Only gestures that accompany SV(O), TOPIC-COMMENT, BA, and SV, respectively, are proportions within the category of Isolating gestures separately, as opposed to proportions of the total gestures accompanying each category of utterance, as was the case in Table 8b.

Table 8c brings out two additional patterns in the Mandarin data. First, in regard to their relative proportions of Isolating versus Incorporating gestures, SVO utterances and BA-constructions are similar. The same is true of TOPIC-COMMENT and SV utterances.

We know from the preceding, however, that the Mandarin SVO and BA utterance

categories differ in terms of which type of Isolating gesture each is more likely to associate with. The SVO utterances associate with a higher proportion of Element Only gestures; BA-constructions with a higher proportion of Action Only. Likewise, despite similar proportions of Isolating gestures, the relative proportions of Element Only and Action Only gestures that accompany TOPIC-COMMENT and SV utterances differ. Of the Isolating gestures that accompany TOPIC-COMMENT utterances, 69% are Element Only; of those that accompany SV utterances, only 29% are Element Only.

Second, considering the total number of gestures elicited per utterance category, and comparing the TOPIC-COMMENT and SV utterance categories, it is clear that the majority of SV utterances are accompanied by just a single gesture, and the majority of those are Action Only. In contrast, on average, TOPIC-COMMENT utterances are accompanied by slightly more than two gestures. Adding up by sub-category the gesture counts listed in Table 8c, it appears that this multi-gestural accompaniment to TOPIC-COMMENT utterances could consist of combinations of Element Only gestures together with gestures from either the Action Only or the Incorporating categories. A summary of the findings in regard to sequences of gesture in utterances accompanied by two or more gestures will be given presently.

Returning to Table 8b to consider the data on gestures that accompany English SV and SVO utterances, we see that the difference in proportions of Isolating versus Incorporating gestures that accompany English SVO utterances (24% versus 76%) is somewhat greater than the difference in the corresponding proportions of these gesture types that accompany SV utterances (43% versus 5%). Since the English speakers

produce relatively few Isolating gestures overall, the numbers are small when this category of gestures is divided, however the proportions of Element Only and Action Only gestures that accompany both types of English utterances are roughly the same. The English speakers produce about twice as many Action Only gestures as Element Only gestures to accompany utterances of both categories, SVO and SV. Table 8d is the English subset of the data from Table 8b, transformed in the same way as was the Mandarin data in Table 8c. The English speakers' twenty Isolating gestures that accompany SV utterances are 70% Action Only. The eleven Isolating gestures that accompany English SVO utterances are 64% Action Only.

Table 8d. The utterance structure categories of English: In descending order, the proportions of Element Only gestures within the category of Isolating gestures.

	<u>Proportions of</u> <u>Isolating gestures that are:</u>		<u>All gestures:</u>		<i>Total Gestures</i>	<i>Total Utterances</i>
	Element Only	Action Only	Isolating	Incorp.		
SVO (n)	.36 4	.64 7	.24 11	.76 34	45	40
SV (n)	.30 6	.70 14	.43 20	.57 27	47	40
					92	80

When we compare the proportions of the two categories of Isolating gestures that accompany SV and SVO structures in Mandarin and English (Tables 8c and 8d), we see that the relative proportions are almost identical for SV, but show opposite tendencies for

SVO. Of the Isolating gestures that accompany SVO utterances, a high proportion in Mandarin are Element Only; a comparatively low proportion in English are Element Only. Thus, although the ordering of constituents is the same in the two languages in this target utterance sample, the two speaker groups produce different gestural accompaniments to their SVO utterances.

One final pattern that emerges in the data summarized in Table 8b concerns the proportions of pantomimic enactments across languages and utterance structure contexts. In all categories but one, such enactment gestures are in the minority of gesture production. In English SVO utterance contexts they are the majority, representing 60% of all gestures produced. Approximately equal proportions of the gestures in Mandarin SVO and BA utterance contexts are enactments, 37% and 38% respectively. The categories with the lowest proportions of this kind of gestural reference are English SV, Mandarin TOPIC-COMMENT, and Mandarin SV, with 17%, 17%, and 4% pantomimic enactments, respectively.

Single-Gesture Versus Multi-Gesture Utterance Accompaniments

Tables 8 and 8 *a* through *d* dealt with gesture types and gesture content as proportions of total gesture production, by language and utterance category. Coming up next, Tables 9 and 9 *a* through *c* address the issue of the content of the gestures in single-versus multi-gesture accompaniments of utterances as proportions based instead on total utterances, by language and utterance category. Table 9 divides the 80 utterances from each language according to the number and category — Isolating versus Incorporating — of the accompanying gestures. To see how this table works, note that the extreme left

Table 9. The Isolating and Incorporating gestural accompaniments of all 80 utterances of each language, enactments included, all utterance categories, Mandarin versus English.

	ISOL	INC	ISOL INC	INC INC	ISOL ISOL	ISOL INC	ISOL ISOL	ISOL INC	ISOL ISOL	ISOL ISOL
Mandarin	.33	.16	.14	.08	.21	.01	.03	.03	.01	.01
<i>N of utterances</i>	26	13	11	6	17	1	2	2	1	1
English	.20	.65	.09	.05	.01	---	---	---	---	---
<i>N of utterances</i>	16	52	7	4	1					

side shows that 33% of the 80 Mandarin utterances are accompanied by a single Isolating gesture, compared to 20% of the 80 English utterances. The right side of Table 9 shows that two Mandarin utterances are accompanied by four gestures each. In one case, these comprise three Isolating gestures and one Incorporating gesture; in the other they are all Isolating gestures. In between, Table 9 displays the other combinations that were observed, and the proportion each represents of the 80 utterances in each language.

As has already been shown, English speaker single-gesture utterance accompaniments are largely of the Incorporating type. Only twelve English utterances in the data set are accompanied by more than one gesture, compared to forty-one utterances in the Mandarin data set. The proportions based on the small number of English speaker multi-gesture accompaniments indicate that it may be the case that when these speakers add gestures they tend to be the Isolating type, while the Mandarin speakers show a general preference overall for Isolating gestures in both single- and multi-gesture utterance accompaniments. Table 9a makes this last point more explicitly by summarizing the same data of Table 9, but with the multi-gesture accompaniments

Table 9a. The Isolating versus Incorporating gestural accompaniments of all 80 utterances of each language in single versus multi-gesture accompaniments, enactments included, all utterance categories, Mandarin versus English.

	<u>Single gesture</u>			<u>Multi-gesture</u>			
	ISOL	INC	Totals	ISOL INC	INC INC	ISOL ISOL	Totals
Mandarin	.33	.16		.19	.08	.25	
<i>Number of utterances</i>	26	13	39	15	6	20	41
English	.20	.65		.09	.05	.01	
<i>Number of utterances</i>	16	52	68	7	4	1	12

collapsed to provide a better sense of their general composition. As presented in Table 9a, any multi-gesture accompaniment that has both Isolating and Incorporating gestures in it, regardless of how many of each of these there are, is included in the first column of the right half of the table. Any multi-gesture accompaniment comprising only Incorporating gestures is summarized in the center column of the right half. The rightmost column summarizes those multi-gesture accompaniments that comprise only Isolating gestures. The data summarized in this way support the finding that Mandarin speakers throughout rely less on Incorporating gestures than do English speakers, and suggests further that for English speakers Isolating gestures are typically used as add-ons to Incorporating gestural performances.

Table 9b presents the same data of Table 9a, but with the category of Isolating gestures separated into Element Only and Action Only encoding gestures. Among the multi-gesture utterance accompaniments in the English sample, the highest proportion is the conjunction of an Element Only gesture and an Incorporating gesture. In Mandarin the highest proportion is combinations of Action Only and Element Only, followed by

Table 9b. Action Only, Element Only and Incorporating gestural accompaniments of all 80 utterances of each language, in single versus multi-gesture accompaniments, enactments included, all utterance categories, Mandarin versus English.

	ACT	ELEM	INC	ACT INC	ELEM INC	INC	ACT ELEM	ACT ACT	ELEM ELEM	ACT ELEM INC
Mandarin	.24	.09	.16	.03	.15	.08	.21	---	.04	.01
<i>N of utterances</i>	19	7	13	2	12	6	17	---	3	1
English	.20	---	.65	---	.09	.01	.04	.01	---	---
<i>N of utterances</i>	16		52		7	1	3	1		

Element Only plus Incorporating. Thus in both groups it appears that there is a tendency in multi-gesture accompaniments to make explicit gestural references both to action and to some element or elements that participate in the action.

Table 9c makes it possible to determine whether the tendency to produce Isolating versus Incorporating gestures by themselves as single-gesture utterance accompaniments versus together with other gestures as multi-gesture accompaniments varies with utterance structure. This table presents the data that was organized in Tables 9 and 9a according to the categories of Isolating versus Incorporating gestures, broken out by target utterance category. Table 9c shows that Isolating gestural references appear to be promoted in Mandarin SV, SVO, and TOPIC-COMMENT utterance contexts in both single- and multi-gesture accompaniments. This appears not to be the case in either of the English contexts. Further, despite their presumably comparable verb argument structures, the gestural accompaniments of Mandarin SV and SVO utterances differ from those of English SV and SVO. Both of these utterance categories in Mandarin associate with more Isolating gestural references than do English SVs and SVOs, in both single- and

synthetically encoding and so by default would incorporate all or most of the event components that any Isolating gesture might individually highlight.

Encoding of Event Components and Case Roles in Gesture

The data on Incorporating and Isolating, Element and Action gestures that co-occur with utterances either singly or in multiple-gesture productions, give coarse-grained answers to the question of what components of the events expressed in utterances are singled out for gestural depiction or indexing. Isolating gestures may encode event components such as AGENT/FIGURE, ACTION, or PATIENT, GROUND, or others. A finding so far, for instance, is that a relatively high proportion of the gestures accompanying Mandarin TOPIC-COMMENT utterances are Element Only Isolating gestures. On the basis of this alone we cannot know whether such Element Only gestures may disproportionately represent AGENT/FIGURE, PATIENT, GROUND or other components in Mandarin TOPIC-COMMENT utterance contexts. Incorporating gestures may be featurally marked for these same components. The following production from the Mandarin data set contains both an Isolating and an Incorporating gesture.

- (45) [/ ta ke-yi **pa**] [nei-ge / **shui-guan a** /]
 / 他可以 爬 那個 / 水管呵 /
 / he can climb that-CL / water-pipe A /
 / he can climb the / drainpipe y'know

Here the speaker re-tells the cartoon event of Sylvester the cat climbing up a building's drainpipe to reach Tweety. The first gesture synchronizes with the main verb of the utterance, "pa"/"climb." It is a deictic gesture to a location in the speaker's gesture space where previous gestures related to the drainpipe have been located, thus it is an Element Only Isolating gesture that indexes 'drainpipe.' The second gesture synchronizes with

the GROUND component in speech, “shui-guan”/“drainpipe,” and is an iconic depiction of the cat climbing. It is an Incorporating gesture that encodes Action (PATH and MANNER) and Element (FIGURE) in combination.

The summary in Table 10 addresses the issue of which components of the events expressed in utterances are gesturally indexed or depicted in the different utterance contexts. The table is organized to show gestural references to specific event components by language and utterance structure. In Table 10 there is a return to the procedure of calculating proportions based on total numbers of gestures by category, as opposed to numbers of utterances. Pantomimic enactment gestures are summarized separately in this table. Because of their extreme incorporating character, enactments do not selectively focus on individual components of events; rather, such gestures present the entirety of an event in a highly synthetic fashion. Isolating or Incorporating gestures that make specific reference to some particular event component or components make clear that those are a focus of a speaker’s conceptual representation. Pantomimic enactments lack this specificity.

A word is necessary about the different gesture totals on which proportions are based in different parts of Table 10. In the two sections of the table that present overall comparisons of enactment versus non-enactment gestures — Mandarin at the top and English below the middle of the table — the proportions are calculated with reference to the total number of gestures by language and within language by utterance category. These gesture totals for each language, 130 versus 92 for Mandarin and English, respectively, have already been the basis for summaries of data in other tables, as have their respective subtotals by utterance category. In contrast, in the two sections of the Table 10 where gestural marking of particular event components in non-pantomimics is summarized, the proportions are based instead on total instances of gestural marking of

Table 10. Proportions of gestures marked for particular event components, by utterance category, Mandarin versus English. Enactments and non-enactments separated.

	SV	SVO	T-C	BA	Totals
<u>Mandarin enactments:</u>					
<i>Proportion of enactments by category:</i>	.04	.37	.17	.41	.25
<i>Number of enactments by category:</i>	1	10	7	14	32
<i>Total Mandarin gestures:</i>	27	27	42	34	130
<u>Mandarin non-enactments:</u>					
<i>Proportion of non-enactments by category:</i>	.96	.63	.83	.59	.75
<i>Number of non-enactments by category:</i>	26	17	35	20	98
<hr/>					
Action (n)	.65 20	.28 5	.32 13	.50 13	.44 51
Agent/Figure (n)	.23 7	.05 1	.34 14	0 0	.19 22
Patient (n)	n/a	.67 12	.15 6	.42 11	.25 29
Other (n)	.13 4	0 0	.20 8	.08 2	.12 14
<hr/>					
<i>Total component marking in non-enacts:</i>	31	18	41	26	116
<hr/>					
<u>English enactments:</u>					
<i>Proportion of enactments by category:</i>	.17	.60			.38
<i>Number of enactments by category:</i>	8	27			35
<i>Total English gestures:</i>	47	45			92
<u>English non-enactments:</u>					
<i>Proportion of non-enactments by category:</i>	.83	.40			.62
<i>Number of non-enactments by category:</i>	39	18			57
<hr/>					
Action (n)	.55 36	.50 15			.54 51
Agent/Figure (n)	.32 21	.27 8			.30 29
Patient (n)	n/a	.17 5			.05 5
Other (n)	.12 8	.07 2			.11 10
<hr/>					
<i>Total component marking in non-enacts:</i>	65	30			95

the specific components. These totals reflect only marking that occurs in non-pantomimic gestures.

Note that the total instances of component marking summarized in this section of Table 10 exceed the total number of non-pantomimic gestures. This is because each Incorporating gesture in the set makes explicit gestural reference to more than one event component. To understand this aspect of the data summarized in Table 10, consider the Mandarin speaker totals on the top half of the right side of the table. Overall, the Mandarin speakers produced 130 gestural accompaniments to the 80 utterances sampled from their narrations. Thirty-two of these, or 25%, are pantomimic enactments that were kept separate from the calculations on which the event component proportions are based in Table 10. Of the remaining 75%, or ninety-eight non-pantomimic gestures that the Mandarin speakers produced, some are Isolating and some Incorporating. At the midpoint on the right side of Table 10, the total 116 is the overall number of explicit gestural references to particular event components that these ninety-eight non-pantomimic gestures encoded.

The patterning of pantomimic enactment gestures versus non-pantomimics that has been outlined previously in this presentation emerges in Table 10 as well. Accompanying the English SV and SVO utterances, pantomimic enactments and non-pantomimics occur in roughly inverse proportions. The relative proportions of pantomimic enactments and non-pantomimics accompanying English SVOs are inverse also to those accompanying Mandarin SVO and BA utterances. These are the two utterance structures in Mandarin that have the same verb argument structure as English SVO utterances, yet neither of them lines up with English on this dimension of gesture

patterning. Of the two Mandarin structures, SVO and BA, BA-constructions may associate with slightly more pantomimic gestures than do SVO utterances. Across utterance structures and languages, however, English SVO utterances stand out as favoring pantomimic enactments.

Note that from Table 10 we may also derive that, in spite of the fact that the entire corpus of Mandarin utterances is accompanied by 41% more gesture than that of the English (130 versus 92 gestures in total), it is unlikely that the Mandarin speakers gesturally encode proportionally more propositional content than do the English speakers. To see why this is true, consider first that the higher proportion of pantomimic enactments in the English speakers' gestures — 38% to Mandarin speakers' 25%, overall — increases the overall level of gestural component encoding across the English speakers' productions. Pantomimic enactments are more densely encoding than Isolating gestures or observer-viewpoint Incorporating gestures that usually conflate just two components of an event in one gestural production. Second, Table 10 shows that when pantomimic enactments are left out of consideration, and only those explicit gestural references to event components encoded by Isolating and observer-viewpoint Incorporating gestures are tallied, the English speakers' total across these categories comes closer to that of the Mandarin speakers than is the case when just overall number of gestures is considered. The English speakers produce 95 explicit gestural references to event components in a total of 57 non-enactment gestures; Mandarin speakers produce 116 in a total of 98. So whereas Mandarin speakers produced 41% more gestures overall, they produced only 22% more explicit gestural references to event components in just the categories of non-pantomimic gestures — Isolating or observer-viewpoint Incorporating gestures. The English speakers' higher overall proportion of pantomimics further redresses this imbalance, as will be seen presently in the presentation of the data

summarized in Table 11.

In regard to the encoding of event components, several tendencies are identifiable from the summary in Table 10. In English speaker non-enactment gestures, ACTION and AGENT/FIGURE are favored in both SV and SVO utterance contexts in roughly similar proportions, about 50% and 30%, respectively. Explicit gestural references to PATIENTS are relatively low frequency in the gestural accompaniments to English SVO utterances. By comparison, in the Mandarin data set the findings are more varied. First, the overall proportion of PATIENT references is high in Mandarin gestural accompaniments compared to those of English. Most of this is accounted for by specific PATIENT-marking in the gestures that accompany SVO and BA utterances. Note that specific gestural PATIENT references are actually proportionally higher in Mandarin SVO than in BA-construction contexts, a finding contrary to the expectation that PATIENTs might be differentially highlighted in the gestural accompaniments to BA-constructions. In fact, specific gestural references to ACTION in BA-construction contexts slightly exceed those to PATIENT. The proportions of Mandarin speakers' gestures that encode ACTION are relatively high in SV and BA utterance contexts; somewhat lower in accompaniments to SVO and TOPIC-COMMENT utterances. No AGENT/FIGURE-marking occurs in BA utterance contexts in this sample, and it occurs only once in a Mandarin SVO context. In contrast, there is a non-negligible amount of gestural AGENT/FIGURE-marking in Mandarin SV and TOPIC-COMMENT contexts. As well, there is a non-negligible amount of gestural AGENT/FIGURE-marking in English SV and SVO contexts. In this we see

another dimension of gesture patterning in which the two utterance structures in Mandarin, SVO and BA, that have the same two-place verb argument structure as English SVO, do not line up with English.

It should be mentioned here in regard to TOPIC-COMMENT utterances that the proportions of event component encoding in their gestural accompaniments are likely influenced by the heterogeneity of underlying utterance structures encompassed by this sample of twenty utterances. Recall that it was earlier noted that the TOPIC-COMMENT subset of the Mandarin sample is ‘messier’ than the other three subsets, because it includes a variety of verb argument structures. Of this sample, 55% of utterances are organized around two-place verbs, another 25% around one-place verbs. Four utterances in the sample of twenty TOPIC-COMMENT utterances, or 20% of the sample, appear to include pieces of two propositions of either the SV or the SVO variety, meshed together. The comparatively more evenly distributed gestural references to event components in accompaniments to the TOPIC-COMMENT utterances is likely related to this heterogeneity.

Table 10a presents once more a subset of the data summarized in Table 10. This is the subset that comprises all non-enactment gestures, broken out according to the categories of Isolating versus Incorporating gestures. The summary organized in this fashion makes it possible to see whether high proportions of isolating gestural references to specific event components or case roles are a feature of particular utterance contexts. The proportions presented are calculated on the basis of total number of gestures by utterance category. This differs from the procedure used for this same subset of the data in Table 10, where the proportions were based on the total number of instances of gestural expression of event components, rather than on number of gestures.

Table 10a. Proportions of Isolating versus Incorporating gestures marked for particular event components, by utterance category, Mandarin versus English. Pantomimic enactments excluded.

	SV		SVO		T-C		BA	
	Iso.	Inc.	Iso.	Inc.	Iso.	Inc.	Iso.	Inc.
<u>Mandarin:</u>								
Action (n)	.71 15	1.0 5	.25 4	1.0 1	.31 10	1.0 3	.50 9	2.0 4
Agent/Figure (n)	.19 4	.60 3	.06 1	0 0	.40 13	.33 1	0 0	0 0
Patient (n)	n/a	n/a	.69 11	1.0 1	.16 5	.33 1	.50 9	1.0 2
Other (n)	.10 2	.40 2	0 0	0 0	.13 4	1.33 4	0 0	1.0 2
Total marking:	21	10	16	2	32	9	18	8
Total non-enacts:	21	5	16	1	32	3	18	2
<u>English:</u>								
Action (n)	.70 14	1.16 22	.64 7	1.14 8				
Agent/Figure (n)	.10 2	1.0 19	.27 3	.71 5				
Patient (n)	n/a	n/a	.09 1	.57 4				
Other (n)	.25 4	.21 4	0 0	.29 2				
Total marking:	20	45	11	19				
Total non-enacts:	20	19	11	7				

There is a peculiarity in the proportions presented in Table 10a that must be explained. This peculiarity is a result of two factors. First, this summary is based on data from which all the gestures containing pantomimic enactments have been removed.

Some of the enactments removed, however, are themselves part of gestural incorporation in which another event component is gesturally encoded as distinct from and in addition to the enactment encoding. The gestural productions in these cases are of the ‘dual viewpoint’ category, such that the speaker pantomimes an action with body and hand while simultaneously featurally marking one of the components of the pantomimed event in one hand. Such complex gestural expressions are not frequent in the data but they do occur, and in the analyses presented so far they have been eliminated from the gesture counts of non-pantomimics. Nevertheless, for the purposes of Table 10a, the extra event component encoding that is conflated with these gestures is included, since it does count as explicit gestural marking that occurs in the context of Incorporating gestures. What this means for the proportions presented in Table 10a is apparent in the English section of the table where proportions of gestural encoding of event components *higher* than 100% appear in two instances. This occurs twice in the Mandarin section as well, in the data for TOPIC-COMMENT and BA-constructions. For example, relative to a total of nineteen non-enactment Incorporating gestures in English, there are twenty-two instances of gestural encoding of ACTION. This is the result of including in these counts those few instances of additional event component marking that are conflated with enactment gestures.

Table 10a makes evident once again that the proportions of non-pantomimic Incorporating gestures accompanying the Mandarin utterances of each category are low relative to the proportions of Isolating gestures. Reading left to right on the line labelled ‘Total non-enacts,’ we see that there are only five, one, three and two incorporating gestures accompanying SV, SVO, TOPIC-COMMENT, and BA utterances, respectively. Given such small numbers of gestures, it may not be appropriate to interpret their relative proportions of event component expression. Examining therefore just the proportions of

Isolating gestures in each Mandarin utterance category, we see that Mandarin SV utterance contexts associate with isolating ACTION references; SVO contexts PATIENT references; BA contexts both ACTION and PATIENT references; TOPIC-COMMENT contexts perhaps both AGENT/FIGURE and ACTION references.

In regard to the English data summarized in Table 10a, it emerges that, when pantomimic enactments are removed, English SVO utterances turn out to be accompanied overall by fewer Incorporating than Isolating gestures (seven versus eleven gestures, respectively). Thus the preponderance of Incorporating gestures accompanying English SVO utterances is accounted for by the high proportion of pantomimic enactments with which they associate. In contrast, SV utterance contexts associate with a high proportion of Incorporating gestures even when enactments are removed. In regard to isolating gestural references to event components, Table 10a shows that appreciable proportions of the Isolating gestures in English SV and SVO are ACTION references. Incorporating gestures in SV contexts highlight ACTION and AGENT/FIGURES. Since the number of non-pantomimic Incorporating gestures in English SVO contexts is small, the proportions appear exaggerated; however, ACTION is encoded in every gesture, and it may be that AGENT/FIGURE is no more highlighted in SVO contexts than is PATIENT.

Thus, when pantomimic enactments are excluded and gestural highlighting of event components and case roles is considered, it appears that there is more explicit gestural encoding of the PATIENT role in the Mandarin speakers overall than in the English speakers. There are two important factors to consider when interpreting this finding. The first is that Mandarin SVO utterances appear especially to be part of an

environment for this emphasis on PATIENT-marking. The second is that, although the Study Two sample has equal numbers of Mandarin SVO utterances and BA-constructions, recall that the SVO utterances were especially difficult to find, whereas BA-constructions were abundant. BA-constructions appear to be part of an environment for gestural ACTION-marking, and they are much more frequent than SVO utterances in the narrations of the Mandarin speakers. Thus what looks like an emphasis on PATIENT-marking overall in the samples examined for Study Two would likely be minimized if the target utterances were sampled in proportion to their actual rate of occurrence in the narrations.

Event Component Encoding in Single- Versus Multi-gesture Utterance Accompaniments

Finally, Table 11 addresses the issue of the total event component encoding in all utterance contexts, and in single- versus multi-gesture accompaniments. What distinguishes the proportions presented in Table 11 from those of previous tables is that they incorporate the contributions of the pantomimic enactment gestures that accompany utterances in each utterance category. Because enactment gestures are densely-encoding, their inclusion in the tallies of course elevates the proportions of gestural references to all relevant event component categories where they occur. Because they occur disproportionately in different utterance contexts, including them allows us to see how their contribution differentially impacts the proportions of encoding of different event components across contexts. The presentation of the data until now has focused on sorting out which components are specifically highlighted, using which categories of gesture. Such an analytic focus on gestural highlighting leads to an impression of an imbalance in overall amount of component encoding cross language. Table 11 shows that this is probably not an accurate impression; rather, that differences in event

Table 11. (Page 1 of 2) Proportions of gestural event component encoding, including the contributions of enactments, by utterance category, Mandarin versus English.

Single-gesture accompaniments						Multi-gesture accompaniments				Grand
	Isol.	Inc.	Enact.	Total	Isol.	Inc.	Enact.	Total	Total	
Mandarin:										
SV 20 utts.	Act. (n)	.69 9	.23 3	.08 1	.81 13	.86 6	.14 1	---	.50 7	.67 20
	A/F (n)	---	.50 1	.50 1	.13 2	.80 4	.20 1	---	.36 5	.23 7
	Oth. (n)	---	1.0 1	n/a	.06 1	1.0 2	---	n/a	.14 2	.10 3
	Total								16	14
SVO 20 utts.	Act. (n)	.38 3	---	.63 5	.35 8	.17 1	---	.83 5	.27 6	.31 14
	A/F (n)	---	---	1.0 5	.22 5	.17 1	---	.83 5	.27 6	.24 11
	Pa. (n)	.50 5	---	.50 5	.43 10	.50 5	---	.50 5	.45 10	.44 20
	Oth. (n)	---	---	n/a	0 0	---	---	n/a	0 0	0 0
Total					23	22			45	
T-C 20 utts.	Act. (n)	1.0 3	---	---	.60 3	.41 7	.18 3	.41 7	.31 17	.33 20
	A/F (n)	---	---	---	0 0	.56 10	.06 1	.39 7	.33 18	.30 18
	Pa. (n)	1.0 1	---	---	.20 1	.33 4	.08 1	.58 7	.22 12	.22 13
	Oth. (n)	1.0 1	---	n/a	.20 1	.50 4	.50 4	n/a	.15 8	.15 9
Total					5	55			60	

Table 11. (Page 2 of 2) Proportions of gestural event component encoding, including the contributions of enactments, by utterance category, Mandarin versus English.

		<u>Single-gesture utterances</u>				<u>Multi-gesture utterances</u>				<i>Grand</i>	
		Isol.	Inc.	Enact.	<i>Total</i>	Isol.	Inc.	Enact.	<i>Total</i>	<i>Total</i>	
<u>Mandarin:</u> (Continued)											
BA 20 <i>utts.</i>	Act. (<i>n</i>)	.44 4	.11 1	.44 4	.53 9	.28 5	.17 3	.56 10	.36 18	.40 27	
	A/F (<i>n</i>)	---	---	1.0 4	.24 4	---	---	1.0 10	.20 10	.21 14	
	Pa. (<i>n</i>)	---	---	1.0 4	.24 4	.43 9	.10 2	.48 10	.42 21	.37 25	
	Oth. (<i>n</i>)	---	---	n/a	---	---	1.0 1	n/a	.02 1	.01 1	
<i>Total</i>					17					50	67
<u>English:</u>											
SV 40 <i>utts.</i>	Act. (<i>n</i>)	.28 10	.53 19	.19 7	.58 36	.50 4	.38 3	.13 1	.44 8	.55 44	
	A/F (<i>n</i>)	---	.70 16	.30 7	.37 24	.33 2	.50 3	.17 1	.33 6	.36 29	
	Oth. (<i>n</i>)	---	1.0 3	n/a	.05 3	1.0 4	---	n/a	.22 4	.09 7	
<i>Total</i>					62					18	80
SVO 40 <i>utts.</i>	Act. (<i>n</i>)	.17 6	.17 6	.66 23	.38 35	.14 1	.29 2	.57 4	.35 7	.37 42	
	A/F (<i>n</i>)	---	.21 6	.79 23	.31 29	.43 3	---	.57 4	.35 7	.32 36	
	Pa. (<i>n</i>)	---	.12 3	.88 23	.28 26	.17 1	.17 1	.67 4	.30 6	.28 32	
	Oth. (<i>n</i>)	---	1.0 3	n/a	.03 3	---	---	n/a	---	---	
<i>Total</i>					93					20	113

component encoding are an issue of how the encoding is achieved.

Note that the proportions reported in this table are once again based on the total number of instances of gestural reference to event components by language and utterance category, rather than on proportions of total gestures. Note further that each enactment gesture is credited with augmenting the total instances of gestural reference to event components across the board within its production. That is, a single enactment gesture accompanying, for example, a SVO utterance, is coded as augmenting by one the number of gestural references to each of the three event components that such an utterance encodes — ACTION, AGENT, and PATIENT. In contrast, an observer viewpoint incorporating gesture may conflate only the event components ACTION and AGENT, for example, and in that case only the values of those component categories are augmented.

Another issue should be mentioned concerning event component encoding in pantomimic enactments. It is obvious how a pantomime, for example, of the act of hitting something with an umbrella would by default encode ACTION and AGENT roles, as well as potentially the INSTRUMENT role, by virtue of a handshape that is configured in a way appropriate to holding an umbrella. It may be less obvious how, when a speaker pantomimes such an action, the PATIENT role is gesturally encoded as well. The manifestation of this role is a function of the spatial frame in which such a gesture occurs. That is, the pantomimic action of striking downward with an elongated object is not considered to be randomly directed in gesture space; rather, the striking motion is directed at some location in that space. Typically a speaker's gaze is directed at the location as well. This location itself may therefore be construed as encoding the PATIENT.

The different proportions in Table 11 must be read with reference to two different sets of totals. This will be illustrated with an example from the portion of the table that summarizes the data on Mandarin SV utterances. From the top left corner reading down, the event components ACTION, AGENT/FIGURE, and Other are listed. Examining the line of proportions that extends rightward from ACTION ('Act.'), note that it is divided between gestures that occur singly and those that occur together with other gestures in multi-gesture utterance accompaniments. Within the single-gesture accompaniments section, the proportions of gestural references to ACTION are 69% Isolating, 23% are part of Incorporating gestures, and 8% are expressed as part of enactments. These are all proportions of the total number of event component references encoded in gestures that occur singly as accompaniments to Mandarin SV utterances. Thirteen of the Mandarin SV utterances are accompanied by single gestures and every one of these encodes the component of ACTION in some way.

From left to right the fourth column in the *single-gesture accompaniments* section of Table 11 is labeled 'Total' and this column displays the totals of all gestural references to each event component. These totals are themselves totalled at the bottom of the section for each utterance category. In the case of Mandarin SV utterances, there is a total of sixteen references to event components — ACTION, FIGURE, and Other — among the thirteen gestures that singly accompany utterances of the SV category. The proportions of each event component in the Total column are calculated based on the figure at the bottom of the column within the SV section of the table. Thus, 81% of the sixteen total gestural references to event components in SV contexts accompanied by single gestures

are references to ACTION.

Moving rightward, the four columns within the *multi-gesture accompaniments* section of table function in the same way as the set just described. Finally, the rightmost column of the table, labelled 'Grand Total,' shows the combined totals of gestural references to specific event components from both single-gesture and multi-gesture utterance accompaniments. The proportions listed in the Grand Total column are based on the figure at the bottom of this column within the Mandarin SV section of the table.

We see that 67% of all gestural references to particular event components in Mandarin SV contexts are references to ACTION and 23% are references to FIGURE, and this proportion includes references from Isolating and Incorporating gestures, observer viewpoint and pantomimic enactment gestures.

In summary, by reading across, one is given the proportions in which each event component, for instance, AGENT or PATIENT, is encoded by different gesture types in single-gesture versus multi-gesture accompaniments to utterances. By reading down in the 'Total' columns, one is given the proportions of encoding of each event component relative to the other event components in single-gesture versus multi-gesture accompaniments, without regard to the category of gesture in which the component was expressed. Table 11 summarizes these data by language, utterance category, and single-gesture accompaniments versus multi-gesture accompaniments. With this number of factors, Table 11 is large and complex, and the number of instances in some subsets of the cells is too small to inform general observations. When they are broken out in this way, however, certain patterns seem evident in the Study Two data. Six of these patterns are all that will be mentioned here.

One set of inferences that seems legitimate based on the summary in Table 11,

given the number of observations involved, concerns comparisons among the two-place verb utterance structures that were sampled for Study Two, in terms of event component and case role encoding. These utterance structures are the English SVO and the Mandarin BA-construction and SVO utterance structures. Considering the column labelled ‘Grand Total,’ we see that, overall, Mandarin BA-constructions and English SVO utterances are similar in terms of their respective proportions of gestural ACTION-encoding when the contributions of pantomimic enactments are figured in. In Mandarin BA-constructions, 40% of all gestural event component references are to ACTION; in English SVO utterances the proportion is 37%. Compared to these two target structures, Mandarin SVO utterances show a lower proportion overall of ACTION-encoding, 31%. When we consider the proportions in Mandarin BA-constructions and English SVO utterances in relation to ACTION-encoding in single- versus multi-gesture utterance accompaniments, however, we see that these two utterance categories differ somewhat. In Mandarin BA-constructions, a higher proportion of the gestural event component references in single-gesture accompaniments are to ACTION than is true in the case of English SVO. Therefore, in spite of the overall equivalence between the two languages in these utterance contexts of proportions of ACTION-encoding, it seems that ACTION-encoding is more prevalent in single-gesture-accompanied Mandarin BA utterances, while being roughly proportionally equal in single-gesture versus multi-gesture accompaniments to English SVO utterances. This is evidence from a slightly different angle for the emphasis already detected in Mandarin BA-construction contexts for gestural highlighting of

ACTION. If a Mandarin speaker produces just one gesture in a BA-construction context, there is a higher likelihood that it will be ACTION-encoding than is true of a single-gesture accompaniment to an English SVO utterance. In this comparison, although the numbers are small, Mandarin SVO utterances appear to line up with English SVOs in terms of ACTION-encoding; that is, 35% of single-gesture accompaniments to Mandarin SVO utterances encode ACTION, compared to 38% of English SVOs, compared to 53% of Mandarin BA-constructions.

The second pattern is that Mandarin SV utterances are accompanied by the highest proportions of ACTION-encoding gestures in every division of the data. This utterance category has the highest proportions as well of Action Only Isolating gestures in every division of the data, if the high proportion of such gestures in single-gesture accompaniments to Mandarin TOPIC-COMMENT utterances is excluded, on the grounds that that proportion is based upon only three gestures.

A third pattern is evident in the gestural references to PATIENT across the two-place verb utterance categories in the two languages. Considering first the proportions listed in the Grand Total column for Mandarin SVO and BA utterances, and English SVO utterances, we see that English SVO utterances are accompanied by the lowest proportion overall of gestural PATIENT references. Recall that SVO utterances are abundant in the English narration data sampled for Study Two. This means that this level of gestural PATIENT-encoding is probably a prominent pattern overall in the English data. Of the gestures occurring in all single-gesture and multi-gesture accompaniments to the forty English SVO utterances, including pantomimic enactments, 28% are PATIENT-encoding in

some way. Staying with the Grand Total column and considering the proportions listed for the twenty Mandarin BA utterances, we see that 37% of these are PATIENT-encoding in some way; thus somewhat more than English SVO. The twenty Mandarin SVO utterances are accompanied by the highest proportion of all of PATIENT-encoding gestures, 44%. Recall, however, that SVO utterances are relatively rare in the Mandarin narration data sampled for Study Two. This means that this high proportion of PATIENT-encoding in gestural accompaniments of twenty utterances is probably not representative of the pattern overall in the Mandarin narrative.

When we consider just the single-gesture accompaniments, Mandarin SVO utterances are again distinguished as having the highest proportion of PATIENT-encoding gestures, 43%, compared to BA-constructions at 24%, and English SVO utterances at 28%. In multi-gesture accompaniments, Mandarin SVO and BA utterances have roughly equal proportions of PATIENT-encoding gestures, 45% and 42%, respectively, while English multi-gesture accompaniments have this type of gestural encoding in about the same proportion as in single-gesture accompaniments in the SVO utterance category, 30%.

Note also that half of the total gestural PATIENT references in the single-gesture accompaniments to Mandarin SVO utterances are packaged as Isolating gestures, whereas all the PATIENT references in the (small number of) single-gesture accompaniments to BA-constructions are embedded in pantomimic enactments, the latter being less specific forms of reference. In terms of specific gestural references to PATIENT in multi-gesture utterance accompaniments, these two Mandarin utterance structures appear similar. Both

have similar proportions of Isolating PATIENT-encoding Isolating gestures, and PATIENT-encoding pantomimic enactments, so it is in the single-gesture accompaniments that they differ as to PATIENT reference. In single-gesture accompaniments to English SVO utterances, the much higher proportion of PATIENT references are, similarly to Mandarin BA contexts, embedded in pantomimic enactments. Thus in regard to PATIENT-encoding, when all the factors that were coded for Study Two are considered, the gestural accompaniments of Mandarin SVO utterances and BA-constructions do appear to differ, and not in the expected direction; that is, the PATIENT role seems differentially gesturally highlighted in Mandarin SVO contexts compared to BA contexts. The overall pattern of PATIENT references in the gestural accompaniments to English SVO utterances does not really resemble the pattern of either of the Mandarin structures.

Staying with the two-place verb argument structures and considering gestural references to AGENT/FIGURE, a fourth pattern is evident in the data summarized in Table 11. Only English SVO utterances are accompanied by an appreciable proportion of explicit gestural references to the AGENT/FIGURE case role. Among the gestural accompaniments to the forty combined Mandarin SVO and BA utterances, there is only one gestural reference to AGENT/FIGURE that is not embedded in a pantomimic enactment. This is an Element Only Isolating gesture that is part of a multi-gesture accompaniment to a SVO utterance. Adding together the total AGENT/FIGURE references listed in Table 11 for the Mandarin SVO and BA utterances, this is one gestural reference out of a total of twenty-five. In contrast, specific AGENT/FIGURE references in single- and multi-gesture accompaniments to English SVO utterances, whether Element Only

Isolating gestures or as part of observer viewpoint Incorporating gestures, comprise 25% of the total gestural references to AGENT/FIGURE in that utterance context. Comparing this pattern with the one discussed previously, and keeping in mind that the numbers are rather small, it appears that AGENT/FIGURE is somewhat preferentially encoded over PATIENT in English SVO utterance contexts. The reverse appears to be true of the Mandarin SVO and BA structures.

Considering AGENT/FIGURE-encoding across all utterance contexts in both languages, a fifth pattern is evident. Among the *single-gesture* accompaniments to Mandarin utterances of all categories, there is but one specific AGENT/FIGURE reference. This is as part of an Incorporating gesture that accompanies a Mandarin SV utterance. Among multi-gesture accompaniments to Mandarin utterances, as has already been mentioned, there were no specific AGENT/FIGURE references accompanying BA-constructions, and again only one specific AGENT/FIGURE reference accompanying an SVO utterance. Thus in these Mandarin categories, the only gestural references to AGENT/FIGURE are embedded in pantomimic enactments.

The truly appreciable numbers of specific AGENT/FIGURE references in Mandarin are those in multi-gesture accompaniments to TOPIC-COMMENT utterances. In the Mandarin data, specific AGENT/FIGURE references do occur as part of the multi-gesture accompaniments to both SV and TOPIC-COMMENT utterances. The numbers are relatively small, however, in SV contexts — a total of only five specific gestural references in multi-gesture accompaniments to SV utterances. In contrast, in the TOPIC-COMMENT

utterance category, when we add the totals of Isolating and Incorporating AGENT/FIGURE references shown in Table 11, we see that eleven of eighteen total gestural references, or 61% of gestural references in multi-gesture accompaniments to Mandarin TOPIC-COMMENT utterances are to the AGENT/FIGURE role. Specific AGENT/FIGURE reference in multi-gesture utterance accompaniments (the much larger proportion of TOPIC-COMMENT utterance accompaniments) is what most distinguishes this utterance category from the other Mandarin categories examined for Study Two.

In contrast to the general picture regarding specific AGENT/FIGURE-encoding in the Mandarin data, both English utterance structures, SV and SVO, associate with appreciable proportions of specific references to this case role in single-gesture utterance accompaniments, and, it appears, in multi-gesture accompaniments as well, although the numbers are small in the latter category. English SV contexts stand out among all utterance categories in both languages as associating with a particularly high proportion of specific AGENT/FIGURE references, the majority of these as part of observer viewpoint Incorporating gestures. This latter finding fits with the English speakers' overall tendency to produce Incorporating gestures. Calculating from the totals in Table 11, the proportion of total gestural references to the AGENT/FIGURE case role that occur as Isolating gestures or in observer viewpoint Incorporating gestures in English SV utterance contexts is 72%. In contrast, the larger proportion, 75%, of AGENT/FIGURE references in gestures accompanying English SVO utterances are embedded in the pantomimic enactment gestures which that utterance context favors.

The final pattern evident in the Table 11 summary that will be mentioned is one that concerns the issue of overall gestural encoding of propositional content or event

components, and whether this differs between the two speaker groups. In the description of the findings concerning non-pantomimic gestures presented earlier in Table 10, it was noted that, despite the fact that the Mandarin speakers produce more gestures overall than the English speakers, it is unlikely that the Mandarin speakers encode more propositional or event content in their gestures overall than do the English speakers. The proportions listed in the ‘Total’ columns of Table 11 include the contributions of pantomimic enactment gestures. It is therefore possible to tally the instances of gestural reference to event components by language. Calculating from the raw number values in the table, the total for the Mandarin speakers is 202 references; for English speakers, 193. Thus despite differences in the proportions in which particular types of gestures are used by each speaker group, and variations across utterance contexts in terms of which components are singled out for explicit gestural reference, overall the two groups appear to gesturally encode comparable amounts of content.

Summary of Findings on Rate of Gesture and Gesture Semantic Content

The gestural accompaniments of this sample of 160 utterances appear to vary by language and within language by target utterance structure. Table 12 lists in summary form the findings that were presented in narrative form and in each of the tables in the preceding presentation of the results of the planned analyses.

Across all utterance categories, Mandarin speakers show a higher ratio of gestures-to-utterance. Within Mandarin, TOPIC-COMMENT utterances are accompanied by the highest number of gestures with more than two gestures per utterance on average, followed by the object-fronting BA-constructions and the SV(o) utterance structures, in that order. In the English sample, SV and SVO utterances are accompanied by approximately the same number of gestures on average.

Table 12. (Page 1 of 2) Main points from the results of the planned analyses of Study Two, Mandarin versus English.

Speech sampling:

- MA Low incidence of TOPIC-COMMENT and SVO utterance structures suitable for sampling; high incidence of BA-constructions and 'TOPIC chains.'
- EN Abundant SV and SVO utterances suitable for sampling.

Gestures-to-utterance ratio:

- MA The higher ratio in every target utterance category, the highest ratio being in TOPIC-COMMENT utterance contexts, followed by BA-constructions, then SV(O).
- EN Lower gestures-to-utterance ratios in English SV(O) than in Mandarin SV(O).

Isolating versus Incorporating gestures:

- MA Two thirds of gestures are Isolating.
- EN Two thirds of gestures are Incorporating.

Element Only versus Action Only gestures:

- MA Element Only Isolating gestures are high frequency.
- EN Element Only Isolating gestures are low frequency.

Semantic content by utterance category:

- MA
- In TOPIC-COMMENT utterances, emphasis on Element Only Isolating gestures the bulk of these being AGENT/FIGURE-encoding. In almost all cases, the utterance is further accompanied by additional gesture(s) that encode other event components.
 - In SVO utterance contexts, almost no specific AGENT/FIGURE-encoding; none in BA contexts. Specific AGENT/FIGURE-encoding is non-negligible in SV contexts, but the frequency of Isolating gestural references to this role in SV contexts is low compared to that in TOPIC-COMMENT contexts.
 - In BA-construction contexts, roughly equal proportions of Element Only and Action Only Isolating gestures. When Incorporating gestures are included, an overall emphasis on gestural ACTION-encoding in this utterance context emerges.
 - In SV utterance contexts, the emphasis is on Action Only Isolating gestures.
-

Table 12. (Page 2 of 2) Main points from the results of the planned analyses of Study Two, Mandarin versus English.

Semantic content by utterance category (continued):

- MA • SVO utterance contexts, of the Mandarin utterance categories, most associate with PATIENT-encoding in gesture. In BA-construction contexts, the somewhat greater emphasis is on ACTION-encoding.
- EN • In both SV and SVO utterance contexts, there are about twice as many Action Only as Element Only Isolating gestures among the overall small number of Isolating gestures.
- In SVO contexts, pantomimic enactments are high frequency; in SV contexts, low frequency.
- In SVO contexts, specific gestural references to PATIENT are uncommon; to AGENT/FIGURE somewhat less uncommon. The high frequency of enactments in this utterance category, however, elevates the proportions in both of these semantic categories substantially.
- In SV contexts, specific gestural references to ACTION are common. Specific gestural references to AGENT/FIGURE are rather frequent — as frequent overall as in Mandarin TOPIC-COMMENT utterance contexts.

Single- versus multi-gesture accompaniments to utterances:

- MA • About half of utterances are accompanied by more than one gesture.
- There is an overall preference for Isolating gestures in both single- and multi-gesture accompaniments to utterances, with a tendency to pair ACTION- and ELEMENT-encoding gestures in multi-gesture accompaniments.
- In SVO contexts, gestural PATIENT-encoding is most marked in single-gesture accompaniments.
- Gestural AGENT/FIGURE-encoding is almost non-existent in single-gesture accompaniments, but is a common feature of multi-gesture accompaniments to TOPIC-COMMENT utterances, and non-negligible in multi-gesture accompaniments to SV utterances.
- EN • Only 15% of utterances are accompanied by more than one gesture.
- Perhaps a tendency, in the small number of multi-gesture accompaniments to pair Isolating gestures with Incorporating ones.

Overall gestural encoding of propositional or event content:

Comparable in the two languages, when the contributions of pantomimic enactments are included, despite a higher ratio of gestures-to-utterance in the Mandarin speakers.

The findings on the kinds of semantic content encoded by these gestural accompaniments might best be summarized by listing the most prominent tendencies within each language and target utterance category. In the Mandarin sample of gestures, the dominant tendency was to produce Isolating gestures; in the English sample, Incorporating gestures.

Among Mandarin TOPIC-COMMENT utterances, Element Only Isolating gestures that encode the AGENT/FIGURE role predominate. These almost always occur together with gestures of other types and contents. The other non-SV(O) utterance category in Mandarin, the BA-constructions, has less of a tendency to associate with Isolating gestures. A significant proportion of the Isolating gestures that do accompany BA-constructions are ACTION-encoding. The high proportion of Incorporating gestures that occurs in this utterance environment also encode ACTION. The expectation that this utterance structure would associate with a high proportion of PATIENT-encoding gestures was not supported. It is the Mandarin SVO utterance structure in fact that, of the four structures examined, associates with the highest proportion PATIENT-encoding Isolating gestures. Mandarin SV utterances stand out among the Mandarin target utterance types as associating with the highest proportion of Action Only Isolating gestures, and almost no pantomimic enactments.

The gestural accompaniments to English SVO utterances are distinguished by their very high proportion of pantomimic enactments. This high proportion, because of the densely-encoding character of enactments, elevates the proportions of encoding in all case role and event component categories. English SVO utterances are associated with

less specific PATIENT-marking than either of the Mandarin two-place verb structures. Both English utterance categories associate with appreciable proportions of gestural AGENT/FIGURE reference, particularly SV utterances, the latter standing out among all utterance categories and languages as associating with a high proportion of this kind of reference, especially in observer viewpoint incorporating gestures.

Appendix E lists all the single- and multi-gesture accompaniments of utterances by target utterance structure in order of frequency of occurrence.

Other findings

In the preceding sections the results of the planned analyses were presented. These analyses covered the overall amount of gesture, and the semantic content of gestures that accompany different categories of utterances. In the course of extracting the samples from the narrations and making the planned comparisons, other features of speech-gesture pairings that may distinguish the two speaker groups and that are conceivably related to the typological distinction of topic prominence versus subject prominence were observed. These final sections of the Results report these other features.

Although the data analyzed for Study Two consist of extracted sentential utterances that were examined in isolation from the contexts in which they occurred, it was observed while scanning the data for these samples that there may be differences between the Mandarin and English speakers in how their gestures synchronize with speech and in how their speech-gesture productions connect in sequence, as follows: (i) In Mandarin there may be pressure within some speech-gesture productions to produce gesture toward the front end of the ‘production envelope.’ This gestural tendency could be interpreted as in keeping with the establishment of TOPICs in utterance-initial position.

Several observations concerning specific patterns of speech-gesture synchrony are reported; (ii) Mandarin speakers appear to make more use than do the English speakers of repeated and re-configured speech and gesture material as they describe connected sequences of events in narrative, possibly as a means to maintain topic reference.

Speech-gesture synchrony

Speech-gesture timing and semantic content in Mandarin non-SV(O) utterances.

Table 13 shows the speech-gesture composition of the twenty TOPIC phrases of the twenty TOPIC-COMMENT utterances in this sample. A high proportion of these (fifteen utterances, or 75%), are accompanied by representational gestures of some sort. In the majority of these cases, the accompanying gesture (eleven utterances, or 73% of the fifteen) is closely related in semantic content to the spoken TOPIC statement; that is, the two modalities seem largely redundant in meaning. Roughly half of these redundant

Table 13. The semantic contents of the TOPIC statements and their accompanying gestures in Mandarin TOPIC-COMMENT utterances.

Content of TOPIC statement			Content of TOPIC-accompanying gesture
.25	5	AGENT or FIGURE	(no gesture accompanying TOPIC)
.20	4	AGENT or FIGURE	AGENT or FIGURE (redundant with TOPIC)
.15	3	ACTION	ACTION (redundant with TOPIC)
.10	2	GROUND	GROUND (redundant with TOPIC)
.10	2	PATIENT	PATIENT (redundant with TOPIC)
.15	3	AGENT or FIGURE	Co-referring with the following COMMENT
	1	AGENT	pantomimic enactment

20 TOPIC-COMMENT utterances in total.

gestures are iconics, thus depictive. The other half of the gestures that are redundant with spoken TOPIC statements are deictics or localizers that index the spatially-located referent of some spoken nominal expression; for instance, a location in gesture space already set up in the foregoing narration. Three of the TOPICS in this sample of twenty, or 15%, are accompanied by gestures that are co-referential with some component of the immediately following COMMENT.

The following is an example of a TOPIC-accompanying gesture whose semantic content is redundant with that of the spoken TOPIC with which it synchronizes. In this case the referent, “men”/“door,” may be considered a GROUND component of the event representation.

- (46) [ta-men-de men a / sh][ang-mian you yi-ge tian-chuang /
 他們的門 呵 / 上面 有一個天窗 /
 their door TOP / top-side have one-CL sky-window /
as for their door / above it there is a transom /

nei-ge nu zhu-ren] jiu cong tian-chuang shang-mian # jiu chu-lai
 那個女主人 就 從天窗上面 # 就出來
 that-CL woman owner just from sky-window top-side # just exit-come
the (bird's) lady owner then appears/comes out up there in the transom

The gesture that accompanies the spoken topic statement is depictive of the same semantic content as is encoded in the speech, so the two are essentially redundant. The gesture depicts the two sides of the door. The gesture that follows the TOPIC-accompanying gesture times with another spoken GROUND component, “tian-chuang”/“transom,” and is a representation of the PATH of moving through this transom. The second gesture thus precedes its most closely co-referential portion of speech,

“chu-lai”/“come out,” by a fair margin.

Example (46) illustrates a phenomenon encountered frequently in the data where a TOPIC and COMMENT each are accompanied by a gesture, often one that encodes a meaning similar to the unit of speech with which it co-occurs. Table 13 shows that spoken TOPICS frequently co-occur with a gesture of closely-related meaning. Recall from the results of the planned analyses that gestural accompaniments to BA-constructions are frequently ACTION-encoding. When we look at the precise synchrony of these ACTION gestures with constituents in the co-occurring BA-construction, we see examples of how the semantic content of gesture need not be ‘redundant’ with the content of the directly co-occurring speech, but may instead express a complementary meaning. Of the the BA-constructions, a higher proportion than in the case of the TOPIC-COMMENT utterances, 90%, had a speech-gesture timing pattern such that the fronted verb objects co-occurred with a representational gesture of some sort. The analysis of those gestures showed that in a slight majority of cases, 55%, this gestural accompaniment encoded the ACTION ‘suffered’ by the fronted verb object. Somewhat fewer cases, 44%, were Element-marked gestures semantically redundant with the FIGURE/PATIENT referred to by the fronted verb object.

In example (47), below, the speaker uses a BA-construction to convey the cartoon event of Sylvester capturing a monkey:

- (47) [ta jiu ba **nei hou**-zi gei] zhua-lai le
 他就 把那猴子 給抓來了
 he just OBJ that monkey give catch-come PRF
 then he grabs the monkey

The gesture in this case is an enactment of grabbing the monkey with one hand. This gesture times not with the verb expressive of that ACTION, “zhua”/“catch,” but rather with the PATIENT of the ACTION. This was a very common speech-gesture timing pattern in the Mandarin data, and it did not always involve comprehensively-encoding pantomimic gestures. That is, often fronted verb objects were accompanied by Action Only Isolating gestures as well.

Thus the typical gestural accompaniments of the two categories of non-SVO utterance examined in Study two differ somewhat. The gestures that accompany spoken TOPICs are often semantically redundant with the content of the topic statement. Those that accompany BA-constructions are often co-expressive in the sense of having complementary semantic content.

The position of gesture within the production envelope. The semantic content and timing relationships between speech and gesture just mentioned are specific to the Mandarin TOPIC-COMMENT utterances and BA-constructions examined for this study. More general differences as well between the two languages emerged during the sampling process in terms of the typical timing relationship between speech and co-expressive gesture within the speech production unit. In brief, it was observed that, whereas in English it is typical for gestures that represent an action expressed in a sentential predicate to synchronize with a co-referential and prosodically emphasized spoken unit within that predicate, in Mandarin such gestures may often precede their most closely co-referential unit of speech by a fair margin. In Mandarin there appears to be a pressure within the utterance production envelope for the focus of speech and gesture to converge at a position toward the beginning of the envelope. What Mandarin

speakers often produce, and the English speakers apparently seldom produce, is a gesture that conveys a particular semantic content, timed to appear a bit earlier within a production unit than the unit of speech that conveys corresponding semantic content. Such a gesture may time with an utterance-initial topicalized unit of speech, or with a unit that is not linguistically marked as TOPIC, but with which the gesture shares a complementary semantic relationship. The unit of speech with which the gesture could be considered to be more ‘redundant’ in semantic content then often comes later in the production envelope, sometimes spoken in a prosodically de-emphasized form. In the Mandarin speakers, for instance, we sometimes see COMMENT-related gestures timing with topic statements in speech. The example given earlier from a Mandarin speaker is one such instance. In it the speaker produces a topicalized phrase about the cat, “mao,” and accompanies it with a gesture about the action of the cat and a bowling ball rolling

- (48) [mao ne] zai da jie-shang <ehn> gun-dong
 貓 呢 在大街上 <ehn> 滾動
 cat TOP on big street-surface <ehn> roll-move
 the cat <um> rolls around on the street

around together. The units of speech that express the event component of MOTION itself, “gun-dong”/“roll around,” come along only some moments later, prosodically de-emphasized and following a pause. This kind of speech-gesture synchrony is different from what is typical in the English speakers, who show a preference for timing predicate-related gestures rather tightly with the predicates of their spoken sentences.

Further, in a non-negligible number of instances in the Mandarin narrations, these ‘fronted’ gestures co-occur with vocal onomatopoetic forms. These mimetics often seem to express something like MANNER of movement; sometimes, with pitch changes, PATHs. These created forms thus appear to encode event components in a way that is co-

expressive with the content of the co-occurring gestures. This kind of speech-gesture production was observed only once in the seven English narrations examined for Study Two. Thus it appears that Mandarin-speakers' gestures can precede co-referring speech by a margin that is wide in comparison to the relatively tight spoken PREDICATE-gesture synchrony that is common in English-speakers; yet the Mandarin combinations have their own kind of semantic coherence.

Repetition in gesture. The particular direction taken by the analysis of Mandarin speakers' higher gestures-to-utterance ratio that was presented at the beginning of the Results was a consequence of a certain stand-out feature of their gestural performance observed in the exploratory pilot study out of which Study Two grew. What stood out in the Mandarin data considered for that study was the high proportion of Element Only Isolating gestures. Therefore it seemed natural to pursue an analysis guided by the notion that what may account for the Mandarin speakers' higher gestures-to-utterance ratio is a tendency on their part to be gestural 'splitters' as opposed to 'lumpers.' That is, it was thought that these speakers may isolate event components in gestural expression to a greater extent than the English speakers. Then, if they maintain comparable levels of overall expression of event components in gesture, the consequence would have to be a higher amount of gesture. The data summarized so far suggest this conclusion.

However, another feature of the gestural performance of Mandarin speakers emerged, one not visible in the analyses as they were framed here, in the course of considering sequences of their narrations in order to extract the sets of target utterance structures. This is a tendency to repeat gestures; that is, to produce gestures that repeat all or some part of a previous gestural performance.

Repeat gestures are of many types. It appears to be somewhat rare for a gesture to be a wholly unmodified reiteration of a previous one. Typically, gestures are repeated in reduced or elaborated forms. A reduced gesture may for example have a truncated stroke phase, be expressive of fewer event components, or be executed with a modified, less ‘tensed’ hand. An elaborated gesture could simply be bigger, encompassing more of the speaker’s gesture space, or it could be the same size as a previous gesture, but with the addition of some features. For instance, one gesture could express the PATH of a motion event, and a repeat of it could express this same PATH, but with some MANNER-marking added.

The possibilities for such variation among gesture repetitions appear to be many and these are as yet unexplored. What stood out in the narrations examined for Study Two, however, was an apparently higher incidence of repeat gestures of various kinds in Mandarin speech-gesture productions, relative to what was encountered in the English narrations. Because the coding performed for Study Two did not categorize gestures with reference to this dimension, the impact of the apparent high incidence of repeat gestures in the Mandarin narrations on the comparison of the Mandarin and English speakers’ gestures-to-utterance ratios cannot be assessed. It is reasonable to suppose, however, that if Mandarin speakers repeat gestures more often than English speakers, and do so within utterance frames such as the ones sampled for this study, that this could account for all or some part of their higher gestures-to-utterance ratio.

Several features of Mandarin speakers’ repeat gesture performances were noted in the course of sampling the data included in the Study Two data set. For example, the Mandarin speakers frequently use repeated depictive gestures as well as non-deictive indexing gestures (deictics and beats) as a means to index information that has already been brought out, but that will be focal in the discourse yet to come. In example (49) an

Discussion

In spontaneous Mandarin narration data of the sort examined here, object-fronted and TOPIC chained structures are frequently-chosen means of expression. When a Mandarin speaker narrates an event that typically would be described using a two-place verb in English, it is unusual for her to produce a SVO utterance structure; choosing instead one of the other, more favored constructions. There are many more BA-constructions in these narrations than were sampled for this study. In contrast, there are not many more Mandarin SVO utterances in the data set than the sample of twenty extracted; this in spite of the fact that eleven approximately eight-minute Mandarin narrations were scanned for the target utterance structures. It is thus clear that the topicalizing, or focusing, structures of Mandarin examined in this study are the core, high-frequency structures preferred by Mandarin speakers as they construct their narrations.

It seems further clear on the basis of the gestural evidence, especially so from the comparisons of proportions of gestures that selectively depict elements such as AGENT/FIGURE, PATIENT, and GROUND in English and Mandarin, that in the process of thinking-for-speaking, the two speaker groups differ in terms of how they divide up the domain of event components for expression. For the most part, English speakers do *not* divide it up. This speaker group generates a preponderance of incorporating gestures, the favored type being the most extreme kind of incorporation, pantomimic enactment. This gesture production pattern of event component conflation contrasts with the pattern in the Mandarin speakers. The Mandarin speakers make more isolating gestural references to the Element and Action components of events. Most likely in part because of this these speakers generate an overall higher gestural output. It was hypothesized that high

proportions of Element Only gestures would be revealed to be the gestural realization of those processes in spoken Mandarin that separate elements and shift them around into positions of TOPIC or FOCUS. When the gestural accompaniments of spoken topic statements are considered, this hypothesis seems supported. Topic statements are often accompanied by isolating gestural references to the content they express.

In contrast, the PATIENTs of actions are not preferentially highlighted in the gestural accompaniments of BA-constructions. Rather, the utterance structure that according to standard theory is the base, untransformed structure of the language, SVO, is the one that appears to associate with the largest proportion of specific gestural PATIENT reference. When we pair this unexpected finding with the native speaker preference for *SOV* utterances mentioned in the description of BA-constructions in the first part of the chapter, it becomes reasonable to consider that it is the *SVO* structure in fact that should be considered a linguistically-marked, PATIENT-focusing construction. On this view, particular salience in conception of the PATIENT role could be part of the conditions favoring production of a marked, SVO spoken form. This is consistent with the link between word order and the definiteness/indefiniteness of reference distinction in Mandarin. Recall that the BA-marked NPs are always definite, thus are vehicles for information that has already been mentioned in discourse or is shared between the conversation partners. Thus it is not predicted that these PATIENT NPs would usually be the locus of the psychological predicate, or GP; rather, such a NP would have been the new information in some utterance preceding the one in which it appears as a BA-marked constituent. That such ‘given’ PATIENT information in speech in BA-construction

contexts often pairs with ‘new’ ACTION information in gesture is consistent with the pattern of speech-gesture synchrony in Mandarin that was outlined in the Results, and that appears to distinguish Mandarin from English.

We may characterize the general finding by saying that utterance constituents do appear to be conceptually more dissociable in Mandarin. In the process of thinking-for-speaking in Mandarin, these constituents come to occupy structural slots that individually generate speech+gesture productions. In terms of McNeill’s (1992) theory of speech and gesture production, they are individually the locuses of GPs. This is true not only of the non-SVO utterance structures examined here, but of the two structures that overtly mirror English word order patterns as well.

The set of facts about Mandarin utterance production that have pushed linguistic theorizing about the language away from standard assumptions in the tradition of Indo-European linguistic research, thus appear to have gestural correlates. The gestures make clearer the essential nature of utterance structuring dynamics in Mandarin than do the surface ordering of sentential constituents. To see why this is true, consider the Mandarin SVO utterances sampled for this study. According to standard theory, this should be considered to be the base, untransformed utterance structure of the language. Yet SVO utterances in Mandarin are accompanied by more gestures on average than English SVO utterances and by gestures of qualitatively different kinds. Such gesture evidence suggests that the constituents making up such Mandarin utterances may be dissociable in conception even when their sequential patterning mirrors that of cohesive English SVO sentences. Despite the surface similarity between these utterance structures in the two languages, it thus appears not to be the case that utterance planning and cohesion are in fact totally comparable. Thinking-for-speaking apparently differs such that elements of

Mandarin propositions that are overtly organized similarly to cohesive SUBJECT-PREDICATE English sentences, can arise from two ‘idea units’ rather than one. This seems in line with the implications of the claim in the Li and Thompson (1981) analysis, that what makes Mandarin a topic-prominent language is that every utterance, regardless of the ordering of its constituents, or the presence or absence of TOPIC-marking grammatical particles, should be regarded as a topicalized utterance, including ‘SV’ or ‘SVO’ utterances. If whatever occupies utterance-initial position may be considered to be the utterance TOPIC, and whatever follows this the COMMENT, and, further, we see that such constituents are often each the locus of a GP, then a finding of higher ratios of gestures-to-utterance across all utterance categories in Mandarin than in English makes sense. The implication of this may be that it is generally inappropriate to construe the linkage in any Mandarin utterance between an utterance-initial constituent and what follows it as being a cohesive SUBJECT-PREDICATE relationship.

The issue here at root concerns what structures or constituents in each language show themselves to be the sources of units of expressive production in the language. Such speech-gesture ‘pulses’, or production envelopes, reflect the total extent of unpacking of single GPs in speech and gesture production. From the data presented above it would appear that spoken topic statements and fronted objects can be part of the organizing impulse of a production envelope in Mandarin; that is, these syntactic constituents may comprise the speech code portion of a Mandarin GP.

That there can be depictive gestural representations co-referential with topic statements in Mandarin brings up theoretical issues in both GP theory and in the standard theory of topicalization in Mandarin, having to do with the notion of ‘given information.’

According to McNeill, the frequent role of deictics and related indexical gestures is to represent already established discourse referents by pointing to locations in gesture space with which they have been identified. Iconically depictive gestures in contrast are thought to line up largely with non-presupposed information as it is introduced into the discourse. The claim of standard topicalization theory is that topic statements are vehicles for presenting already-established information in discourse. Putting these pieces of the two theories together, we could have expected that iconically depictive gestures would not typically accompany topic statements. The finding that this does sometimes occur either means that the functions of gestural depiction must be reconsidered within GP theory, or it constitutes gestural evidence for the position held by some (Huang, 1994) that the ‘given’/‘new’ information distinction is not relevant to a characterization of topicalization in Mandarin.

The design of Study Two, based as it is on a sample of utterances that were extracted and considered in isolation from their contexts, does not permit determining whether the ‘pause particle’-marked topic statements in the sample differ systematically within category in terms of the information value of their content, in a way that predicts whether their gestural accompaniment will depict (iconics) as opposed to index (deictics). Such issues require a method of investigation by which changes in information value of the constituents of utterances connected in sequence are tracked, to see if these correlate with changes in gestural mode of representation.

Given the greater overall gesture output in the Mandarin speakers, neither is it *only* the TOPICs and fronted objects that are pairing with gesture in the course of producing single ‘sentential’ utterances. In spite of comparable expression of narrative content, the gestural data make clear that propositional content more often expressed in a single production envelope in English is often broken out across two or more production

envelopes in Mandarin, with particular components isolated for separate gestural realization. These facts of gesture production mark the utterance structuring dynamics of Mandarin as fundamentally different from those of English. The SV(O) structures of the English speakers in this sample were accompanied more often by incorporating, fused, gestural representations of the same narrative events that frequently associated with two or more gestures in the Mandarin speakers.

Note that it is likely not true that Mandarin speakers, in their higher ratio of gestures-to-utterance, gesturally encode a larger proportion of the propositional content conveyed in speech than do the English speakers. The Incorporating gestures, particularly enactments, favored by the English speakers are densely-encoding gesture forms. When a speaker produces an Incorporating gesture, she doubles or triples the encoding of event components, over the single components encoded in Isolating gestures.

Summary and Conclusions

The goal of Study Two was not to undermine or reframe existing and widely-accepted linguistic analyses of the topic prominence feature that distinguishes Mandarin from English, but rather to take an exploratory sample of spontaneous natural language data in the service of furthering the development of a psychological theory of linguistic communicative production; one centered on the hypothetical GP unit. Since the standard typological analysis is located at the level of sentential syntax, the attempt was to carry out Study Two at that level. The standard analysis provided the templates with reference to which the Mandarin data were sampled. With propositional content held roughly comparable, any systematic co-variation between utterance structure and the patterning of gesture in relation to speech would be taken as evidence of different patterns of thinking-for-speaking.

Gestural differences were in fact observed. It was further found that these gestural differences are present even when linguistic structures in the two languages are compared whose ordering of constituents is overtly the same. Yet the question of how central to Mandarin utterance formation dynamics the target utterance structures are, renders interpretation of the gesture findings of Study Two equivocal, or at least weakens the evidence by bringing into question the representativeness of portions of the data. Useable TOPIC-COMMENT and SVO utterances were difficult to locate in the narration data. At this juncture, tentative statements only seem warranted along the following lines: (i) speakers of Mandarin appear to gesture at a higher rate than do speakers of English, and (ii) the gestures of the Mandarin speakers separate and isolate the components of events to a greater extent than the gestures of English speakers. The additional observations listed in the Results suggest further that, (iii) speech-gesture synchrony may differ in the two languages in systematic ways, and (iv) some portion of the higher ratio of gestures-to-utterance in Mandarin may be accounted for by gesture repetition. Further research may establish the latter two findings as being generally more informative in regard to the essential utterance structuring dynamics of the language than those based on the set of extracted sentences considered in isolation from their contexts.

Among the Isolating gestures produced by the Mandarin speakers in TOPIC-COMMENT utterance contexts, the proportion of Element Only gestures is greater than that of Action Only gestures. These results are generally consistent with the hypothesis that the topic prominence feature of Mandarin may increase the saliency in conception of the referents indexed by topic statements. Although there was an appreciable amount of gestural PATIENT reference in the high-frequency BA-construction contexts, these contexts were actually distinguished by their tendency to associate with gestural ACTION

references. Therefore the hypothesis that this utterance structure in Mandarin is the expression of a singular focus in conception on PATIENTs is not supported; rather, the frequent association of gestured ACTION with spoken PATIENT in these cases suggests a dual, fused focus.

Specific PATIENT focus appears to be a factor instead in Mandarin SVO contexts, a finding that, together with the native speaker sense that a SVO utterance is a marked structure, calls for rethinking assumptions concerning what constitutes the linguistically ‘marked’ form, in the case of an utterance organized around a two-place verb. What is generally taken to be the base, unmarked form of the language, the SVO structure, may actually be a relatively low-frequency, marked, focusing construction.

That in this sample Mandarin speakers produced more Element Only than Action Only gestures is most likely a function of two factors: (i) the majority of sentential TOPICs in the sample were nominals; the fronted objects in BA-constructions are of course always nominals; (ii) speech-accompanying gestures are frequently highly redundant in expressive content with the units of speech they accompany. In other words, since the gestural accompaniment of a spoken nominal is likely to encode features of the nominal referent, and most of the TOPICs were nominals, as were the fronted objects, a higher proportion of Element Only Isolating gestures is predicted. It is possible that, had more of the sentential TOPICs in the sample been verbals, a larger proportion of the Isolating gestures produced by the Mandarin speakers may have been of the Action Only type.

The gesture evidence thus suggests that Mandarin and English speakers’ communicative productions spring from organizationally dissimilar idea units, in a manner that is not bound by the particular surface conjunctions among ‘sentence’ constituents.

The status of the results of Study Two. It must be asked just how constrained the results of Study Two may have been by the design of the sampling procedure used, making them perhaps a less than completely reliable reflection of differences between Mandarin and English utterance structuring dynamics. For instance, it may not be globally true that Mandarin speakers overall produce more gestures when covering the same narrative content as English speakers, nor that Mandarin speakers more often ‘split’ where English speakers ‘lump’ event content in gesture. The sampling of utterance structures that went into this study was selective in the sense of the need to isolate clean exemplars of the target SV(O) and non-SV(O) utterance structures in the two languages. These needed to be cross-linguistically comparable in propositional content and excized from free-flowing narrations. As was alluded to in the earlier section on Mandarin syntax, what appears to be the most common topicalizing process in these narrations, TOPIC chaining, was of necessity not sampled at all. So there is a chance that the most robust utterance-structuring dynamics of the language may not be reflected in the data sets used, as a result of application of narrow sampling criteria. A conclusion on this issue awaits further analysis of the chains encountered in these narrations. Such an analysis will place this type of utterance in perspective with the more often discussed TOPIC-COMMENT structures.

Study Two’s sampling procedure depended upon what has been called here the ‘standard’ linguistic analysis of Mandarin Chinese as a tool for sorting the data. This analysis is sensitive to the facts of Mandarin utterance structure where it differs from the SUBJECT-PREDICATE mechanics central to many Indo-European languages. As such it is a distinct improvement over previous analyses that grew from IE-relevant syntactic and discourse categories. Still, the standard analysis has as one of its foundation assumptions the idea that the topicalizing processes of the language will be analyzable at the level of

sentential syntax.

The assumption is that utterances launched by statements of TOPIC, and demarcated in one of the ways outlined in the earlier section on TOPIC-COMMENT structures, necessarily have limits and extents comparable to those of the syntactically-defined SV(O) sentences of many Indo-European subject-prominent languages. This assumption may turn out to have impeded progress toward an accurate understanding of the utterance structuring dynamics of Mandarin. Examination of the spontaneous utterances encountered in the Mandarin narration data used for this study makes it necessary to question the utility of the grammatical unit of analysis ‘sentence’ in describing the topicalization patterns in the language. In these data, the topicalization patterns observed are not typically so neatly delimited as are the examples one finds in the expositions in grammars of Mandarin, or as are created for grammaticality judgments in linguistic elicitations.

From grammatical treatments of Mandarin one sometimes gets the impression that the field of Chinese linguistic research has tacitly regarded one of its tasks to be that of producing an analysis of topicalization at this level. Transformational generativists have approached the task by positing covert sentence-level constituents that control overt utterance structures. Researchers of the functionalist perspective have focused instead on the semantic relationships that make some sentences easily interpretable and others anomalous. Neither theoretical approach however has questioned the notion that isolable, full-sentential TOPIC-COMMENT constructions would be frequent in naturalistic discourse. Thus, although Chinese linguistics has left such notions as ‘major’ and ‘minor’ sentences behind, there is still the attempt to fit Mandarin topicalization phenomena into a sentential mold. The results of Study Two call this foundational assumption into question. Recall that from among the eleven cartoon narrations, each

several minutes long, it was problematic to locate twenty ‘textbook’ TOPIC-COMMENT utterances that were fluent and gesture-accompanied. That such sentences, thought to be a defining utterance structure of Mandarin, were difficult to extract forces us to ask how representative of typical speech-gesture production such marked topicalized utterances as were found may in fact be. Comrie (1994) outlines these issues at a general level:

“A linguist who is familiar with the structure of language A and who approaches the analysis of a very different language B might adopt one of two extreme positions, ... One extreme would be to assume that language B is going to have essentially the same structure as language A, so that features of language A that are not apparent in language B will be held to be ‘covert’ features of the structure of language B. The opposite extreme is to assume that only ‘overt’ features of language B are relevant to its analysis, thus denying any relevance of properties of language A that are not immediately apparent in language B. ... both approaches are likely to be misguided. There are deep-seated differences between languages of different types, and simply carrying over the structure of one language to another language can lead us to misanalyse language B, to apply categories to it that are simply not relevant to its structure.”

Comrie’s own analysis of topicalization in Chinese is another that tries to work out mechanics of it within the syntactic ‘sentence,’ a unit that may need to be subjected to the same scrutiny that he advocates be devoted to other linguistic categories before porting them among languages.

It is true that the extraction process was restricted to fluent utterances and also to a limited number of utterances from each of the eleven speakers. Some speakers may be more likely to produce TOPIC-COMMENT structures, or to adopt a register in which they are more likely to occur (assuming their frequency of occurrence varies with speech register as seems to be the case in English). Still, TOPIC-COMMENT and SVO utterances were more elusive in the data as a whole than might be expected, given that these structures are considered definitional of Mandarin syntax. In contrast, the narration data of English, classified as a subject-prominent SVO language, yielded many more SVO utterances than were needed to complete the data sets. Such a contrast in the degree of

difficulty of sampling the target utterance structures in these two languages is in a way evidence for the very differences between Mandarin and Chinese that motivated this research; that is, there is something about the utterance-structuring dynamics of Mandarin that is truly different from English, yet it operates at a level other than that of sentential syntax.

The other observations that accumulated in the process of scanning the Mandarin narration data were also presented in the Results as tentative findings that point further research directions. In Mandarin the speech-gesture production process appears configured to put the main pulse of a production unit, more often than is true of English, toward the beginning edge of the production envelope. The remainder of the Mandarin utterance can then often unfold as an intonationally de-stressed unit. For the English speaker, it is the predicate phrase itself within a SUBJECT-PREDICATE construction that seems to drive most of the utterance-structuring machinery. The main speech-gesture pulse of an utterance will typically have its locus there in an English sentence.

In Mandarin gesture, repetitions and holds demonstrate how in that language at times gesture alone may shoulder the burden of carrying TOPIC forward from one production into the next via reduced repeats, post-stroke holds, and so on. Or gesture may pair with speech in particular ways not typical of English to accomplish this same function. The semantic content of gesture is thus reprised within the narration, in reduced or otherwise modulated form, to reflect changes in the information status of that content. Pressure to create or maintain a dynamic discourse focus on TOPIC (broadly construed as ‘aboutness,’ and not necessarily as a vehicle for given information) in Mandarin versus on the sentence PREDICATE in English, may underlie much of the positioning of gesture relative to speech in the two languages, the relative frequency and character of gesture repetitions and holds, and the character of the semantic relationship between speech and

gesture in the two languages.

The differences between Mandarin and English speakers' gestures observed in Study Two may be a function of the topic prominence feature that is theorized to typologically distinguish the two languages. A sampling process framed by a somewhat too narrow linguistic analyses, however, may have somewhat skewed the results reported here, and this is an issue that requires further research.

Future research: an alternative use of narration data. In spite of not having as a goal the re-characterization of the phenomenon of topic prominence in Mandarin, the method employed in Study Two required the examination of connected spontaneous speech-gesture data. The additional observations that accumulated during that process suggest further research directions that seem likely to lead to such a re-characterization. The additional observations suggest features of speech-gesture production that can be examined rigorously in an analysis based on a broader notion of 'comparability of propositional content' than the sentence-meaning notion on which the study described here was based. The cartoon content is ideally structured to support such an analysis, given its episodic format and the simple goal structure within each episode and in the cartoon as a whole. It seems likely that this stimulus structure is what prompts most narrators to construct narratives that are themselves highly episodically structured. Most narrators attempt to re-tell the events within each episode and to re-tell these in their proper sequence. These characteristics of the speakers' narrations mean that it should be a relatively straightforward matter to equate expressive content across speakers and speaker groups at the level of whole cartoon episodes.

On the basis of such data, it will be possible to track elements of discourse content, in both their spoken and gestured realizations, through extended sequences. It

would be possible then to determine, in the case of TOPIC chains for instance, whether or not, as COMMENT pivots to TOPIC along such a chain, each pivot constituent is likely to be accompanied by a gesture. If this turns out to be true, then in theory each of these spoken pivot constituents and its co-occurring gesture arises from something like a single idea unit of discourse; each pivot point is a GP. Not much is known about how such pivot constituents can be structured semantically, the ways in which they may relate to what precedes and follows them, nor what their discourse scope can be. The pivot points in chained sequences do seem however to be a different kind of unit from the clause or sentence that in English often associate with a single gesture. Yet high-frequency structures of spoken Mandarin such as these, together with their associated gestures, link to cover apparently the same narrative ground as is covered by the SUBJECT-PREDICATE structures of English. Certainly, when the analytic unit ‘utterance’ is not limited to the four narrowly defined structures examined in Study Two, but is opened to include this kind of production, we can expect that the findings reported here, for example the values reported for Mandarin gestures-to-utterance ratios, will be at least somewhat transformed.

The exercise of examining spontaneous Mandarin narration data that was embedded in the procedures of Study Two has made it possible to see what the ‘standard,’ sentence-level analysis of topicalization in Mandarin has been trying to capture. There is something quite different about utterance structure and utterance sequencing in Mandarin as compared to English. The standard analysis has promoted an alternative to the traditional SUBJECT-PREDICATE framework used in the analysis of many Indo-European languages. What may have foiled this alternative, however, could be some of the standard elicitation methods of linguistics, and the assumptions they incorporate. Some of these methods are dependent upon grammaticality judgments of sentence-like whole utterances in isolation. A more promising approach would be to

compare sequential analyses of speech-gesture productions that connect in sequence over extended portions of Mandarin and English spontaneous narration. The portions to be compared must still cover comparable narrative ground. The alternative way just outlined of framing the research would shift the composition of the Mandarin data set from one consisting of the kinds of utterance structures that have to date received the most focused attention of linguists, to one consisting of the kinds of structures that Mandarin speakers actually generate in casual, spontaneous narration.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Speech and gesture pattern together in extremely orderly ways. The two modalities function as a fully integrated system at multiple levels. Studies One and Two entered this system via the ‘speech code’; or rather, via theories about the speech code. In retrospect, it is possible to see that in choosing this path of entry, both studies were virtually assured of missing important essentials of the patterning they were intended to explore. The essential limitation in both studies was that the explorations were organized around the unit of the sentence, extracted from its context and considered in isolation from other utterances. In the case of linguistic aspect, it seems likely that in languages like English and Mandarin, the category of aspect is not properly construed simply as one featural setting among several on a verb that is the nucleus of a sentence. Rather, ‘verb aspect’ must certainly be a feature of a larger frame than that encompassed by the reach of single verbs. Slobin and Berman’s (1994) notion of ‘tense-aspect frame’ captures part of the sense intended here, as do theoretical treatments of aspect that seek to extend its scope beyond the ‘internal temporal constituency’ of verbs. The gestural evidence of Study One seems to support these larger frame analyses.

In the case of the comparison of sentence structures across the two languages, there are the concerns outlined in Chapter IV concerning whether ‘sentence’ as a unit of analysis is appropriate for parsing much of what comes under the rubric of

‘topicalization’ in spontaneous spoken Mandarin. Once we distance ourselves from a sentential-syntactic analysis of topic prominence, according to what terms should we understand the widely-noted differences between Mandarin and English in how utterances are structured? There is no answer to this as yet.

More fundamentally, a goal of the exploratory research reported here was to inform the further development of a model of language production based on the hypothetical GP unit of analysis. The focus on isolated sentential utterances constrains how we may apply toward this research goal the findings concerning variations across sentence structure in speech-gesture patterning. Recall that Bates and MacWhinney (1991) state that a model of language production must provide an account of “the process by which native speakers select a set of expressions to convey meaning” in the context of real-time language use. The virtue of a production model organized around the hypothesized GP is that it can in theory constrain and account for the direction taken from one production unit to the next in discourse. Therefore, when any speech-gesture analysis is pared down to decontextualized utterances, prospects dim for explicating how variations in lexico-syntactic resources cross-language interact with the on-line creation of communicative packages in language production. It becomes difficult to show how, as language production progresses from GP to GP, at each step the number of narrative moves is limited that differentiate new from given information in a sufficiently contextualized fashion, while satisfying a given language’s formal linguistic requirements, and at the same time shortening the distance between the current narrative location and the narrative goal.

That a given event component is selected for gestural encoding in speech-gesture ‘pulse’ *b* is, according to GP theory, largely dependent upon what was encoded in speech-gesture pulse *a*, and is further influenced by what is planned for pulse *c*. Further, it is

important not to overlook the fact that the particular *spoken* form that serves as a vehicle for the expression of this event component is as well just as much a function of the preceding context and the upcoming narrative goal.

GP theory thus necessitates an analysis centered on context, and context is what was largely skirted by the methods and analyses of Study Two. To take just one example, it is interesting that SVO utterances in Mandarin appear to be part of an environment in which the PATIENT role is often differentially highlighted. But why should this be so? Why is it not always the case? A speaker's choice of the SVO utterance format, together with the gestural highlighting of the PATIENT case role in any one instance, are both a function of what came before the utterance, and at the same time of what is planned to immediately follow it. This production pattern cannot be fully understood without reference to both.

It has been mentioned at several points throughout this treatment that the two areas of difference in linguistic form between Mandarin and English that provided the impetus for these studies, verb aspect and topic prominence, are contentious areas of linguistic research. The exercise of going through the spontaneous narration data examined here, and attempting to explore them in terms of units of analysis borrowed from this linguistic research, has made clear both why these areas of research are contentious, and as well something of what they have been trying to capture. In both cases, analysis of speech-accompanying gesture shows promise as a means to carry research in these areas forward by contributing additional data on the thinking-for-speaking beyond what may be gleaned from speech alone.

Representational gestures of the sort examined here provide clues to how speakers sort out and organize the domain of referents during real-time speech production. The observations reported suggest that gesture patterns in relation to features of linguistic

form at certain levels of linguistic structure. This demonstrates the utility of the analytic study of gesture for studying processes of thinking-for-speaking. Cross-language research on speech together with the gesture that accompanies it provides a means of determining the extent to which conception is related to linguistic representation in areas of grammatical difference between languages such as the ones examined for Studies One and Two. Since languages differ in how they encode the features of an event representation, spoken utterances in different languages will differ in the features they select or highlight. The gestural component of a speech-gesture unit, reveals the structure and contents of the conception that are active in conjunction with these linguistic representations.

The research presented here lends support to the thesis that, at least in some respects, thinking-for-speaking proceeds independently of specific surface linguistic forms. The gestured indices of aspectual view in event representations are quite similar in speakers of Mandarin and English, despite surface differences in how the two languages handle linguistic aspect. The differences that have been noted between the two languages in this regard must then in certain ways be somewhat superficial. In contrast, the typological feature of topic prominence that distinguishes English and Mandarin appears to generate patterns of thinking-for-speaking that are related to the positioning of utterance components as TOPIC and/or FOCUS constituents. The quite different character of gestural performance in relation to the latter widely-discussed area of grammatical difference between the two languages, suggests that it is centrally involved in basic utterance organization processes.

In the course of carrying out both Studies One and Two, the necessary direction for future research became apparent. This will be one in which modulations in the target parameters are tracked across larger units of discourse than that which can be captured by the unit of a sentence.

REFERENCES

- Aske, Jon. 1989. Path predicates in English and Spanish: a closer look. Papers from the 15th Annual Meeting of the Berkeley Linguistics Society. 1-14. Berkeley, California: BLS.
- Bates, Elizabeth and Brian MacWhinney. 1989. Functionalism and the competition model. The crosslinguistic study of sentence processing, ed. by Brian MacWhinney and Elizabeth Bates, 3-73. Cambridge: Cambridge University Press.
- Binnick, Robert. 1991. Time and the verb: A study of tense and aspect. Cambridge: Oxford University Press.
- Bolinger, Dwight. 1986. Intonation and its parts: melody in spoken English. Stanford, California: Stanford University Press.
- Butterworth, Brian. and Geoffrey W. Beattie. 1978. Gesture and silence as indicators of planning in speech. Recent advances in the psychology of language: formal and experimental approaches, ed. by Robin N. Campbell and Philip T. Smith, 347-360. New York: Plenum Press.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. Subject and topic, ed. by Charles N. Li, 457-489. New York: Academic Press.
- Chappell, Hilary. 1991. Causativity and the *ba* construction in Chinese. Partizipation, ed. by H. Seiler and W. Premper. Tübingen: Gunter Narr Verlag.
- Chao, Yuen-ren. 1968. A grammar of spoken Chinese. Berkeley and Los Angeles: University of California Press.
- Choi, Soonja and Melissa Bowerman. 1991. Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. Cognition. 41.83-121.
- Comrie, Bernard. 1981a. Aspect. Cambridge: Cambridge University Press.
- Comrie, Bernard. 1981b. Language universals and linguistic typology. Oxford: Blackwell.
- Comrie, Bernard. 1985. Topics, grammaticalized topics, and subjects. Papers from the 11th Annual Meeting of the Berkeley Linguistics Society. 265-279. Berkeley, California: BLS.

- Comrie, Bernard. 1994. Cross-linguistic typological variation, grammatical relations, and the Chinese language. Paper presented at the fourth international symposium on Chinese languages and linguistics, Academia Sinica: Institute of History and Philology, Taipei, Taiwan.
- Duncan, Susan. 1992. The conceptual representation of temporal aspect in English and Mandarin Chinese: gestural indices. Paper presented at the Annual Conference of the Linguistic Society of Belgium in Antwerp, Belgium.
- Eifring, Halvor. 1995. Clause combination in Chinese. Leiden, The Netherlands: E.J. Brill.
- Erbaugh, Mary S. Unpublished manuscript. The Chinese language and its effect on cognition: a debate over chopsticks versus forks.
- Erbaugh, Mary S. 1990. Mandarin oral narratives compared with English: the pear/guava stories. Chinese language teachers association journal.
- Erbaugh, Mary S. 1993. The acquisition of Mandarin. The cross-linguistic study of language acquisition, vol. 3, ed. by Dan I. Slobin. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Gigerenzer, Gerd and Daniel G. Goldstein. 1994. Mind as computer: the birth of a metaphor. Creativity research journal.
- Grace, George W. 1987. The linguistic construction of reality. New York: Croom Helm in association with Methuen.
- Hadar, Uri and Brian Butterworth. 1992. Iconic gestures, imagery and word retrieval in speech. Psychological Review. 99. 116-148.
- Havilland, John. 1993. Anchoring, iconicity and orientation in Guugu Yimithirr pointing gestures. Linguistic Anthropology. 1. 3-45.
- Huang, James C-T. 1984. On the distribution and reference of empty pronouns. Linguistic Inquiry. 15:4. 531-74.
- Huang, Yan. 1994. The syntax and pragmatics of anaphora: A study with special reference to Chinese. Cambridge: Cambridge University Press.
- Kendon, Adam. 1972. Some relationships between body motion and speech. Studies in dyadic communication, ed. by Aron Wolfe Siegman and Benjamin Pope, 177-210. New York: Pergamon Press.
- Kendon, Adam. 1980. Gesticulation and speech: two aspects of the process of utterance. The relation between verbal and nonverbal communication, ed. by Mary Ritchie Key, 206-227. The Hague: Mouton Publishers.
- Kendon, Adam. 1995. Andrea Di Jorio - the first ethnographer of gesture? Visual anthropology. 7. 375-394.

- Kita, Sotaro. 1990. The temporal relationship between gesture and speech: a study of Japanese-English bilinguals. MS. Department of Psychology, University of Chicago.
- Kita, Sotaro. 1993. Language and thought interface: a study of spontaneous gestures and Japanese mimetics. Ph.D. dissertation. University of Chicago.
- Krauss, Robert M., Palmer Morrel-Samuels, and Christina Colasante. 1991. Do conversational hand gestures communicate? *Journal of personality and social psychology*. 61.743-754.
- Lakoff George and Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- La Polla, Randy J. 1995. Arguments against 'subject' and 'direct object' as viable concepts in Chinese. *Bulletin of the Institute of History and Philology*. Nan Kang, Taiwan: Academia Sinica.
- Levelt, Willem J.M. 1989. *Speaking*. Cambridge, Massachusetts: M.I.T. Press.
- Levy, Elena and David McNeill. 1992. Speech, gesture, and discourse. *Discourse processes*. 15.277-301.
- Li, Charles N. and Sandra Thompson. 1976. Subject and topic: a new typology of language. *Subject and topic*, ed. by Charles N. Li, 457-489. New York: Academic Press.
- Li, Charles N. and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley and L.A.:University of California Press.
- Li, Ping. 1991. The acquisition of aspect in Mandarin Chinese. Dordrecht.
- Li Zhen-qing. 1988. *Practical Chinese dialog II*. Taipei: Mandarin Training Center, National Taiwan Normal University.
- Liu, Hua, Elizabeth Bates, and Ping Li. 1992. Sentence interpretation in bilingual speakers of English and Chinese. *Applied psycholinguistics*. 13.451-484.
- Longacre, Robert. 1970. Sentence structure as a statement calculus. *Language*. 46. 783-815.
- Lucy, John A. 1992. *Language diversity and thought: a reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- McCawley, James D. 1989. Notes on Li and Thompson, *Mandarin Chinese: a functional reference grammar*. *Journal of the Chinese language teachers association*. 24:1. 19-42.
- McCullough, Karl-Erik. and Susan Duncan. 1993. Gesture and linguistic typology in Mandarin Chinese and English. Paper presented at the Linguistic Society of America Annual Meeting, Los Angeles.

- McNeill, David. 1985. So you think gesture is non-verbal? *Psychological Review*. 92. 350-371.
- McNeill, David. 1987. *Psycholinguistics: a new approach*. New York: Harper and Row.
- McNeill, David. 1992. *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, David and Elena Levy. 1993. Cohesion in speech and gesture. *Discourse Processes*. 16.363-386.
- McNeill, David and Susan Duncan. To appear. The growth point. *Gesture: an emerging field of research*, ed. by Adam Kendon, David McNeill and Sherman Wilcox.
- Miao, X. 1981. Word order and semantic strategies in Chinese sentence comprehension. *International journal of psycholinguistic research*. 8.23-33.
- Morrel-Samuels, Palmer and Robert M. Krauss. 1992. Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of experimental psychology: learning, memory and cognition*. 18.615-622.
- Ramsey, R.S. 1987. *The Languages of China*. Princeton, NJ: Princeton University of Press.
- Reinhart, Tanya. 1982. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*. 27.53-94.
- Saussure, Ferdinand de. [1916] 1959. *Course in general linguistics* (W. Baskins trans.). Reprint. New York: Philosophical Library.
- Sheng Ma, Jing-heng. 1985. *A study of the Mandarin Chinese verb suffix ZHE*. Taipei: The Crane Publishing Company, Ltd.
- Shi, Dingxu. 1989. Topic chain as a syntactic category in Chinese. *Journal of Chinese linguistics*. 17:2.223-261.
- Slobin, Dan I. 1987. Thinking for speaking. *Papers from the 13th Annual Meeting of the Berkeley Linguistics Society*, ed. by Jon Aske, N. Beery, L. Michaelis, & H. Filip, 480-519. Berkeley, California: BLS.
- Slobin, Dan I. In press, a. Typology and rhetoric: verbs of motion in English and Spanish. *Grammatical constructions: their form and meaning*, ed. by M. Shibatani and Sandra A. Thompson. Oxford University Press.
- Slobin, Dan I. In press, b. Learning to think for speaking: native language, cognition, and rhetorical style. *Rethinking linguistic relativity*, ed. by John Gumperz and Stephen C. Levinson. Cambridge, UK: Cambridge University Press.
- Slobin, Dan I. and Ruth A. Berman. 1994. *Relating events in narrative: a cross-linguistic developmental study*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Supalla, Ted, Elissa L. Newport, Jenny Singleton, Sam Supalla, T. Metlay and G. Coulter. 1993. *Test Battery for American Sign Language: Morphology and Syntax*, Burtonsville, Maryland: Linstok Press.
- Talmy, Leonard. 1975. Semantics and syntax of motion. *Syntax and semantics*, ed. by J. Kimball, 457-48. New York: Academic Press.
- Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. *Language typology and syntactic description, volume III: grammatical categories and the lexicon*, ed. by Timothy Shopen, 57-149. Cambridge, UK: Cambridge University Press.
- Talmy, Leonard. 1991. Path to realization: a typology of event conflation. *Proceedings of the 17th annual meeting of the Berkeley Linguistics Society*, 480-519. Berkeley, CA: Berkeley Linguistics Society.
- Tai, James H-Y. 1985. Temporal sequence and Chinese word order. *Iconicity in syntax*, ed. by John Haiman, 49-72. Amsterdam/Philadelphia: John Benjamins.
- Tsao, Feng-fu. 1987. A topic-comment approach to the *ba* construction. *Journal of Chinese linguistics*. 15:1. 1-50.
- Tsao, Feng-fu. 1990. *Sentence and clause structure in Chinese: a functional perspective*. Taipei: Student Book Co., Ltd.
- Tzeng, Ovid and Daisy Hung. 1984. Psychological issues in reading Chinese characters. *Neurological studies in processing Chinese languages*, ed. by H. Kao and R. Hoosain, 341-368. Hong Kong: University of Hong Kong Press.
- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, New York: Cornell University Press.
- Vygotsky, Lev S. 1986. *Thought and Language*, ed. by A.Kozulin. Cambridge, Massachusetts: MIT Press.
- Xu, Liejiong and D.Terrence Langendoen. 1985. Topic structures in Chinese. *Language*. 61. 1-27.
- Zinchenko, V.P. 1985. Vygotsky's ideas about units for the analysis of mind. *Culture, communication, and cognition. Vygotskian perspectives*, ed. by James V. Wertsch, 78-95. Cambridge: Cambridge University Press.

APPENDIX A

ASPECT SAMPLE

Key to the transcription:

Grammatical features:

PRG	progressive aspect marker (“zai”/在)
DUR	durative aspect marker (“-zhe”/著)
PRF	perfective aspect marker (“-le”/了)
CMPL	completive verb particle
-CL	noun classifier
-MOD	adjectival, possessive, or adverbial marker (“de”/的 or 得)
OBJ	object-fronting particle (“ba”/把)
PAS	“passive” marker (“bei”/被)
BA	topic-marking particle/discourse particle (“ba”/吧)
NE	topic-marking particle/discourse particle (“ne”/呢)
MA	topic-marking particle/discourse particle (“ma”/嘛)
ONOM	onomatopoeia

Speech features:

*	self-interrupt
/	unfilled pause
< ... >	filled pause
#	breath pause
%	non-speech sound (a <i>laugh</i> unless otherwise indicated)

Gesture features:

[...]	extent of gesture
bold	stroke phase
<u>underline</u>	pre- or post-stroke hold
^	beat

Note: the letters at the left margin are subject identification codes, followed by the source of the particular sample — “c” for cartoon narration, “v” for vignette response, and “m” for movie narration.

Mandarin Aspect Sample

1. *Progressive-marked speech samples:*

- hhf/c [ye zai wang **lou-xia kan kan kan jiu kan-dao yi-zhi da ^ye mao #**]
 也在往樓下看看看就看到一隻大野貓
 also PRG toward building-down look look then see-CMPL one-CL big wild cat
 (he) is also looking, looking, looking down then sees a big wild cat

Progressives (continued):

hhf/c [jie-guo ta zheng-**zai zou zou**] zou
 結果他正在走走走
 result he PRG walk walk walk
 so he's walking walking walking

hhf/c shi nei-z[hi jin-si-**que gen ta-de nei-ge*** / nei-ge / [zhu-ren %]
 是那隻金絲雀跟他的那個* / 那個 / 主人 %
 be that-CL gold-silk-sparrow with he-MOD that-CL* / that-CL / owner %
 it's the canary with that* / that / owner of his %

zai kai ne]i-liang dian-che %
 在開那輛電車 %
 PRG drive that-CL train %
 driving the train %

hhx/v nei-ge* yi-zhi ci-wei [/ **zou / zai zou**]
 那個* 一隻刺蝟 / 走 / 在走
 that-CL* one-CL hedgehog / walk / PRG walk
 the* a hedgehog / walks / is walking

csx/v kan-dao yi-ge hen ke-ai-de xiao wa-wa [zai **xuan-zhuan**]
 看到一個很可愛的小娃娃在旋轉
 see-CMPL one-CL very cute-MOD small doll PRG revolve-turn
 (I) see a very cute little doll turning circles

csx/v zhe* zhei-ge yan-hui-gang ke-[neng zai **bo-li pian shang zai hua-dong**]
 這* 這個煙灰缸可能在玻璃片上在滑動
 this* this-CL ashtray maybe on glass sheet surface PRG slide-move
 this* this ashtray might be sliding on a sheet of glass

csx/v yi-ge xiao yi-zi [% **zai yi-dong**]
 一個小椅子 % 在移動
 one-CL small chair % PRG shift-move
 a small chair % is moving

csx/v [ye xiang* / **ye xiang zai**] gun
 也像* / 也像在滾
 also appear / also appear PRG roll
 seems to be rolling, too

Progressives (continued):

- csx/v ran-hou zai wang qian* / <uh> [/ **wang** zuo-bian] hua-dong
 然候在往前* / <uh> / 往左邊滑動
 next PRG toward front / <uh> toward left-side slide-move
 then (it) is moving forward / <uh> to the left
- wc/c [ta zai shang-mian yi-bianr zou yi-bianr tiao hao-xiang zai tiao hua-er]-zi wu
 他在上面一邊走一邊跳好像在跳華爾茲舞
 he on top-side one-side walk one-side jump appear PRG dance waltz dance
 he alternately walks and jumps on top (of the wires) seems to be dancing a waltz
- wc/m [nei shi-hou jiu zai si-kao **zai**] dou-zheng #
 那時候就在思考在鬥爭 #
 that time just PRG think PRG struggle #
 at that time she is thinking and struggling #
- wc/m [ran-hou nei-ge nan-de zhan nei] si-chu [**si-chu**] xun-zhao
 然候那個男的站那四處四處尋找
 next that-CL male-MOD stand there everywhere everywhere search
 then the guy stands there looking all around (for her)
- cy/c ran-hou nei xiao [niao **zheng-zai nar*** # zh]eng-zai yang-tai / nei-ge nar
 然候那個小鳥正在那兒* # 正在陽臺 / 那個那兒
 next that small bird PRG there* # PRG balcony / that there
 the bird is (sitting) there right at that point # just then on the balcony there
- xmh/v yi-ge xiao wa-wa zai zou-dong
 一個小娃娃在走動
 PRG walk-move
 is moving
- xmh/v nan-sheng zai cao-di shang zai gun-dong
 男生在草地上在滾動
 man on lawn-top PRG roll-move
 a man is rolling on the grass
- xmh/v ci-wei zai zou-dong
 刺蝟在走動
 porcupine PRG walk-move
 a porcupine is walking

Progressives (continued):

- xmh/v ge-cao-ji zai chu-cao
ge-草機在除草
lawn mower PRG weed grass
a lawn mower is mowing the lawn
- ww/c gen nei-ge lao tai-p[o ran-hou **Tweety Bird zai # kai dian-che<eee>**]
跟那個老太婆然候 Tweety Bird 在 # 開電車
with that-CL old lady next Tweety Bird PRG # drive train
with the old lady then it's Tweety Bird # driving the train
- ww/c [ran-hou **zai dian nei-ge<eee>**] Sylvester
然候在電那個 Sylvester
next PRG electric that-CL Sylvester
then is electrocuting Sylvester
- ml/c [jie-guo ta **kan-dao ^HO yi-zhi ^mao zai kan ta xiang chi ta //**]
結果他看到 HO 一隻貓在看他想吃他
result he see-CMPL ONOM one-CL cat PRG watch him want eat him
so he sees YIKES there's a cat watching him (who) wants to eat him
- zwh/c jiu fa-xian [shi nei-ge lao tai-tai **zai kai nei-ge dian-che**]
就發現是那個老太太在開那個電車
then discover be that-CL old lady PRG drive that-CL electric-car
then he discovers that it's the old lady driving the train
- zwh/v yi-ge chi shi zai zhao-zi -mian][/ **zai dong** /]
一個尺是在桌子面 / 在動 /
one-CL ruler be at table -face / PRG move
a ruler is moving on top of a table
- zwh/v yi-[ge mo-xing-de nei-zhong tuo-tu]o-che[shi* / zai*][/ **zai zhao-shang dong**]
一個模型的那種拖拖車是* / 在* / 在桌上動
one-CL model-MOD that-kind pull-pull-vehicle be* / on* / PRG table-top move
one of those model tow trucks is on* is moving on the table
- zwh/v shi yi-zhi mo-xing-de xiao ji / <uh> [**zai tiao**]
是一隻模型的小 ji / <uh> 在跳
be one-CL model small chicken / <uh> PRG jump
it's a small model chicken jumping
- zwh/v wan-ju bei-bei / ke-neng shi nu-hai {...} dai yi-huar [/ ran-hou **zai pa**]
玩具背 bei-bei / 可能是女孩 {...} 待一會兒 / 然候在爬
toy baby / perhaps be female-child {...} wait a while / next PRG crawl
a babydoll / maybe a girl {...} waits a little bit / then she's crawling

2. *Durative-marked speech samples:*

hhf/c [# ta yong hen da-de wang-yuan-**jing ding-zhe ta kan /**

他用很大的望遠鏡盯著他看

he use very big-MOD telescope stare-DUR him look

he uses a very big telescope to stare at him /

ran-hou jiu kan-zhe ta kan-zhe ta ran-hou /

然後就看著他看著他然後

next just look-DURhim look-DUR him next

then just watches him and watches him then

hhf/c [ran-hou ta **jiu la-zhe nei-gen shen**][g-zi

然後他就拉著那根繩子

next he just pull-DUR that-CL rope

then he pulls on a rope

hhf/c xia][ng na / dian-che **jiu zhei-yang yi-zhi zhui-zhe ta #**]

像 那 / 電車就這樣一直追著他

resemble that / train just this-manner one-direct chase-DUR him

then apparently the / train chases him continuously like this

hhf/c [# ta jiu yan-zhe **jie-dao yi-zhi**] [pao

他就沿著街道一直跑

he just along-DUR road straight run

so he runs straight along the road

csx/v wo shi kan-dao [yi-ge xiao mi-feng **rao-zhe yi-ke**] shu zai xuan-zhuan

我是看到一個小蜜蜂繞著一棵樹在旋轉

I be see-CMPL one-CL small bee circle-DUR one-CL tree PRG circle-turn

I see a little bee circling a tree and turning

csx/v ta hao-xiang shi rao-zhe quan-quan / rao-zhe [ta-de* / ta-de*

他好象是繞著圈圈 / 繞著他的* / 他的*

he appear be circle-DUR circle-circle / circle-DUR he-MOD / he-MOD

apparently it goes around a circle / goes around his* / his*

/ **rao yi-quan /**]

/ 繞一圈 /

/ circle one-circle /

/ goes around a circle /

lj/c ran-hou jiu ting-zhe [DING-LING-GUANG-LANG-DING-

然後就聽著 ONOM

next just listen-DUR ONOM

then (one) hears DING-LING-CLANG-DING-LING-CLANG

Duratives (continued):

- LING-^GUANG-LAN]G nei-ge za-ping-zi-de sheng-yin
 ONOM 那個砸瓶子的聲音
 ONOM that-CL smash-bottle-MOD sound
 CLANG the sound of breaking bottles
- lj/c nei niao long-zi wai-mian jiu you yi-ce][**ng bu bao-zhe** /]
 那鳥籠子外面就有一層布抱著
 that bird cage outside just have one-CL cloth cover-DUR
 the outside of the bird cage is covered with a layer of cloth
- wc/c [ta jiu shi / **dai-zhe yi-ge wang-yuan-jing kan**] / gang-hao kan-dao nei-ge
 他就是 / 帶著一個望遠鏡看 / 剛好看到那個
 he then be / wear-DUR one-CL telescope look just-then see-CMPL that-CL ...
 he is looking through a telescope / and just as he sees the ...
- wc/c ya [ran-hou lao-yuan **kan-zhe wang-yuan-jing**]
 ya 然後老遠看著望遠鏡
 yeah next old-distant look-DUR telescope
 yeah then (he) looks through the telescope for a long time
- wc/c ta jiu zai [nar bei-zhe **shou # [tuo-zhe -ge yi-ba]-de**
 他就在那背著手 # 拖著個易巴得
 he just at there behind-DUR hand # drag-DUR -CL tail-MOD
 so there he is with his hands behind his back dragging his tail
- zai nar zou-lai-zou-qu zou-lai xia-ruan pao <heh> #**]
 在那走來走去走來 xia-ruan 跑 <heh>
 at there walk-come-walk-go-walk-come purposelessly run <heh>
 walking back and forth there <heh>
- wc/c shi nei-ge niao [qu le ne jiu **bao-zhe jiu pao ni zhi-dao b**][a
 是那個鳥去了呢就抱著就跑你知道吧
 be that-CL bird go PRF NE just carry-DUR just run you know BA
 it's the bird and goes, then runs carrying it, y'know?
- mw/c zui-hou zhei-ge # [lao tai-tai ye shi **dai-zhe ta dao**]
 最後這個老太太也是帶著他到
 most-after this-CL # old lady also be carry-DUR him arrive
 at the end, the # old lady arrives carrying him
- mw/c [ta jiu yao cai-zhe zhei-ge dian-wang cai-guo-qu]
 他就要踩著這個電網踩過去
 he just will step-DUR this-CL electric-net step-cross-go
 then he's gonna step across on the wires

Duratives (continued):

- mw/c [da-gai jiu **yi-zhi gen-zhe zhei-ge dian-wang jiu*** jiu wang qian zou]
 大概就一直跟著這個電網就* 就往前走
 probably just all-the-way follow-DUR electric-net just* just toward forward walk
 then he probably just follows the wires and then walks forward
- ww/c ni kan-dao Sylves[ter **na-zhe yi***] [yi-tiao sheng-zi xiang Tai-shan nei-yang-zi]
 你看到 Sylvester 拿著一* 一條 繩子像 Tarzan 那樣子
 you see-CMPL Sylvester hold-DUR one* one-CL rope resemble Tarzan that-way
 you see Sylvester holding onto a rope looking like Tarzan
- ww/c ni kan-dao Sylves[ter **na-zhe yi***] [yi-tiao sheng-zi xiang Tai-shan nei-yang-zi]
 你看到 Sylvester 拿著一* 一條 繩子像 tai-shan 那樣子
 you see-CMPL Sylvester hold-DUR one* one-CL rope resemble Tarzan that-way
 you see Sylvester holding onto a rope looking like Tarzan
- cy/c yin-wei [nei- hou **shi chuan-zhe yi-fu**] tiao-wu #
 因為那猴是穿著衣服跳舞 #
 because that- monkey be wear-DUR clothes dance #
 because the monkey is dancing wearing clothes #
- cy/m [jiu nei-ge **yuan-yuan-de** nei-ge wu-ding] t[a **shun-zhe** nei-ge pa-sha]ng /
 就那個圓圓的那個屋頂他順著那個爬上 /
 just that-CL round-round-MOD that-CL roof he along-DUR that-CL climb-up /
 then he climbs up along the curved roof
- zwh/c ran-hou n[a-zhe wang-yuan-jing **ran-hou jiu** fa-xian le] nei-ge niao
 然候拿著望遠鏡然候就發現了那個鳥
 next hold-DUR telescope next just discover PRF that-CL bird
 then he's holding the telescope and then discovers the bird
- ml/c nei mao jiu k[ai-shi **pa dao / <aiii> yan-zhe nei-ge s**][hui / pa shui-guan
 那貓就開始爬到 / <aiii> 沿著那個水 / 爬水管
 that cat just begin climb to / <aiii> along-DUR that-CL water / climb water-pipe
 then the cat begins to climb / <aiii> along the drainpipe
- ml/c ta kan-dao yi-zhi hou-zi <ho> / chuan-zhe nei [# **yi-jian yi-fu**] #
 他看到一隻猴子 <ho> / 穿著那 # 一件衣服
 he see-CMPL one-CL monkey <ho> / wear-DUR that # one-CL clothes #
 he sees a monkey <ho> / wearing those # an outfit #
- llm/c ran-hou nei-ge mao ne / jiu yi-tian dao wan]
 然候那個貓呢 / 就一天到晚
 next that-CL cat NE / just one-day to late
 then the cat / from morning 'til night

Duratives (continued):

- [jiu na-**zhe wang-yuan-jin***] [/ -jing si-chu kan
就拿著望遠鏡* / 鏡四處看
just pick up-DUR telescope* / -scope four-place look
then looks everywhere through a telescope
- llm/c [you yi-* / you / yi][-ge long-zi li-mi][an **guan-zhe yi-zhi mao /**]
有一* / 有 / 一個籠子裡面關著一隻貓
have one* / have / one-CL cage inside contain-DUR one-CL cat /
there's a* / there's / a cat in a cage /
- llm/c ta ji[u / **yan-zhe xie-**][po zhei-yang gun gun gun gun /
他就 / 沿著斜坡這樣滾滾滾滾 /
he just / along-DUR slant-slope this-manner roll roll roll roll /
then he / rolls along the slant like this /
- llm/c nei-zhong / dian-che [jiu-shi shang-mian hui **la-zhe nei-ge dian**]-che ma
那種 / 電車就是上面會拉著那個電車嘛
that-kind / train just-be top-side will pull-DUR that-CL train MA
the kind of train that is pulled by a thing on top, you know?

3. *Perfective-marked speech samples:*

- hhf/c bu zhi-dao wei-shen-ma nei-zhi mao hao-xiang] [**fa-xian** le #]
不知道為甚麼那隻貓發現了
not know why that-CL cat appear discover PRF
(I) don't know why, but the cat seems to discover (the bird)
- hhf/c [ta jiu bang le yi-ge **sheng-zi**]
他就綁了一個繩子
he just tie PRF one-CL rope
then he ties a rope
- hhf/c OK ran-hou / ta jiu [hen gao-xing jiu kai **le men jin-qu**]
OK 然候 / 他就很高興就開了門晉去
OK next / he just very happy open PRF door enter-go
OK, then / he happily opens the door and goes in
- hhf/c ran-hou diao-xia-lai yi-[tan zhei-**kuai you fei le** #]
然候掉下來一彈這塊又飛了
next drop-down-come ASP-upon spring this-piece again fly PRF
then (he) drops down and upon springing on this piece flies again

- hhf/c jiu zai zhuan-jiao nei-bian d[a-gai shi ba ta **da-hun** le /]
就在轉角那邊大概是把他打昏了
just at corner that-place probably be OBJ him hit-dizzy PRF
(he) probably knocks him senseless around the corner
- cy/c ta jiu [ba nei **qiu** cong nei] guan-zi li sai jin-qu le
他就把那球從那管子裡塞進去了
he then OBJ that ball from that pipe inside squeeze-enter-go PRF
he squeezes the ball into it.
- cy/c jiu cong niao long-zi] [li **fei-dao** jiu pao] hui nei-ge wu-zi li qu le
就從鳥籠子裡飛到就跑回那個屋裡去了
then from bird cage inside fly to* then run return that-CL house inside go PRF
from inside the bird cage flies to* then he immediately runs back into the house
- cy/c ta bao-zhe nei niao long-zi jiu* [**jiu chu**-men-qu le]
他抱著那鳥籠子就* 就出門去了
it carry-DUR that bird cage then* then out door go PRF
It carries the bird cage and goes out.
- cy/c [ta jiu **jian**] le yi-ge te-bie zhong-de shen-me dong-xi <ah>
他就撿了一個特別重的甚麼東西 <ah>
he then pick-up PRF one-CL extremely heavy MOD what something AH
then he picks up something very heavy
- cy/m ran-hou ta j[iu* jiu **zhua** le yi]-ge ren nei-yang-zi
然後他就* 就抓了一個人那樣子
next he just* just* seize PRF one-CL person that-way
then* then he seizes the man
- lw/c cong xin you co[ng nei-**ge pai**-shui]-guan pa-shang-qu le
從新又從那個排水灌爬上去了
from new again via that-CL drainpipe climb-up-go PRF
starting over he again climbs up through the drainpipe
- lw/c jiu zhuang-dao nei-ge qiang shang [le **jiu diao**-xia-qu le]
就撞到那個牆上了就掉下去了
just crash-to that-CL wall-surface PRF just drop-down-go PRF
then he crashes into the wall and drops down
- llm/c -hou nei- gui-fu ren ne jiu zai* [/ jiu* jiu ba ta **gei gan**-chu-l]ai le
候那貴婦人呢就在* / 就* 就把他給趕出來了
next that honor person NE just at* / just* just OBJ him give hurry-exit-come PRF
then the old lady you know, then she hurries him out

Perfectives (continued):

- llm/c ran-hou nei-ge mao jiu zai zhong-jia][n bei da / -**shang le**]
 然候那個貓就在中間被打傷了
 next that-CL cat just at middle PAS hit / -injure PRF
 then the cat, caught in between, is hit and injured
- llm/c nei ta ji[u na le yi-gen **xiang jia**][o
 那他就拿了一根香蕉
 that he then take PRF one-CL banana
 now, he takes a banana
- llm/c zhu ren jiu zai xia-mia][n ba ta you **gei da-hun le**
 主人就在下面把他打昏了
 owner person just at down-side OBJ him again give hit-dizzy PRF
 the owner is down there and hits him again
- suo-yi nei xiao niao**] [you mei-you bei zhua-dao
 所以那小鳥又沒有被抓到
 therefore that little bird again not-have PAS catch-CMPL
 so the little bird again was not captured
- llm/c ta ji[u **pao-diao le**]
 他就跑掉了
 he just run-away PRF
 then he runs away
- csx/v ran-hou mu-ma / [**shuai-xia-lai le**]
 然候木馬 / 摔下來了
 next wood-horse / tumble-down-come PRF
 then the hobby horse tumbles down
- csx/v ran-hou / zhi-hou ta bei yi-dong-[de **hua-xia-lai le**
 然候 / 之候他被移動得滑下來了
 next / after he PAS move-MOD slide-down-come PRF
 then / after that he is slid down
- csx/v shi bu-shi hou-mian-[de **yi dia**]o-xia-lai le
 是不是候面的易 掉下來了
 be not-be back-side-MOD tail fall-down-come PRF
 is it the tail on his backside that falls off?
- csx/v ran-hou kan-dao yi-ge qua[n-**quan** /] ta ting-xia-lai le
 然候看到一個圈圈 / 他停下來了
 next see-CMPL one-CL circle / it stop-down-come PRF
 then (it) sees a circle / and stops

Perfectives (continued):

- lj/c ran-hou cong nei-ge / shan][g-mian **HUA**-de] / reng-xia-qu le
 然候從那個 / 上面 ONOM 得 / 扔下去了
 next from that-CL / top-side ONOM-MOD / throw-down-come PRF
 then he throws it down from on top WHOOSH
- lj/c ta nei [niao **jiu** fei*] / <nn> fei zou le
 他那鳥就飛* / <nn> 飛走了
 he that bird just fly* / <nn> fly away PRF
 then the bird flies* flies away
- lj/c ran-hou [**ta zi-ji** chu]an-shang le
 然候他自己穿上了
 next he himself don-on PRF
 then he puts (them) on himself
- lj/c ran-hou ta te gao-xing [jiu ba **nei-ge** l* la-kai] le /
 然候他特高興就把那個拉* 拉開了
 next he especially happy just OBJ that-CL pu* pull-open PRF
 then he's really happy so he pu* pulls it off
- lj/c [<ah> **jiu*** /] jiu qu le
 <ah> 就* / 就去了
 <ah> just* / just go PRF
 <ah> then* / then (he) just goes
- wc/c ta ben-lai hao pang hao pang-de shen-][cai shi-**jing shi-jing** ji-bian le]
 他本來好胖好胖的身材勢勁勢勁擠扁了
 he originally very fat very fat-MOD body forceful forceful squeeze-flat PRF
 he starts out fat but then is forcefully squeezed flat
- wc/c [lai **dian-che** le]
 來電車了
 come train PRF
 the train comes
- mw/c xiao niao ne [/] tao-zou le
 小鳥呢 / 逃走了
 little bird NE / flee-leave PRF
 the little you know, flees
- mw/c [shuai dao ta **zui** li qu le /]
 摔到他嘴裡去了
 fall-to his mouth inside go PRF
 (it) falls into his mouth

Perfectives (continued):

mw/c [ta **jiu** pao le]
 他就跑了
 he just run PRF
 then he runs

mw/c [cong dang-zhong ye **xi-qu** le] jiao-xun
 從 dang-zhong 也吸取了教訓
 from { ... } also absorb PRF lesson
 from this (we) also absorb a lesson

mw/c lao tai-tai yi-kan] [jiu # rang ta jin-q]u le
 老太太一看就 # 讓他晉去了
 old lady ASP-upon just # make him enter-go PRF
 as soon as the old lady sees him, she lets him in

ww/c nei-zhi mao jiu hao-xiang ba nei-ge qiu c][**hi-xia-qu** le #]
 那隻貓就好像把那個求吃下去了
 that-CL cat just appear OBJ that-CL ball eat-down-go PRF
 then apparently the cat swallows the ball

ww/c ke-shi ta bei nei-ge lao tai-p[o kan-dao **le** /]
 可是他被那個老太潑看到了
 but he PAS that-CL old lady see-CMPL PRF
 but he is seen by the old lady

ww/c ran-hou ta you] [/ **hui-jia** le]
 然後他又 / 回家了
 next he again / return home PRF
 then he again goes back home

zwh/c [/ jiu ba nei-ge mao **gei reng-chu-lai** le /]
 就把那個貓給扔出來了
 just OBJ that-CL cat give throw-exit-come PRF
 the (she) throws the cat out

zwh/c <uh> kan-dao ta le jiu ba ta gei <e[ehh> da-**chu-lai** le #]
 <uh> 看到他就把他給 <eehh> 打出來了
 <uh> see-CMPL him PRF just OBJ him give <eehh> hit-exit-come PRF
 <uh> (she) sees him and <eehh> hits him away

zwh/c ran-hou zh[ei lao **tai-tai** jiu guo-la]i le shuo #
 然後這老太太就過來了說
 next this old lady just across-come PRF say # ...
 then the old lady comes over and says # ...

Perfectives (continued):

- zwh/c lao tai-tai jiu na nei-ge s[an % yi-xia **jiu ba** ta gei da-y]un le
 老太太就拿那個傘 % 一下就把他給打暈了
 old lady just pick up that-CL umbrella % once just OBJ him give hit-dizzy PRF
 then the old lady picks up the umbrella and knocks him dizzy
- ml/c nei-zhi niao #] [/ **jiu** # / wang shang-mian* / wang shang-mian qu le /]
 那隻鳥 # / 就 # / 往上面* / 往上面去了
 that-CL bird # / just # / toward top side* / toward top side go PRF
 the bird # / then # / goes up* / goes up
- ml/c [/ **jiu pao le** /]
 就跑了
 just run PRF
 then (he) runs
- ml/c su[o-yi / ta **na le wang-yuan-jing kan**]
 所以 / 他拿了望遠鏡看
 therefore / he pick up PRF telescope look
 so he picks up a telescope and looks
- ml/c jie-gu[o # **ta pao-chu-qu le**]
 結果 # 他跑出去了
 result # he run-exit-go PRF
 so # he runs out
- ml/c nei-zhi m* / mao [HO / / you xiang **le zhei-ge**] miao-ji #
 那隻貓* / 貓 HO // 又想了這個妙計 #
 that-CL ca* / cat HO // again think PRF this-CL excellent plan #
 the ca* / cat WOW // thinks up another excellent plan #
- zp/c ta jiu ba n][ei-ge xiao hai gei **cang-q**]i-lai le
 她就把那個小孩給藏起來了
 she just OBJ that-CL small child give cage-rise-come PRF
 then she locks the child up in a cage
- zp/c ba ta nei-ge cong sh][ang-mian gei ta **da-xia-qu le**]
 把他那個從上面給他打下去了
 OBJ him that-CL from top-side give him hit-down-go
 (she) knocked him down from above
- zp/c zhan zai nei-ge niao long-zi li-mian kan-dao [ba ta **da-xia-qu le**]
 站在那個鳥籠子裡面看到把他打下去了
 stand in that-CL bird cage inside see-CMPL OBJ him hit-down-go PRF
 (she) stands inside the bird cage, sees him and hits him down

Perfectives (continued):

- zp/c da le yi-x[ia ran-hou **hou-zi jiu pao-diao le**]
 打了一下然後猴子就跑掉了
 hit PRF one-down next monkey just run-away PRF
 (she) hits him down and then the monkey runs away
- zp/c nei-zhi mao ne zai chuang sh[ang **ting-dao le t**][a shuo -ju hua #
 那隻貓呢在窗上聽到了他說句話
 that-CL cat NE at window top listen-CMPL PRF he say -CL speech
 the cat you know, on the window sill, overhears him say something

English Aspect Sample1. *Progressive-marked speech samples:*

- viv/c well # <uh> [Sylveste][r's sitting in the **mail boxes**]/
- viv/c she crawls up in the tr[ansom **she's looking down**] at him outside
- viv/c and h[e's **trying to**] [**masquerade**] #
- viv/c and / [he* **so he's running** along] the wire
- lau/c um / [the bird **is driving* #**] [**the canary is driving** the streetcar #]
- san/c [/ **and you see him rolling down** the street] into a bowling alley #
- san/c the next you see # Sylvester # [making all **these calculations** /]
- san/c the trolley car comes] [**along and he's trying to r**][**un a**][**head**][**of it #**]
- jan/c so [he's **looking** through binoculars]
- jan/c [he's **listening** ^so] [**he goes**]
- jan/c drawing bo[ard and **he's making** e]laboarte plans
- jan/c [he's **trying to walk across those** to get in^to the window] #
- cel/c [but she's **sort of playing** along] #
- cel/c and / he'[s out in the **middle of nowhere balancing** very care][fully
- cel/c [**and then Granny is**] [in there **ringing the bell** next to him]

Progressives (continued):

- adi/m and [he's **kind of like mulling around**]
- adi/m alice has to [sort of] # [come i][n to **whe**][re her **father is reading***<uh>/smo*]
[king **a pipe / next to** the fire]
- adi/m [alice is **sitting there brooding about**] [/ this man / bein]g chased by the police
- adi/m [that* that pain][ting is sort of **lying around the police station**]
- den/c sees Sylvester [is / **peeking around**] lookie for Tweetie Bird
- den/c the final scene / * involves Sylvester [/ tightrope **walking** /][the* <uh>
- den/c the trolley car / is <uh> # riding alo[ng / **literally chasing after Syl**][vester
- cur/v [a fisher-**price person is walking around**] [on a roll of masking tape
- cur/v <uh> it looked like a small gr[een plane **was flying around a tree**]
- cur/v [a fisher -price **cow is wagging it's**][tail
- cur/v # a co][w is being **moved in the back** of a truck]

2. *Durative-marked speech samples:*

- viv/c [and / **as he's coming up** [and the bowling ball's coming down]]
- viv/c he comes out the bottom of the drai[npipe and he's **got this big bowling ball**
inside h][im
- viv/c inside h[im and **he rolls on dow**]n into a bowling alley
- viv/c <and uh> / h[e goes crawling up the drai]npipe again
- lau/c the canary # throws* # puts a # bowling ball # into # [[the **drain spout as the**
cat is climbing up /]
- lau/c [/ and they go # riding down the street] giving the cat # electric shocks
- lau/c <um> / the / weight / [follows **him** / #]
- san/c and so [# he's walking along] [and all of sudden the trolley comes] along
- san/c [you know and they **keep goin**]g
- san/c [and they* / **they just keep** going on] like that #

Duratives (continued):

- jan/c go **up inside the rain gutter** / [you know][/**barrel**][**ing up** through it]
- jan/c and then it just sort of ends [with them **going** down the street] while / Sylvester gets / electrocuted #
- cel/c hits him on the head and he winds up **rolling down the street** et
- cel/c bottom of the rain barrel and **rolls down the street** [#
- cel/c # yes it / [makes him all **look* # he bounces around** <heh>]
- cel/c [and the monkey's **walking around going # heeheeheehee** and the whole thing]
- cel/c and sh[e*/he **holds out the can** [and she **drops*/the penny in the*/]** in the little / cup
- adi/m #<um> h* but he sort of like as he storms outta the date he [['s kind of **like / feels bad**] ['cause he's really in love with a]]lice #
- adi/m [but while the chase is **going on**]n alice is sitting there brooding
- adi/m # there's a long chase scene a[nd **while the chase / is going on** /]
- adi/m [[/ but **while all this is going on** /][# scotl*]] the other detectives of scotland yard / are / talking to the landlady right?
- adi/m # and/ <uh> scotland yard / **goes through their records**] / and find th* the match
- den/c [down the drainpipe, **rolls down the street into the bowl**]ing alley. Strike. #
- den/c goes up into the apart[ment **making like the monkey**] #
- den/c chasing after Syl][vester and every **time it meets up with him as he's [tight*] /as he's tight-rope walking** #] there's a* / a shock / ok

3. *Perfective-marked speech samples:*

- viv/c inside him and he rolls on down into a **bowling alley**
- viv/c and as soon he gets outside the building # h[e **opens it up**] #
- viv/c [n immediately **he's thrown**] # out
- viv/c and then granny comes and [**throws him** off the] window sill

Perfectives (continued):

- viv/c [he **lands on the ground**]
- lau/c and he just [swings **smack into the** building]
- lau/c [and **that's the end**]
- lau/c and he [gets **kicked**] out /
- lau/c gets the bird c][age he **throws out**] [the **suitcase**]
- lau/c and # goes back the down* / the* / the back staircase # ou[t / **into** the alley]
- san/c and the doorman # [**throws** him out] into a bunch of garbage #
- san/c and Sylvester I guess [**swallows**] [it
- san/c [and **you hear**] [all the pins knocked down #
- san/c [and he **gets into** the apartment] #
- jan/c the first thing he tries [is to **just go in**] the ^front ^door of the hotel #
- jan/c but [the manager of] the ho[tel **throws him out**]
- jan/c [**and ^kicks** him out] in the street #
- jan/c [she **slugs** him]
- jan/c [and **he's thrown out ^into the street** again] #
- cel/c [you assume that he swallows / this bowling ball]
- cel/c and / when the [grandmother sort of **catches on**] to him
- cel/c [and she drops the * / the penny in the*/] in the little / cup
- adi/m she # <uh> [/ sort o][f / blows him off /]
- adi/m s[o <um> she goes out with this man] #
- adi/m and [frank **gets really upset and**] [**storms ou**][t of / the date
- adi/m the bum gets [**real**]ly scared as[d dashes away / you know]
- adi/m he climbs u][p to the roof /][and / #] [/ **falls** through* /]
- den/c <uh> th[e bowling ball **fal**][ls into Sylvester's mouth
- den/c goes out into the h[allway **dumps**] [the suitcase /

Perfectives (continued):

den/c [and that's **the end** of the cartoon]

den/c Sylvester] [gets **a bell**][boy uniform

mar/m / you had this scene i[n which the police / **broke** into his ho]use

mar/m the police / broke into his ho]use / <um> and [**arrested** hi]m

mar/m # after this initial arrest goi[ng **out**] / to # <uh> meet / <uh> this woman?

mar/m things didn't work out # so / sh[e'll settle **for** /]this guy / ok

mar/m gives her the money for the bill and [/ just kind of / **leaves**]

cur/v a red cylinder /] [**fell off**] [a swing

cur/v [a yellow plank **fell** / flat]

cur/v the / rear / wing / of a plane [**fell off** /]

cur/v a telephone book [**fell flat** /]

cur/v it did tumbles # in a backwards dia* diagonal direction % [and then **stopped** /]

mik/c and <uh> get [/ **across**] the road into the other building

mik/c get / across [the road **into** the] other building

mik/c [and **he has sort of** /] climbed up / the* you know {unint.} u* up through the thing

mik/c and* and <uh> # he sees the* the cat there [who has **climbed u**]p and <uh>

mik/c the bird goes into the room /][**flie**][s away

jus/v I guess it's like a window in the roof / [sort **of opened**] up

jus/v a ruler was standing up and then [/ **fell o**]ver

jus/v and / the one in front sort of [/ **hopped up on**] to the other one's back

jus/v and the [wreath / **sort of** / fell] off

APPENDIX B

UTTERANCES CONTRIBUTED BY EACH SPEAKER TO EACH TARGET ASPECT CATEGORY

Mandarin Speakers:

	cy	hhf	hhx	lj	llm	ml	mw	wc	ww	zp	zwh	lw	czx	xmh
PRG	1	3	1	-	-	1	-	3	2	-	5	-	5	4
DUR	2	4	-	2	4	2	3	4	2	-	1	-	2	-
PRF	5	5	-	5	5	5	5	2	3	5	4	2	4	-

Total = 100 utterance-gesture pairs

English Speakers:

	cel	den	jan	lau	mik	san	viv	adi	cur	mar	jus
PRG	3	3	4	1	-	3	4	5	5	-	-
DUR	5	-	2	3	-	3	4	5	-	-	-
PRF	3	3	5	5	4	4	5	5	5	5	4

Total = 100 utterance-gesture pairs

APPENDIX C

SENTENCE STRUCTURES

Key to the transcription:

Grammatical features:

OBJ	object-fronting particle (“ba”/把)
BA	topic-marking particle/discourse particle (“ba”/吧)
NE	topic-marking particle/discourse particle (“ne”/呢)
MA	topic-marking particle/discourse particle (“ma”/嘛)
-CL	noun classifier
-MOD	adjectival, possessive, or adverbial marker (“de”/的 or 得)
PRG	progressive aspect marker (“zai”/在)
DUR	durative aspect marker (“zhe”/著)
PRF	perfective aspect marker (“le”/了)
CMPL	completive verb particle
ONOM	onomatopoeia
PAS	“passive” marker (“bei”/被)

Speech features:

*	self-interrupt
/	unfilled pause
< ... >	filled pause
#	breath pause
%	non-speech sound (laugh unless otherwise indicated)

Gesture features:

[...]	extent of gesture
bold	stroke phase
<u>underline</u>	pre- or post-stroke hold
^	superimposed beat

Mandarin Sentence Structure Sample

Topic/comment structures:

hhf	ta jiu # [pao lai- bu-ji <u>ma jiu di</u>][an-che <u>peng-yi-x</u>][ia
	他就 # 跑來不及嘛 就電車碰一下
	he just # run come-not-reach MA just electric-train bump-one-down
	then he # runs doesn't make it you know, and the train bumps (him)

Topic-comments (continued):

- hhf ran-hou nei-ge* %la[ugh nei-ge **lao tai-tai jing-ran**] [ta bu shi kai-me]n
 然後那個* % 那個老太太驚然他不是開門
 after that-CL* % that-CL old lady unexpectedly she not be open-door
 then that * % unexpectedly, the old lady doesn't open the door
- ml [/ yi-**zhi** ma]o / [du-**zi** hen e / xiang zhao dong-xi chi /]
 一隻貓 / 肚子很餓 / 想找東西吃 /
 / one-CL cat / stomach very hungry / want look-for thing eat /
 / a cat / (he's) very hungry / wants to look for something to eat /
- ml [**nei-zhi m*** /] mao [HO / / yong xiang le zhei-ge] miao ji #
 那隻貓* / 貓 HO // 用想了這個妙計
 that-CL c* / cat HO / again think PRF this-CL excellent plan #
 that c* / cat HO / thinks of another excellent plan #
- zwh ran-hou zhei-ge / [mao ta jiu xiang **qu** /]
 然後這個 / 貓他就想去 /
 next this-CL / cat he just want go /
 then the / cat, he plans to go /
- zwh ran-hou ta jiu* zhei-ge # [mao ne jiu **qu zhui** zhei-ge] niao
 然後他就* 這個貓呢就去追這個鳥
 next he just* this-CL # cat NE just go chase this-CL bird
 then he, the # cat, chases the bird
- zwh [[zhei-ge re][n ne ba zhe][i-ge ta* / ta-de]][xiang-**zi** song-shang-qu
 這個人呢把這個他* / 他的箱子送上去
 this-CL person NE OBJ this-CL he* / his-MOD case deliver-up-go
 the guy sends his suitcase up
- llm ra][n-hou nei-ge **mao ne / jiu**] [**yi-tia**][n-**dao-wan**]
 然後那個貓呢 / 就一天到晚
 next that-CL cat NE / just one-day-until-late
 then the cat, all day long
- [jiu na-**zhe wang-yuan-jin***] / -jing si-chu kan
 就拿著望遠鏡* / 鏡四處看
 just hold-DUR telescop* / scope everywhere look
 looks all over through binoculars
- llm [ta-**men-de men a** / sh][ang-**mian** you yi-ge tian-chuang % nei-ge* nei-ge nu
 他們的門呵 / 上面有一個天窗 % 那個* 那個女
 their door A / top-side have one-CL sky-window % that-CL* that-CL woman
 their door y'know, above it there is a transom % the* the (cat's) lady

Topic-comments (continued):

- zhu-ren] jiu cong tian-chuang shang-mian # jiu chu-lai
 主人就從天窗上面 # 就出來
 owner just from sky-window top-side # just exit-come
 owner then comes out from up there
- llm zhu [zai **ta jia dui-mian** ne]
 住在他家對面呢
 live at his house opposite-side NE
 living opposite his house
- [# you yi-* **you yi**][-ge **long-zi li-mi**][an **guan-zhe yi-zhi mao** /]
 # 有一* 有一個籠子裡面管著一隻貓 /
 # have one* have one-CL cage inside contain-DUR one-CL cat
 there's a cat kept in a cage
- zp nei-ge hou-zi ne ran]-^hou lao [tai]-tai [you gei na **yu-san**] da le yi-xia
 那個猴子呢然後老太太又給拿雨傘打了一下
 that-CL monkey NE next old lady again give pick-up umbrella hit PRF once
 the monkey, the old lady picks up an umbrella and whacks him
- zp zhei-ge sh[i-hou ne zhei-zhi **mao** ne] [jiu zhuang-ban-**cheng** ne]i-ge guan-li-yuan
 這個時候呢這隻貓呢就裝扮成那個官吏員
 this-CL time NE this-CL cat NE just disguise into that-CL service person
 this time the cat disguises himself as the deskclerk
- zp ta [ba niao **long-zi n**][e **nei-ge da-kai yi-kan** ne] shi yi-ge lao*
 他把鳥籠子呢那個打開一看呢是一個老*
 he OBJ bird cage NE that-CL open ASP-upon NE be one-CL old*
 when he opens the bird cage and looks, it's an old*
- shi nei-ge lao tai-tai
 是那個老太太
 be that-CL old lady
 it's the old lady
- zp dan-shi ne zhei-shi ta [**zou-yi-zou**] [ne hou-mian lai] le -li[ang **wu-gui dian-che**]
 但是呢這時他走一走呢後面來了一輛無軌電車
 however NE this-time he walk-one-walk NE behind-side PRF -CL no-track train
 but now as he's walking a trolley comes from behind
- wc [nei dian-che / **ta nei dian-xian bu he ta nei-ge** /]
 那電車 / 他那電線不合他那個 /
 that train / it that wire not connect it that-CL /
 the train, don't the wires connect to it /

Topic-comments (continued):

[bu-yao zhei-yang cong **dian-xian-shang** zou-g]uo ma
 不要這樣從電線上走過嗎
 not-will this way from wire-top walk-past MA
 he goes via the top of the wires like this y'know

mw [zhei-ge / niao long-zi / ta shi ^fang zai nei-ge chuang-tai shang*
 這個 / 鳥籠子 / 他是放在那個窗臺上*
 this-CL / bird cage / it be put one that-CL window sill top*
 the / bird cage / it's positioned on the window sill*

zhei-ge liang-tai shang #
 這個亮臺上 #
 this-CL balcony on #
 on the balcony #

mw [nei-ge **xiao** niao ne] [/] tao-zou le
 那個小鳥呢 / 逃走了
 that-CL little bird NE [/] flee-go PRF
 the little bird [/] flees

mw [jie-guo zui-hou **da-kai** yi-kan NE] [shi **lao tai-tai** zai li-tou]
 結果最候打開一看呢是老太太在裡頭
 result at end open ASP-upon look [P] is old lady in inside
 so when he opens it and looks, it's the old lady inside

lj nei-ge xiao niao % ne ye ^na yi-ge wang-yuan-jing ^/ [jiu || kan-d]ao ta
 那個小鳥 % 呢也拿一個望遠鏡 / 就看到他
 that-CL little bird % NE also hold one-CL telescope / then see to him
 The little bird % is also watching through binoculars / and sees the cat.

lj ta / nei [**niao jiu fei**] / <nn> fei-zou le
 他 / 那鳥就飛 / <nn> 飛走了
 he / that bird just fly / <nn> flies away
 then the bird flies away

Object-fronted structures — 'ba' (把):

hhf ran-h[ou ta jiu ba nei-ge # / **da-kai***]
 然後他就把那個 # / 打開
 after he just OBJ that-CL # open*
 then he opens it

Object-fronted (continued):

- hhf ran-ho[u ta jiu **ba nei-ge**] [hou][-zi de yi-fu **ba**][-xia-lai]
 然後他就把那個猴子的衣服剝下來
 then he just OBJ that-CL monkey MOD clothes peel-down-come
 so then he peels the monkey's clothes off
- cy ta jiu [ba nei **qiu** cong nei] guan-zi li sai-jin-qu le
 他就把那球從那管子裡塞晉去了
 he then OBJ that ball from that pipe inside squeeze-in-go PRF
 he squeezes the ball into the pipe
- cy ta shuo zhei* / zhei yi-ge xiang-jiao[**ba nei hou yin-dao nei-ge** / qiang] guai-jiao
 他說這* / 這一個香蕉把那猴音到那個 / 牆拐角
 he say this* / this one-CL banana OBJ that monkey lure-to that-CL wall corner
 he says* / this is a banana to lure the monkey to a corner of the wall
- ml [jie-guo **ta ba ta**] [**quan-bu na** /// <eehh> ba ta] ...
 結果他把他全部拿 /// <eehh> 把他
 result he OBJ it all take /// <eehh> OBJ it ...
 so then he takes everything /// <eehh> and ...
- ml dai-zi jiu diu le / jiu b[a / nei-z]hi niao* niao [**long** / na-dao p]ang-bian
 帶子就丟了 / 就把 / 那隻鳥* 鳥籠 / 拿到旁邊
 suitcase just throw PRF / just OBJ / that-CL bird* bird cage / take-to side-place
 (he) throws the suitcase out, then (takes) the bird* bird cage / off to the side
- zwh <uh> kan-dao ta le jiu ba ta gei <e[h> da-**chu-lai** le #]
 <uh> 看到他就把他給 <eh> 打出來了 #
 <uh> see-to him PRF just OBJ him give <eh> hit-out-come PRF
 <uh> (she) sees him and knocks him away
- zwh ran-hou jiu na yi-ge [xiang-**jiao ba nei-ge** hou-z]i % # yin-guo-qu
 然後就拿一個香蕉把那個猴子 % # 引過去
 next just pick-up one-CL banana OBJ that-CL monkey % lure-across-go
 the (he) picks up a banana and lures the monkey across (the road)
- llm jiu yong yi-g[e hen **zhong-de dong-x**][i ba ta **diu-guo-qu** #]
 就用一個很重的東西把他丟過去 #
 just use one-CL very heavy-MOD thing OBJ it throw-past-go
 then he throws a heavy object over there
- llm ta jiu na le yi-gen xiang-jia[o qu ba nei-ge **hou-zi**] [yin-guo-qu_]
 他就拿了一根香蕉去把那個猴子引過去 /
 he just pick-up one-CL banana go OBJ that-CL monkey lure-across-go
 then he lures the monkey across with a banana

Object-fronted (continued):

- zp jiu ba [nei xiang-**jiao** gei ta] chi
 就把那香蕉給他吃
 just OBJ that banana give him eat
 then gives him the banana to eat
- zp [ta **jiu ba n**][ei-ge xiao hai gei **cang-q**]-i-lai le
 他就把那小還給藏起來了
 he just OBJ that-CL little child give hide-rise-come PRF
 then she hides the child (in something)
- wc sh[ang-**mian** nei-ge jiu ba nei-ge*] jiu [**gei du**-shang le
 上面那個就把那個* 就給堵上了
 top side that-CL just OBJ that-CL* just give stop-up PRF
 the thing on top then stops up* stops it up
- wc [ta jiu ba nei **hou**-zi gei] zhua-lai le
 他就把那猴子給抓來了
 he just OBJ that monkey give catch-come PRF
 then he captures the monkey
- ww ran-hou # t[a jiu **ba niao** l][**ong-zi**] [na-zou]
 然後 # 他就把鳥籠子拿走
 next # he just OBJ bird cage take-go
 then he takes the bird cage away
- ww [/ ran-**ho**][u **ta ba ta da**-kai-de shi-h][ou #
 / 然後他把他打開得時候 #
 next he OBJ it open-MOD time #
 then when he opens it #
- mw xiang ba zhei-ge lao tai-tai jia-de [nei-ge **xiao niao**] <ehn> [qiang-**dao** shou]
 想把這個老太太家的那個小鳥 <ehn> 搶到手
 want OBJ this-CL old lady house-MOD that-CL small bird <ehn> snatch-to hand
 (he) wants to get hold of the bird that is in the old lady's house (get hold of)
- mw [yi-xia **jiu ba zhei-ge** /] [shuai dao **zui-li** qu le /]
 一下就把這個 / 摔到嘴裡去了
 one-down just OBJ this-CL / fall to mouth-in go PRF /
 then at once it plops into (his) mouth
- lj [ba nei **bu**] [**jie*** jiu jie-qi-lai]
 把那布結* 就結起來
 OBJ that cloth tie* just tie-rise-come
 then ties* ties up the cloth

SVO Structures:

- lj nei-bian jiu diao-xia-q[u le PAH **jiu ba ta** tan-qi-lai] le %
 那邊就掉下去了 PAH 就把他彈起來了 %
 that-side just fall-down-go PRF ONOM just OBJ him catapult-rise-come PRF
 that side falls. PAH then catapults him up %
- hhf nei-zhi [jin-si-que **ye na-z**][**he yi-zhi x^iao xi^ao-de wang-yuan-^jing**]
 那隻金絲雀也拿著一支小小的望遠鏡
 that-CL canary also hold-DUR one-CL small small-MOD binoculars
 the canary is also holding a pair of very small binoculars
- hhf jiu / zhan zai nei-ge / jie-de zhuan-ji[ao ta **jiu n**][a-chu **yi-ge xiang-jiao a /**]
 就 / 站在那個 / 街的轉角他就拿出一個香蕉呵 /
 just / stand at that-CL / street-MOD corner he just take out one-CL banana A /
 so he just / stands on the / street corner then he takes out a banana, y'know? /
- cy ran-hou [zi-ji **chuan shang nei-ge yi-fu**]
 然後自己穿上那個衣服
 then self dress on that-CL clothes
 and puts the clothes on.
- cy [ta jiu **jian**] [le yi-ge te-bie **zhong-de shen-ma dong-xi a**]
 他就建了一個特別重的甚麼東西呵
 he then puts PRF one-CL extremely heavy-MOD what thing A
 then he puts an especially heavy thing
- zwh [ke-neng shi **xiang ch**][i nei-ge niao <huhn> zen-ma jia]ng /
 可能是想吃那個鳥 <huhn> 怎麼講 /
 maybe be think eat that-CL bird <huhn> how say
 maybe (he) is thinking about eating the cat or something
- llm [# ran-hou ta jiu **zhua dao nei-zhi xiao niao le #**]
 # 然後他就抓到那隻小鳥了 #
 # next he just snatch-CMPL that-CL little bird PRF #
 then he captures the little bird
- llm [/ ta ke-yi **pa**] [nei-ge / **shui-guan a /**]
 / 他可以爬那個 / 水管呵 /
 / he can climb that-CL / water-pipe A /
 he can climb the / drainpipe y'know
- zp [you hou-zi lai* lai **da-rao ta**]
 有猴子來* 來打擾她
 have monkey come* come bother her
 there's a monkey bothering her

SVO (continued):

- zp dan-shi [ta yi **kai-kai men yi**]-hou mei-you ren
但是她一開開們以候沒有人
however he ASP-upon open-open door after not-have person
but when she opens the door, there's no one there
- wc lao tai-tai [na -ge da **bang**] hao-xiang gei ta da-xia
老太太拿 個大棒好象給他打下
old lady pick-up -CL big board appear give him hit-down
the old lady picks up a big board and knocks him down
- wc bu [dui shi zhei-**ge xiao niao**] [na-zhe -ge **da yuan qiu**]
不對是這個小鳥拿著 個大圓球
not correct be this-CL little bird hold-DUR -CL big round ball
wrong, the little bird picks up a big round ball
- wc ta jiu chuan [yi-ge **lao tou-de nei-ge* nei hei-se-de**] yi-fu
他就穿一個老頭兒的那個* 那黑色的一服
he then put on one-CL old head-MOD that-CL* that black-MOD clothes
then he puts on a black suit like an old guy would wear
- ww [/ ran-ho **ta** chuan xi][ao **hou-zi-de**] / yi-fu #
/ 然後他穿小猴子的 / 衣服 #
next he wear little monkey-MOD / clothes #
then he puts on the little monkey's / clothes #
- ww ran-ho[u lao-tai-**po jiu da** ta-de t]ou
然後老太婆就打他的頭
next old lady just hit his-POS head
then the old lady hits his head
- mw [ta jiu na le **yi-ge da tie tuo** /]
他就拿了一個大鐵 tuo /
it then bring PRF one-CL big iron top /
it brings a big iron top /
- mw [ta **jiu she-ji nei-ge shen-du** #]
他就設計那個深度 #
he then calculate that-CL depth #
he then calculates the distance [between the buildings] #
- mw [ta jiu **chuan-shang nei-ge yi-fu** ma]
他就穿上那個衣服嗎
he then wear-on that-CL clothes MA
then he puts on the clothes

SVO (continued):

- lj # nei-ge xiao niao jiu na [yi-**ge da bao-ling** qi]u
 # 那個小鳥就拿一個大保玲球
 # that-CL little bird then picks up one-CL big bowling ball
 # Then the little bird picks up a big bowling ball,
- lj ran-hou # ta* / [ta na || yi xiang]-jiao HAH
 然後 # 他* / 他拿一香蕉 ONOM
 then # he* / he holds one banana ONOM
 Then # he* / he holds out a banana HAH
- lj ran-hou ta jiu cong qian-bao [li-bian **na**-chu yi-fen qian lai]
 然後他就從錢包裏邊拿出一分錢來
 then she then from wallet inside take out one penny money
 Then she takes out a penny from her wallet

SV Structures:

- hhf/c [ta jiu **zuan**-jin-q]u
 他就鑽晉去
 he just bore-enter-go
 he just bores into it
- hhf [# ta jiu / **gun gun gun gun gun** g][un
 # 他就 / 滾滾滾滾滾滾
 # he just roll roll roll roll roll roll
 # he just rolls and rolls
- cy ta jiu shun-[zhe nei guan-zi pa]-shang-qu le #
 他就順著那管子爬上去了 #
 it then along-DUR that pipe climb upwards PRF #
 So it climbs up along the pipe #
- cy <ehn> nei-ge [mao **zhuang-de** hou-zi j][iu **jin-qu** /]
 <ehn> 那個貓狀得猴子就晉去 /
 <ehn> that-CL cat disguise-MOD monkey then enter /
 <ehn> the cat disguised as a monkey goes into [the room]
- ml [/ nei-**zhi niao** #][jiu # wang shang-mian* / wang shang-mian qu le /]
 / 那隻鳥 # 就 # 往上面* / 往上面去了 /
 / that-CL bird # then # toward top-side* / toward top-side go PRF /
 / then the bird upward* / goes upward /

SV (continued):

- ml peng-qia[o / **nei-yi-kuai shi-tou** #][gang hao you **luo** zai ta-de shen]-ti
碰巧 / 那一塊石頭 # 剛好有落在他的身體
coincident / that-one-CL rock #] [just then again land on his body
coincidentally, the rock # right at that moment falls on top of him
- zwh [/ ran-hou jiu / **pao dao nei-ge**] lou-xia
/ 然後就 / 跑到那個樓下
/ next (he) just / run to that-CL building-bottom
then he runs to the base of the building
- zwh n[a ta jiu shun **nei-ge shang-mian pa-shang-qu**]
那他就順那個上面爬上去
NA he just along that-CL top-side climb-up-go
now, he climbs up along the surface of it
- zwh jie-guo / [/ yi **jin yi** lou] /
結果 / 一晉一樓 /
result / ASP-upon enter one floor /
then, upon entering the first floor /
- llm [ta jiu **tan-shang-qu** /]
他就談上去 /
he just spring-up-go /
then he catapults up /
- llm # ta ji[u **pao-diao** le]
他就跑掉了
he just run-away PRF
then he runs away
- zp [ta xiang cong nei-ge men li-mian # **zuan-shang-qu**]
他想從那個們里面 # 鑽上去
he want from that-CL door inside # barrel-up-go
he wants to barrel up from inside that door
- zp [ran-hou hou-zi **jiu pao-diao** le]
然後猴子就跑掉了
next monkey just run-away PRF
then the monkey runs off
- zp [jiu-shi fang-**zi guan-li-yu**][an **shang-qu**]
就是房子官吏員上去
just-be building clerk up-go
the desk clerk goes up

SV (continued):

- wc ji-[bian le GRR ta **jiu**][pao-shang-qu le]
 擠扁了 GRR 他就跑上去了
 squeeze-flat PRF ONOM he just run-up-go PRF
 then he runs up squeezed flat
- mw ta jiu pa* [pa-shang-lai] #
 他就爬* 爬上來 #
 he then climb* climb-up-come#
 he then climbs* climbs up#
- mw [ta **jiu** pao le] /
 他就跑了 /
 he then run PRF /
 then he runs away /
- mw
- [da-gai jiu **yi-zhi gen-zhe zhei-ge dian wang jiu** * jiu wang qian zou]
 大概就一直跟著這個電網就* 就往前走
 probably just straight with-DUR this-CL electric net just * just toward forward go
 then he probably goes walking along the electrified net* walking forward
- lj ta [yan nei-glle p][a pa pa] nei-ge pa-shang-qu #
 他沿那個爬爬爬那個爬上去 #
 he along that-CL climb climb climb that-CL climb-up-go #
 he climbs and climbs up along that
- lj ta kai-shi cong nei-ge / x[ia-shui-**dao** nei-ge wai-mian][**pa**-shang-qu] de
 他開始從那個 / 下水到那個外面爬上去得
 he begins from that-CL / drainpipe that-CL outside climb-up-go MOD
 he starts climbing up the outside of that / drainpipe,

English Sentence Structures Sample

SVO Structures:

- jan but [the manag]er of the ho[tel **throws him out**]
- jan I dunno [she **slugs** him] or throws him out the window
- jan and he drops a [**bow**]ling ball [**into the rain spout**]
- jan [so he **lures** the guy's monkey away] #
- jan [and **Sylvester**] [**puts on** /] [**the / little** ^monkey suit]

SVO (continued)

jan and he takes [/ a **suitcase**]
 lau he tries / [clim][bing up the / drain spout / of the building #
 lau she [**chases**] [**him out** again]
 lau he throws **out**] [the **suitcase**]
 lau [and he **gets the bird cage** /]
 lau [and when he uncovers # the birdcage] #
 lau <um> [/ the / weight / follows **him** / #]
 viv Tweetie Bird runs and gets a bowling b[all and drops **it down** the drainpipe] #
 viv [he swallows it] #
 viv [and that / / catapults him up] #
 viv and [he **grabs** Tweetie Bird #
 viv he [**opens it up**] #
 viv well [he entices the monkey] away from the organ grinder with a banana
 san and the doorman # [**throws** him out] into a bunch of garbage #
 san [the **grandma** # **hits him** over the head] with the umbrella #
 san [Sylveste][r I guess swal][lows it
 san and he t][hrows **the ball into the top**] #
 san [and he takes them out] #
 san [and # takes the birdcage **down some stairs**] out to the alley
 cel and [drops a bowling bal][**l down the rain b**][arrel
 cel [so it **hits him on**] [the head
 cel so [he entices the monkey around the corner with] a banana
 cel [/ he **holds out the can**]
 cel [and he / takes off his hat to her]

SVO(continued):

- den <uh> the lady goes what a cute monkey [gives the* / gives Sylvester] a penny #
- den Sylvester [gets **a bell**][boy uniform
- den Sylvester [**gets the bags**]
- den goes out into the h[allway **dumps**] the suitcase
- den [takes the **cover off** the] bird cage
- den the little old lady's sitting there in the bird c[age / chas]es him down the street with the umbrella /
- mik then <um> the next thing he tries to do if I remember rightly # is to [/ **climb up the drainpipe** /] to get to the bird /
- mik [gets a bowling ball and / **throws it down the*** /] the* the drainpipe /
- mik [# the cat] [**tempts the monkey away with a banana / right?**]
- mik [an' then he **climbs up** / the / drain]pipe again
- mik so he [/ goes in 'n' **picks up the* / the* <um> / the* the cage**]

SV Structures:

- viv [and he **comes out the bottom of the drai]npipe**
- viv [and **he rolls on down into a [bowling all]]ey**
- viv and a[t this time the five hundred pound weight comes down **and lands on him**]
- viv [he comes **swinging through** on a rope] /
- viv so [he goes **scrambling up**] to <um> # the room
- viv and / [he* so **he's running** along] the wire
- lau <and uh> / he tries # just [walking **in**] the front door of the hotel #
- lau <and uh> # he [# **climbs** into the room]
- lau and [it **goes into his**] mouth /
- lau [and he rolls # **down the drain** spout]

SV (continued):

lau [he / goes **up** / **to** the window] and he grabs the bird

lau [and then when he comes **down** again #]

san 'n so Sylvester # runs [across **the street** into] the apartment #

san and # [he goes **up through** the pipe] this time #

san [and # **it goes** into Sylvester] #

san [so **he goes in** /]

san [n' # you see him swinging down **across the** rope] /

san and so [# so **he's walking** along]

jan he first thing he tries is [to **just go in**] the [front door o]f the hotel #

jan so / he [go* / th* / he goes back down /]

jan [this time he tries to go **up inside the rain gutter** /]

jan [and it **goes down**]

jan [he's trying to walk across those] to get into the window

cel [and he winds up **rolling down the stre**]et

cel [# **and he comes rolling out of the / bottom of the rain bar**][rel

cel [so he has **to climb in the window somehow** #]

cel <um> / [trying to swing **across by** a rope #]

cel [so he tries **to tightrope walk across**] [**to the** window] #

cel [and here comes **one of the** trolleys]

den <uh> th[e bowling ball **fal**]ls into Sylvester's mouth

den Sylvester **falls back**] down the drainpipe

den [**goes up**] [**into the apartment**]

den the final scene / # involves Sylvester [/ **tightrope walking** /]

mik [above with the bo* **the bowling ba**][ll going down]

mik an' [then* then you just see* you just see **it collide BANG like this**.]

mik [# an' the* **an' the*** the cat comes right down the bottom #]

mik [# so the bird **flies into** th]e* / into the / apartment again #]

mik [OK? so **he's tippy-toei**][ng along **the tram** wires #]

mik [so he's **sort of running along** an'] everything

APPENDIX D

UTTERANCES CONTRIBUTED BY EACH SPEAKER TO EACH TARGET UTTERANCE CATEGORY

Mandarin Speakers:

	cy	hhf	lj	llm	ml	mw	wc	ww	zp	zwh	Total
SVO	2	2	3	2	-	3	3	2	2	1	20
SV	2	2	2	2	2	3	1	-	3	3	20
T-C	-	2	2	3	2	3	1	-	4	3	20
BA	2	2	2	2	2	2	2	2	2	2	20
Total	6	8	9	9	6	11	7	4	11	9	80

English Speakers:

	cel	den	jan	lau	mik	san	viv	Total
SVO	5	6	6	6	5	6	6	40
SV	6	5	5	6	6	6	6	40
Total	11	11	11	11	11	12	12	80

APPENDIX E

NUMBER OF UTTERANCES WITH SPECIFIC GESTURAL ACCOMPANIMENTS: SINGLE VERSUS MULTIPLE GESTURES; ISOLATING VERSUS INCORPORATING *

<u>Single Gesture Utterances:</u>				<u>Multi-gesture Utterances:</u>			
<u>Isolating</u>		<u>Incorporating</u>					
Mandarin:							
SV	9	action	3	action+figure/ground	4	action-figure **	
			1	enactment	2	action-ground	
					1	action+figure-ground	
SVO	5	patient	5	enactment	3	enactment-patient	
		3			action	1	action-patient
						1	enactment-enactment
					1	agent-patient	
					1	patient-patient	
T-C	3	action			2	agent-action	
		1			patient	1	action-patient
		1			ground	1	enactment-patient
						1	action-figure
						1	action-result
						1	figure-figure
						1	figure1-figure2
						1	action+ground-ground
						1	action+ground+figure- action+ground
						1	action-figure-location
						1	enactment+patient-agent- enactment
						1	enactment+ground-agent- enactment
						1	enactment-ground-agent- enactment
						1	action-agent-patient-patient

Appendix E (continued). Number of utterances with specific gestural accompaniments: single versus multiple gestures; Isolating versus Incorporating.

		<u>Single Gesture Utterances:</u>		<u>Multi-gesture Utterances:</u>		
		Isolating	Incorporating			
Mandarin:						
BA	4	action	3	enactment	3	enactment-patient
			1	enactment+action	2	action-patient
					1	action-enactment
					1	action+patient+ground- patient+ground
					1	enactment-enactment
					1	enactment-enactment+action
					1	enactment+action-patient
					1	enactment-action-patient
					1	action-patient-patient
English:						
SV	10	action	14	action+figure	2	action+figure-endpoint/ground **
			4	enactment	1	action-action
			2	enactment+action+figure	1	action-figure
			1	enactment+action	1	action+figure-figure
			1	action+location	1	action-location
			1	action+recipient+ground	1	enactment-endpoint
SVO	6	action	23	enactment	2	enactment-agent
			3	action+agent	1	action-patient
			3	action+patient+ agent/ground	1	action+patient-agent
					1	enactment-action+enactment
<hr/>						
* The ordering of components in sequence is not necessarily the order in which they occurred in production.						
** Key: '+' indicates gestural incorporation of multiple event components in non-enactment gestures; '-' is separation of two gestures; '/' means two components are indistinguishable in one gesture form.						