



Viewpoint in Language

What makes us talk about *viewpoint* and *perspective* in linguistic analyses and in literary texts, as well as in landscape art? Is this shared vocabulary marking real connections between the disparate phenomena? This volume argues that human cognition is not only rooted in the human body, but also inherently “viewpointed” as a result; consequently, so are language and communication. Dancygier and Sweetser bring together researchers who do not typically meet on common ground: analysts of narrative and literary style, linguists examining the uses of grammatical forms in signed and spoken languages, and analysts of gesture accompanying speech. Using models developed within cognitive linguistics, the book uncovers surprising functional similarities across various communicative forms, arguing for specific cognitive underpinnings of such correlations. What emerges is a new understanding of the role and structure of viewpoint and a groundbreaking methodology for investigating communicative choices across various modalities and discourse contexts.

BARBARA DANCYGIER is Professor in the Department of English at the University of British Columbia, Vancouver.

EVE SWEETSER is Professor in the Department of Linguistics at the University of California, Berkeley.



Viewpoint in Language

A Multimodal Perspective

Edited by

Barbara Dancygier

and

Eve Sweetser

For Irene,
a traveller in
shared landscape /
and whose perspective
is always a self!
much love,
Eve



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press,
New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107017832

© Cambridge University Press 2012

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Viewpoint in language : a multimodal perspective / [edited by] Barbara Dancygier and
Eve Sweetser.

p. cm.

Includes index.

ISBN 978-1-107-01783-2 (hardback)

1. Language and languages – Style – Psychological aspects. 2. Perspective
(Linguistics) 3. Speech and gesture. 4. Discourse analysis, Literary. 5. Cognitive
grammar. I. Dancygier, Barbara. II. Sweetser, Eve.

P301.5.P75V53 2012

401'.41 – dc23 2012004943

ISBN 978-1-107-01783-2 Hardback

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party internet websites referred to
in this publication, and does not guarantee that any content on such
websites is, or will remain, accurate or appropriate.

Contents

<i>List of contributors</i>	<i>page</i> vii
<i>Preface</i>	ix
<i>Acknowledgments</i>	xii

Introduction: viewpoint and perspective in language and gesture, from the Ground down EVE SWEETSER	1
--	---

Part I Intersubjectivity and subjectification

1 Irony as a viewpoint phenomenon VERA TOBIN AND MICHAEL ISRAEL	25
2 Subjectivity and upwards projection in mental space structure LILIAN FERRARI AND EVE SWEETSER	47
3 Negation, stance verbs, and intersubjectivity BARBARA DANCYGIER	69

Part II Gesture and processing of visual information

4 Interactions between discourse status and viewpoint in co-speech gesture FEY PARRILL	97
5 <i>Maybe what it means is he actually got the spot:</i> physical and cognitive viewpoint in a gesture study SHWETA NARAYAN	113

Part III Multiple viewpoints in American Sign Language

6 Reported speech as an evidentiality strategy in American Sign Language BARBARA SHAFFER	139
--	-----

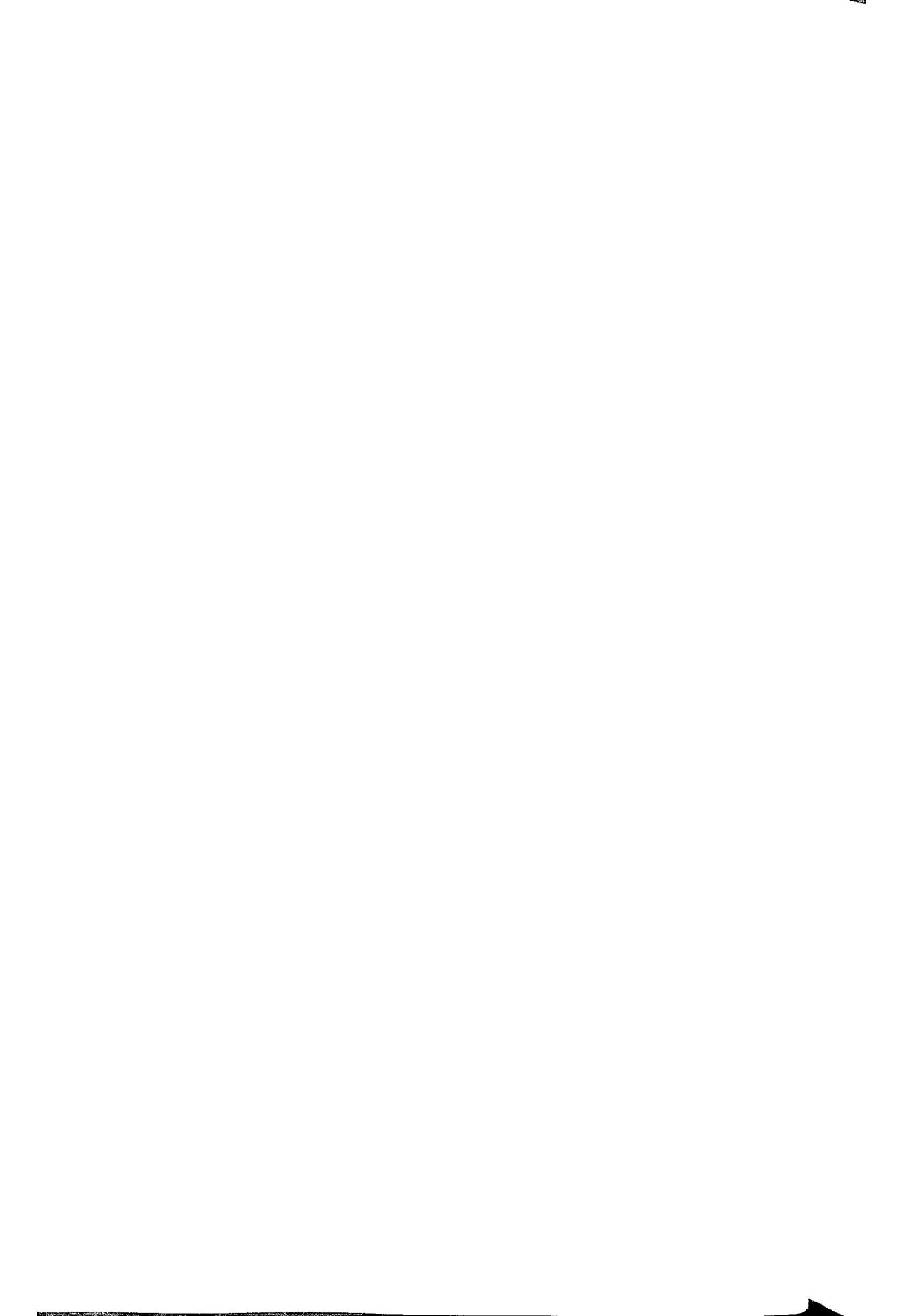
- 7 Two ways of conceptualizing space: motivating the use of static and rotated vantage point space in ASL discourse 156
TERRY JANZEN

Part IV Constructions and discourse

- 8 The constructional underpinnings of viewpoint blends: the *Past + now* in language and literature 177
KIKI NIKIFORIDOU
- 9 “Wait till *you* got started”: how to submerge another’s discourse in your own 198
LIEVEN VANDELANOTTE
- Conclusion: multiple viewpoints, multiple spaces 219
BARBARA DANCYGIER
- Notes* 232
Index 238

Contributors

- BARBARA DANCYGIER, Department of English, University of British Columbia
- LILIAN FERRARI, Department of Linguistics, Federal University of Rio de Janeiro
- MICHAEL ISRAEL, Department of English, University of Maryland
- TERRY JANZEN, Department of Linguistics, University of Manitoba
- KIKI NIKIFORIDOU, Department of English Language and Literature, University of Athens
- SHWETA NARAYAN, Department of Linguistics, University of California, Berkeley
- FEY PARRILL, Department of Cognitive Science, Case Western Reserve University
- BARBARA SHAFFER, Department of Linguistics, University of New Mexico
- EVE SWEETSER, Department of Linguistics, University of California, Berkeley
- VERA TOBIN, English Department, University of California, Santa Barbara
- LIEVEN VANDELANOTTE, University of Namur (FUNDP) and University of Leuven



Preface

We felt we had to bring this volume together, when we realized how many colleagues – with differing methodologies and differing data – were converging on similar perceptions about viewpoint. The first of these perceptions is that viewpoint is far more pervasive in human cognition and language than has been recognized – no matter what the content of our perception, cognition, or linguistic expression, the content is never independent of viewpoint, and viewpoint expression is a crucial and constant job of human communication. Another shared perception is that despite the copious literature on, for example, narrative viewpoint, we are still unearthing new and subtle aspects of the relationships between different viewpoints or complex viewpoints, within discourses – and we need new ways to talk about these relationships. And yet another is the need to return to, and question, the relationship between physical perceptual viewpoint and abstract – for example, narrative – viewpoint.

By bringing together authors from different scholarly communities, we have been able to highlight parallels between viewpoint structures in domains that were not previously thought of as part of one investigation. Knowing that irony, for example, is a particular kind of viewpoint embedding, can help us relate it more precisely to other kinds of viewpoint embedding found in narrative. Noting that American Sign Language narratives, or co-speech gestures, actually use physical viewpoint to represent narrative viewpoint, we may view with new seriousness the place of perceptual viewpoint in “abstract” literary viewpoint. Mental Spaces Theory is a shared framework linking many of the contributions to the volume. The insights of our authors, however, are not dependent on a mental spaces framework; rather the framework serves the need expressed above, of providing clear and general formal expression of many varied, complex viewpoint relationships.

In our first section, on intersubjectivity and subjectification, three chapters present ways in which speakers create complex viewpoint structures by simultaneously evoking a network of contrasting and conflicting spaces. Vera Tobin and Michael Israel present a new analysis of irony, covering a full range of categories, from literary to situational irony. Irony cannot reside in one single mental space or viewpoint; it resides in that final, all-knowing viewer, who “sees

it all” and assesses contrasts, a single subjectivity that incorporates a network of contrasting understandings of a situation. Lilian Ferrari and Eve Sweetser offer a treatment of historical semantic subjectification – for example, deictic markers becoming articles, or deontic modals taking on epistemic meanings – via incorporation of situational inference into conventional lexical semantics. They suggest that the result of such incorporation is “more subjective,” precisely in the sense that the meanings incorporated are located in higher mental spaces, more distant from the space of the content being described. And Barbara Dancygier shows how negation not only evokes alternative spaces (the positive and negative alternatives), but gives very complex, distributed allocation of viewpoint when combined with stance verbs. *I don't think he can't write songs*, for example, may rebut an imagined *interlocutor* who might think that the speaker claims this songwriter cannot write; this is an extremely different meaning from *I think he can write songs*, and far more intersubjectively complex.

Our second section moves into the domain of co-speech gesture, where physical viewpoint crucially represents various kinds of speaker perspective and construal. Fey Parrill shows how Narrator Viewpoint, Observer Viewpoint, and Character Viewpoint are gesturally represented in narratives recounting the story of a cartoon, and develops an understanding of the relationship between the linguistic viewpoint and that manifested in the co-speech gesture. Shweta Narayan offers a view of how interlocutors’ gestures manifest intersubjective sharing of viewpoint; in her case study, alignment of gestures demonstrates cognitive alignment with the interlocutor, while abandonment of that alignment for a new gestural perspective marks new cognitive construal of the content from an independent perspective. Both of these studies help advance gesture studies, by showing the very complex network of different spaces that can be “blended” with the physical gesture space. Speakers and gesturers never have access only to Narrator Viewpoint, or only to the interlocutor’s viewpoint – gesture has meaning only in the context of the complex active network of cognitive construals, just as we saw with irony, lexical change, and narrative viewpoint. In addition, these studies remind us that the rich data offered by speech-accompanying gesture has much to tell us about linguistic as well as cognitive viewpoint.

Our third section offers two new contributions to the understanding of viewpoint in American Sign Language. Barbara Shaffer lays out the mechanisms for marking reported speech in ASL, showing that signed languages also have evidentiality-marking strategies, being grammaticalized, in fact, by mechanisms very similar to those observed in spoken languages. Terry Janzen describes an ASL viewpoint-shifting structure that involves an imagined 180-degree rotation of the signer’s body, so that the signer can alternately physically represent first one interlocutor in a described conversation, and then the other (facing) interlocutor. These chapters document new aspects of ASL grammar, but they also show again the ways in which viewpoint-indicating mechanisms

necessarily evoke relations between multiple construals and views – to negotiate how a single signer represents reported speech, the signer’s body and its surrounding space must be flexibly mapped, in a changing manner, onto different parts of the active network of meaning structure.

In our fourth section, we return to more traditional territory – to literary texts and narrative viewpoint, and also to grammar and grammatical marking of viewpoint. Kiki Nikiforidou discusses the well-known use of the English past tense with a proximal deictic *now* in narrative (e.g. *having spent several years as a graduate student in California . . . she now thought of herself as spiritually an American*). She argues that (1) it indicates a perspective shift to a vantage point inside the narrated event, and (2) that it should be viewed as having characteristics of a high-level discourse-structuring grammatical construction. Lieven Vandelanotte presents evidence for a distinct type of indirect discourse, namely “distancing indirect speech and thought,” which incorporates a quoted speaker’s discourse, while maintaining the viewpoint and deictic center aligned with the quoting speaker – for example, the sarcastic *you was going to do wonders, you was*, where *you* refers to the very addressee whose views are quoted. Like Nikiforidou’s example, this can be seen as a discourse-level construction, which affects diverse grammatical choices, such as pronouns, deictics, or referential expressions. And also like Nikiforidou’s example, it involves blending of different viewpoints, with aspects of the linguistic form marking each perspective.

Our “bookends,” Eve Sweetser’s Introduction and Barbara Dancygier’s Conclusion, tie the collection together in different ways. Sweetser lays out an initial landscape of current approaches to viewpoint – including work in neuroscience, anthropology, and other related areas, as well as the fields represented in this volume. The goal is to give an overview of how new understandings of linguistic and communicative viewpoint can be integrated into both linguistics and the broader study of embodied cognition. Dancygier draws together the strands of the volume by summing up the varied types of relationships between communicative forms and expressed viewpoints that are touched on by our authors. She emphasizes not only the complexity of the cognitive phenomena represented by the deceptively simple linguistic and gestural structures examined here, but also the generalizability of models of mental space relations, such as compression and blending, in understanding the representation of (inter)subjectivity.

We hope that our readers will appreciate how inevitably these varied chapters belong together – a new, interdisciplinary understanding of viewpoint and perspective is emerging from these strands of scholarship, breaking down boundaries between literary and linguistic scholarship, between sign and spoken and written language studies, between gesture studies and language studies. We trust, of course, that some of those readers will carry the field forwards to new syntheses, which we cannot yet see from our current perspective.

Acknowledgments

As editors, we are both grateful to our contributors, who have done a wonderful job of interacting with each other, giving us a variety of viewpoints: they have built up a remarkable scholarly community to a new level. Other colleagues, like Paul Dudis, Gilles Fauconnier, Ron Langacker, Scott Liddell, Cornelia Müller, Rafael Núñez, Esther Pascual, Jo Rubba, Elena Semino, Elizabeth Closs Traugott, Mark Turner, and Arie Verhagen, have crucially shaped both the field and our personal understandings of viewpoint. And, as ever, we both express our ongoing deep appreciation of our families and their many-faceted support.

Barbara further appreciates the varied ways in which her graduate students in Vancouver helped her to shape ideas and link them to various scholarly contexts. Their keen interest and willingness to look beyond their own areas of research have provided a much-needed context for the work represented in this volume. Special thanks also go to Lieven Vandelanotte and José Sanders, whose expertise and friendship were so very valuable in preparing the ground for this project.

Eve wishes to acknowledge particularly her co-author Lilian Ferrari, for a joyful cooperation. Her colleagues and students at Berkeley, and particularly in the Berkeley Gesture Group, deserve thanks for their many insights and their willingness to engage with her very viewpoint-pervaded model of language and thought; so do participants in various International Gesture Studies Conference sessions in 2008 and 2010. And finally, the smart and focused members of her Fall 2010 graduate seminar on Viewpoint at UC Berkeley did an invaluable beta-reading of the whole manuscript.

Introduction: viewpoint and perspective in language and gesture, from the Ground down

Eve Sweetser

In Wallace Stevens' poem "The anecdote of the jar", the narrator has put a jar on a Tennessee hilltop, with apparently dramatic results. The jar, we are told, *took dominion everywhere*, imposed order on *the slovenly wilderness*, and forced the wilderness to *surround* the hill on which it sat. Before the jar, it seems there were no human objects in the wilderness. And natural features like hills and valleys have no inherent preexisting relation between landmarks and "surroundings", since the relation between figure and ground is an aspect of a viewer's perceptual construal, rather than of the perceived objects.[†] The jar – a human artifact and a durable reminder of the human presence of the narrator – has proliferated new human-centered construal of space. It has made the land into a landscape.

A stretch of country with a human in it is no longer just a stretch of country – it is also a human's egocentric conceptualization of that physical area. Viewpoint permeates human cognition and communication – predictably, since we never have experience of the world except as a viewpoint-equipped, embodied self among other viewpointed, embodied selves. Language reflects this fact of embodiment: linguistic structure shows no way entirely out of viewpoint to an objective pre-experiential description of the world. But it also shows in complicated and fascinating ways the possibility of a single mind accessing multiple different viewpoint affordances on the same scene. Without such cognitive flexibility, humans could not cooperate and communicate at the high level that is apparently unique to our species, and universal to neurally and developmentally typical members of the species (Tomasello 1999, 2008). For this reason, viewpoint is a phenomenon of special interest to almost anyone studying cognition or communication: linguists, cognitive scientists, literary analysts, philosophers, and many more.

The title phrase in this introduction, *from the Ground down*, articulates something that the rest of this book confirms: cognitive perspective starts with bodily viewpoint within a real physical Ground of experience. In a mental space

[†] Readers are invited to enjoy the full text of the poem, as one of the best examinations of viewpoint and perspective we know. Copyright issues sadly prevented us from including it in the paper.

network, embedded spaces are thought of as being metaphorically “below” their “mother” spaces in the spatial network. By this metaphor, viewpoint flows downwards like water from the world as we directly experience it, to our embedded spaces of thought and speech – our invisible abstract conceptualizations and our construals for linguistic communication.

Viewpoint itself is universal. And alongside a great deal of cross-linguistic variation in how viewpoint structures are linguistically categorized and represented, there are also evident cross-linguistic patterns. Such patterns should come as no surprise: human neural architecture and experience put important constraints on the ways we are able to access perspectival construals. Both complex linguistic marking systems, and high-level literary and artistic manipulation of viewpoint, are built on shared early experience – for example, Primary Scenes (Johnson 1996, 1999a, 1999b; Grady 1997a, 1997b) link humans’ experience of visual viewpoint with locational proprioception, and with spatiomotor strategies for access and reaching objects.

Even more interestingly, we are not just *capable* of multiple viewpoints; we are in fact incapable of keeping to one single viewpoint of space, or of cognitive structure, when other humans are present. A situation involving multiple humans is necessarily structured, for participants and for human observers, via complex multiple viewpoints. This is at least partly because of some of the ways in which our brains process other humans’ activities. Research on monkeys has shown that some of the same neural activation patterns involved in grasping, touching, and acting on objects is also involved in viewing object manipulation by other primates (Rizzolatti and Arbib 1998; Rizzolatti *et al.* 2001) – these are so-called mirror neurons. Mirror neuron data come from monkeys, but have motivated work with human subjects, which has shown that there are in fact quite widespread patterns of such activation – motor cortex areas involved in actions are also partially involved in observing or hearing or reading about similar actions (Pulvermüller *et al.* 2001; Hauk *et al.* 2004; Buccino *et al.* 2005). Thus, as Katherine Young (personal communication; cf Young 2002) says, our bodies are naturally and constantly *occupied* by the neural patterns of surrounding human bodies – but not, apparently, in the same way by the spatial and force-dynamic relationships of inanimate objects. We are constantly aware of our bodily proximity to objects around us, but when another human is present, we are also unavoidably aware not only of our own human bodily affordances, but of (his or) hers as well – what she can reach, what she can see, and so on.

So, from the start, a human neural system is constrained to experience Ego’s body as special and different: when our mirror motor neurons fire in response to watching someone pick up an apple, the non-mirror ones do not fire; so we neither use our own muscles to grasp a phantom apple, nor hallucinate that we are picking up an apple. On the other hand, the shared mirror neuron firing leads

us naturally to categorize our own actions with those of others – and therefore to use the same verbs (*pick up, kick*), regardless of the identity of the agent. This is, in a sense, a remarkable achievement: the visual and motor and tactile experience of picking up an apple is radically different from that of viewing such an action. Such “dual” experience of others’ actions may well be part of the underlying basis for humans’ universal ability to treat and understand other humans as conscious social agents like themselves, and thus to develop Theory of Mind. Without Theory of Mind – and the confidence that others also have Theory of Mind – complex viewpoint structures would be impossible. Other species do not, as far as we can tell, build complex counterfactual spaces, but they also do not (and cannot) worry about specifically human concerns, such as whether Joe has figured out that Chris and I are attracted to each other (cf Zunshine 2006, for discussion of how these space embeddings play out in literary texts).

More surprisingly, we build viewpointed, sensory simulations in response to linguistic stimuli as well – and not just simulations of actions described, but of the situations involved (Richardson *et al.* 2003), and multimodally (Pecher *et al.* 2003, 2004). Subjects are faster at recognizing a vertically oriented picture of a nail than a horizontal one, after being primed with the sentence *He hammered the nail into the floor*, which implies a vertical orientation of the nail being hammered; the reverse is true after being primed with a sentence like *He hammered the nail into the wall* (Stanfield and Zwaan 2001). Subjects also simulate the temporal structure and completion of described actions differently depending on the linguistic aspect (Bergen *et al.* 2010) expressed. These are only a few examples of the increasing evidence (for overviews, see Barsalou *et al.* 2005; Barsalou 2010) that language prompts us to “run” simulations in the brain – simulations that are necessarily viewpointed, because the experiences on which they are based are necessarily viewpointed. This volume focuses on a much wider range of ways in which language and bimodal communication represent viewpoint – many of which have not yet been tested for simulation correlates in the brain. But since the volume is a contribution to the study of language in the context of *embodied* (or *grounded*) cognition, it is important for these works to be set in the context of evidence that communication activates the embodied neural system’s representation of viewpoint.

The Mental Spaces framework provides a way to represent these viewpoint phenomena. A mental space is a partial and local conceptual representation, which can be mapped onto or combined with other such spaces to build complex conceptual structure. Mental spaces differ from other constructs, such as possible worlds, in being cognitive. A Mental Space analysis of linguistic meaning does not presuppose that there is some reality to which a speaker’s understanding can be compared; all we humans have is our cognitive models of the world, based on embodied experience. These are very powerful. For example,

suppose we read a news story about a rock singer who is suing a magazine for libel in Italy. We have imagined mental spaces for lawsuits, journalism, libel, rock singers' lifestyle, and Italian culture, structured by frames with which we are familiar. Combining these, we build a richer cognitive representation than the news story could possibly evoke alone. Indeed, like most language, and as cogently remarked by Fauconnier (1985 [1994], introduction to 2nd edn. of *Mental Spaces*), the news story's actual words and constructions are only prompts to readers to engage in space-building – they would be useless without the speaker's knowledge base and space-combining (or *blending*) abilities.

Since a given mental space is always attached to some perceiver or cognizer, mental spaces necessarily structure viewpoints. Let us suppose that you and I are sitting next to a table, and both of us independently reach very much the same assessment of the table's physical size, hardness, and other physical characteristics. We are unlikely to notice that we have different construals of the table – so we feel as if we have access to "reality," when in fact we each have access to similar experiential data of the same object via similar embodied neural systems. But language is there not only to represent these unproblematic convergences in conceptual structure, but also to represent all the more complex situations that can arise as humans construe situations in varied ways. Verbs such as *say* and *think*, which explicitly mark expressed or unexpressed cognitive states, necessarily build mental spaces, as Fauconnier has said.

Because humans can embed mental spaces, and hold contradictory spaces in mind at once, they can produce not only represented speech and thought, but also negation and counterfactual conditionals. They can also notice discrepancies between accessible spaces in complex networks, producing effects such as irony – discussed by Tobin and Israel in this volume – and humor.

Linguistic viewpoint

Let us begin by surveying some of the range of linguistic forms that are markers of viewpoint. We shall label as *linguistic viewpoint* all the different ways that content is linguistically presented and construed differently, depending on (at least) the following range of factors noted by linguists.

- (1) Where the Speaker and Addressee are assumed to be, and what they are thought of as being able to see, be able to reach, and so on. English uses such as *here*, *there*, *this*, *that*, *next door*, depend for their reference on implicit information about the Speaker's and/or Addressee's presumed locations and their spatial relationships and access to objects designated. (Fillmore 1997 [1971])
- (2) When the Speaker and Addressee are assumed to be. Just as with spatial terms, grammatical tenses and linguistic usages, such as *now*, *then*, *tomorrow*, *last year*, depend for interpretation on the presumed time of utterance,

writing, reading, or other communicative act. Deixis is neutral as to *scale* of construal. English *here* could mean the room we are in, the town or state we are in, or the planet we are on, each of which is more accessible than the contrasting *there*. Similarly, I can say *this pencil*, *this side of the room*, *this side of the Atlantic*, or *this side of the galaxy*. The less recognized deictic *home* works similarly; in a science-fiction story it can easily mean a character's "home" planet. (See Fillmore 1997 [1971]; a mental space treatment of tense is laid out in Fauconnier 1997 and Cutrer 1994.)

- (3) What the Speaker and Addressee are assumed to know, think, presuppose, and be able to calculate mentally about whatever mental space is involved. Examples of markers that give such clues are:
 - (a) Determiners. The choice of *a* as opposed to *the* says something about the Speaker's assumptions about the Hearer's ability to identify a referent. A Mental Space approach to determiners is laid out in Fauconnier 1985 [1994].
 - (b) Pronouns, address forms, or honorific markers. The choice of a formal/distant rather than an informal/close second-person pronoun in languages with such a distinction (*tu* as opposed to *vous* in French) says something about the Speaker's construal of the social interaction, as do address choices such as *Professor Smith* versus *Mary* versus *Ma'am*.
 - (c) Connectives and evidential markers. Choosing *if* as opposed to *when* or *since* indicates the Speaker's lack of full positive epistemic stance commitment to the relevant mental space (Fillmore 1986, 1990a; Dancygier 1998; Dancygier and Sweetser 2000, 2005). A "hearsay" evidential marker (or non-grammatical marking such as *I hear that*) indicates the Speaker's lack of direct experience of the event referred to (Chafe and Nichols 1986).
 - (d) Presuppositional lexical items. The classic example is *stop*: saying either *Chris stopped smoking* or *Chris didn't stop smoking* indicates the Speaker's assumption that Chris smoked (Stalnaker 1974). A Mental Space analysis of lexical presuppositional structures can be found in Ferrari and Sweetser (this volume).
- (4) What the Speaker and Addressee feel about the contents of the relevant spaces – how they evaluate them affectively, culturally, and so on. Such evaluation includes:
 - (a) Framing (Fillmore 1982, 1985). Calling a given behavior *thrifty* as opposed to *stingy* may not actually depict a contrasting set of behaviors, but certainly indicates that the Speaker frames the (possibly identical) behavior of reluctance to spend money as in the one case prudent and reasonable resource conservation, and in the other case unreasonable and possibly selfish refusal to use resources as appropriate.

- (b) Affective markers. Starting a sentence with *hopefully* marks the Speaker's positive emotional assessment of the eventuality mentioned, just as *maybe* marks her epistemic assessment. Dancygier and Sweetser (2005) have proposed the need to distinguish constructional semantics of *positive emotional stance* from Fillmore's (1990b) related concept of positive interest, in space-building constructions such as *if only*.

Many of these forms might better be analyzed as in some way negotiating Speaker and Hearer viewpoints, more in line with Verhagen's (2005) *intersubjectivity* than with the hypothesis that they simply mark Speaker viewpoint. Definite articles, for example, clearly mark some idea of shared cognitive accessibility, or Speaker's assessment of accessibility to the Hearer; negation and stance verbs (see Dancygier, this volume) negotiate stance between Speaker and Hearer. Intersubjectivity will be discussed in more detail in the next section.

But the list above is the tip of the iceberg. Since the whole point of mental spaces is precisely that humans can manage to separate (or blend) cognitive representations from different cognizers or experiences, naturally language can also express what *imagined* participants can reach, touch, perceive, know, think, presuppose, calculate, and feel about relevant spaces – not just what present speakers and hearers may be cognizing. And we may add that if we take embodied cognition seriously, all hearers and readers are imagined hearers/readers – we have no direct access to their cognitive states, so we are always speaking or writing to a reader or a hearer whose knowledge states, presuppositions, affect (and so on) we are estimating or imagining.

Like markers of thought and speech, linguistic markers of affective and perceptual states are themselves builders of mental space structure. A headache or fatigue can only be directly known by the experiencer, so third-person sentences such as *His head ached* or *He felt tired* would require a non-direct evidential marker in some languages (a hearsay marker perhaps, indicating that he told me about his headache). As a result, the use of such sentences without hearsay marking can also indicate character viewpoint in a third-person narrative, since a separate narrator (not blended with the character) would not be able to speak directly about these aspects of character experience. Similarly, verbal aspect correlates with particular relationships to experiential states; imperfective aspect, which marks viewing an event from inside its temporal extent, is correlated with character viewpoint in fiction for this reason. *It snowed that night* is not how the character experiencing the snow would put it; she would say or think, *It is snowing*, or, in free indirect style accommodated to the narrator's past tense use, *It was snowing* (for a mental spaces analysis of free indirect style, see Sanders and Redeker 1996). Vandelanotte's and Nikiforidou's chapters in this volume examine some of the special literary stylistic exploitations of linguistic viewpoint markers in literary texts.

Metaphor adds yet more layers. Spatial closeness and distance, for example, have basic correlations with social intimacy and “distance” – the perfect basis for a Primary Metaphor in Grady’s and Johnson’s terms (Johnson 1996, 1999a, 1999b; Grady 1997a, 1997b). Unsurprisingly, therefore, most cultures use deictic terms to indicate social relationship and differentiation, as well as spatial relation to Ego. Since temporal distance is correlated in turn with epistemic distance, a further Primary Metaphor (Fleischman 1989) motivates the use of temporal markers to indicate cognitive and social reticence.

In short, viewpoint is marked by just about anything that builds a particular individual’s mental space construal in ways specific to that individual’s cognitive and perceptual access. Authors of fiction exploit this constantly (Banfield 1982) – in a novel where narrative viewpoint shifts between characters, labeling a character *Mommy* as opposed to *Chris* can be quite sufficient to let us know which other character’s viewpoint is currently on stage. Linguists have often preferred to focus on explicit grammaticalized markers whose function is to express shifts in viewpoint, such as pronouns or deictic markers. But in fact, any linguistic form choice is evidence concerning some mental space – and hence about the relevant cognizer’s viewpoint.

Subjectivity and deixis

Many analysts have noted that very often the presence and activity of the speaker or conceptualizer is left largely *implicit* in linguistic forms. Many utterances are not “about” the speaker – for example, *Joe walked into the café* seems to be primarily about the event of Joe’s movement into the café. However, the past tense of the verb tells us that we should construe the reference event as taking place before the time of utterance. The use of *I*, *we*, or *you* can be seen as foregrounding, or bringing into explicit content, parts of the Ground (as Langacker [1987, 1990, 1991] has labeled the Speaker’s or Conceptualizer’s communicative setting) – these pronouns both refer to elements of the Ground, and require access to that Ground, to identify the referent.

Performative verb forms make certain aspects of the speech act Ground into the explicit content of the utterance. An oath such as *I, Barack Hussein Obama, do solemnly swear that I will faithfully execute the office of President of the United States . . .* is intended to make it maximally clear who the speaker is, and exactly what speech act is being performed. Even the referent of *I* is explicitly specified, which is a highly unusual legal requirement for a spoken utterance. Most everyday interaction involves plenty of implicit shared Ground, which need not be mentioned. One friend could accept another’s emailed dinner invitation by saying, *I accept your invitation*, but might more likely say something else, such as, *Thanks, we can all come*, or *Great – what should I*

bring? The prospective host would not have any trouble identifying the speech act of acceptance, or (given the email address) the identity of *we* or *I*.

Langacker (1987, 1990, 1991) characterizes as **subjectivity** the “offstage” *implicit* presence of the conceptualizer and the Ground in construal. The greater this implicit presence, the more subjective is a construal. Thus, for example, *Joe is sitting across the table from me* does require reference to the Ground; there is no mention of the speech act itself, but the speaker participant is explicitly referenced (brought “onstage”) by using the first-person pronoun *me*. On the other hand, *Joe is sitting across the table* could be interpreted as meaning “across the table from me,” even though there is no explicit first-person pronoun. It could also be interpreted as meaning “across the table from whoever is referentially accessible in the discourse context” – again, without explicit reference to Speaker, Hearer and discourse context. These construals are thus more subjective (or less objective) than the construal involved in *Joe is sitting across the table from me*.

Similarly, Traugott (1982, 1989) uses the term **subjectification** to refer to a unidirectional trend in semantic change towards meanings that are increasingly rooted in the discourse context of the Speaker–Hearer interaction.¹ The development of an epistemic modal meaning from a root modal meaning is a case of subjectification. The root modal sense of *should*, for example, refers to some obligation incumbent on the agent, but not necessarily involving the conversational participants: *He should be home by now* (e.g. because he is ten years old and his parents have rules). Epistemic *should*, on the other hand, refers specifically to the speaker’s epistemic evaluation of the rest of the utterance’s content: *He should be home by now* (meaning “I judge it probable that he is home by now”). Aspect markers frequently develop historically into tense markers, for similar reasons: tense, unlike aspect, involves reference to a Speaker’s Now. Although Traugott does not stress this, she clearly does mean *implicit* reference to the Speaker–Hearer interaction: modals and tense markers do not mention the action of speaking or the participants explicitly, but rather depend on them for interpretation. Stein and Wright (1995) expand subjectivity to the point where it might well cover all of viewpoint; it is an interesting question whether one can draw a line between the two, but here we maintain that it is a useful distinction.

Neither of these understandings of subjectivity and subjectification deal primarily with the complexity of relationships within the Ground: the Ground, or the Speaker–Hearer interaction, is treated as a whole, all of which has special implicit status unless parts of it are brought explicitly onstage by direct mention. Verhagen’s (2005) concept of *intersubjectivity* (also brought up in Traugott and Dasher 2002) further elaborates the Ground for us, reminding us that it includes sharing and negotiation of viewpoint between Speaker and Hearer. And Sanders *et al.* (2009) have proposed an understanding of the Communicative Ground

itself as a mental space Network (the Basic Communicative Space Network), including representations of the Speaker's and Addressee's epistemic spaces, as well as the space of speech acts and speech act interaction between participants, and one or more content spaces.

Deixis (Levinson 1996a, 1996b, 2003; Fillmore 1997 [1971]) refers to the conventional use of linguistic forms whose meaning depends on the (implicit) Ground. As Levinson has pointed out, the typology of deictic systems shows some structures to be strongly favored. One very common system for spatial deixis is a three-way one, where A ("this") means "near the Speaker, or nearer to S than to the Hearer," B ("that") means "near H and not near S," and C ("yon") means "not near either S or H." Factors that turn out to be relevant are whether the located entity is manually accessible to S/H, whether it is visible to them, and whether it is possessed by, or in the custody of, one of them.

The question of custody or ownership complicates the spatial understanding of deixis and forces a more social construal. Hanks (1990, 1996) points out the use of a proximal Mayan deictic to refer to the cooking hearth that is further away but used by the Speaker, and a distal one to refer to the hearth that is currently right next to the Speaker but used/owned by someone else. English deictic verbs also make reference to locations conventionally or habitually associated with S or H, rather than only to S/H's physical locations at speech time. In English it would be usual to say *Can you come to my party?* as an invitation, even if I am not at home when I invite you, and even if the party is not being held at my home.

An under-studied aspect of deixis is the phenomenon of deictic *displacement* or deferral: a clear example is that in English, the correct response to the invitation *Can you come to my party?* is *Sure, I'd love to come*, not *Sure, I'd love to go*. The invitation accepter might later say to a third party, *I'm going to Sandy's party on Friday*, and would be unlikely to say *come* in this context. The acceptance utterance thus participated in the inviter's deictic structure, displacing the accepter's deictic center to the inviter's. Notice that *I* and *you* are not similarly displaceable – a Speaker would not, and could not, accommodate an Addressee by using personal pronouns relative to his deictic structure rather than hers. Nor are these particular viewpoint extensions of spatial deictic *come* paralleled in the COME and GO verbs of all languages. Emanatian (1992) details a remarkable example from Chagga, where there is displacement also in temporal uses of *come* as a future-marker, rather as if English speakers said *I'm coming to V* with the meaning of a GONNA future.

Figure 0.1 shows a treatment of the English addressee-centered usage of *come* as a blend. Note that the Speaker of *Can I come to your party?* has not completely recentered her deictic field on the Addressee – *I* still refers to the Speaker, and *you* to the Addressee, so the S/H referential system is not displaced. But the spatial deictic coordinate space, which is most canonically

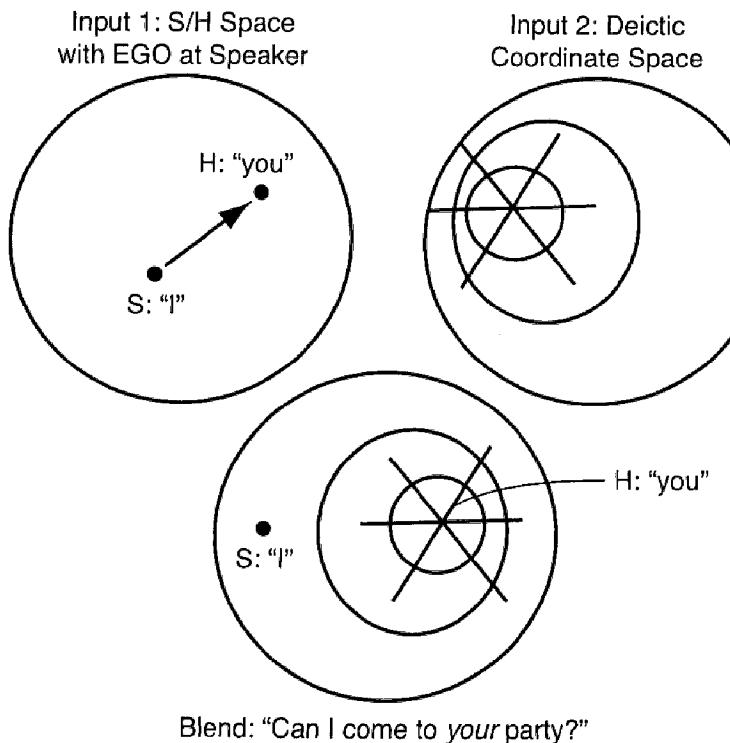


Figure 0.1

centered on Ego (hence on the Speaker), in this blend is mapped onto the Addressee as center.

The crucial point here is that our everyday construal of personal viewpoint is a blend. It is a blend that is so common that it is hard to notice it. We normally experience our own bodies simultaneously as loci of our conscious Selves or Egos, agents of our speech and action, spatial sources of our fields of perceptual access and manual reach, interfaces of social interaction, and more. But as stated above, we also naturally create such models for other individuals around us – aided, very possibly, by our mirror neurons, which respond to other humans' grasping actions (for example), as well as to our own. Once that is accomplished, a speaker can naturally describe motion away from herself with *come*, if she is profiling the deictic field structure relative to another participant (and leaving out mention of her own deictic field). It seems much harder to displace *I* from the speaker's identity – one cannot say *Can you come to my party?* meaning *Can I come to your party?* Embedded viewpoints are certainly possible, keeping separate mental spaces clearly separate (*You said, "Please come to my party"* – where *my* refers to the addressee within the addressee's quoted speech space), but one cannot build this blend within a single space. As Nikiforidou points out in this volume, similar blends occur in temporal language, so that a narrator's tense center may be combined with a character's

temporal adverb to produce examples like *He was now living in retirement*; the opposite blend is in this case possible too (spoken English past-reference narratives contain examples like, *So then the teacher leaves the room and the kids start talking*).

We also need to distinguish between a **viewer's perspective** and a full **deictic center** involving the speaker and the speech event. In English, *come* (and to a lesser degree, the less marked *go*) involve deictic centers, while verbs such as *arrive* involve a viewer's perspective, which need not be that of the deictic center. *Arrive* clearly does take the perspective of the arrival point, in some sense. But *Sue has arrived in Los Angeles* could be spoken either by someone in Los Angeles or by someone in San Francisco, and it is perfectly acceptable to use *arrive* with either a proximal or a distal deictic (*Sue has arrived there/here*). *Sue has come to Los Angeles*, however, is normally interpreted as spoken by someone in Los Angeles; and *come there* is distinctly strange. This is because not only *come*, but *here* and *there* invoke full deictic centers, unlike other perspectival adverbials, such as *across the table*. As Langacker (1987, 1991) notes, *Sue jumped across the table* is normally interpreted as meaning “across the table from Sue's initial pre-jumping location”. But *Sue sat across the table* means “across the table from someone whose perspective is taken” – the speaker and hearer are obvious options, but if we just mentioned Mary sitting at the table, then Mary is a good candidate.

Purely social viewpoint is built into many aspects of language, and indexes Speaker–Hearer interaction as surely as deictic physical viewpoint does. It is not normally referred to as deixis, and is generally studied by sociolinguists rather than semanticists. Yet consider the following attested scenario. A senior male university professor at the Sorbonne is about to introduce an honored speaker in a special lecture series; the speaker is his friend, his colleague, and a senior professor from a university in South America. As they chat before the talk, they address each other by first names and use the informal French second-person pronoun *tu*. Moments later, as he addresses her in an official welcome over the microphone, he is calling her *Professor X*, and using the formal address pronoun *vous*. An American introducer would not of course be able to change address pronouns, but would probably similarly switch from *Mike* to *Professor Jones*, in moving from private conversation with the speaker to the public introduction: the audience is not part of the personal interaction that ratifies the first-name use. Similarly, an American academic would normally refer to faculty colleagues and office neighbors by first names, but in doing undergraduate advising, that same faculty member would probably suggest that the student consult “Professor X” to get references for her paper. These choices, then, are intersubjectively (in Verhagen's terms) indexical not only of the relationship between the speaker and the referent, but of the relationship between the hearer(s) and the referent, which may dominate in determining the choice.

Rubba (1996) has pointed out that supposedly spatial deictics, such as *this* and *that*, *here* and *there*, are also used socially: the proximal ones may indicate “my cultural kind of neighborhood,” and the distal ones “their [not my] cultural kind of neighborhood.” Rubba’s data are particularly compelling because, for other reasons, her interviews (concerning the English-only movement in California) all took place in the same room on the University of California, San Diego campus, relatively distant from all the (socially proximal or distal) neighborhoods referred to. But more such instances are readily accessible: those who listen at linguistics conferences will hear groups of same-campus folks occasionally using *here* to mean “our campus,” distant though that campus may be from the conference location.

Japanese honorifics are one much-studied area of social deixis. They are more complex than English or French address terms, since they apparently extend to cover identification of groups with individuals. Thus Dasher (1995) cites corporate telephone conversations where the secretary phones the President’s wife to ask how he is recovering from the flu and says, “How is the ‘honored’ President’s health today?” using the positive honorific for the President. This is an example of what is usually thought of as the central use: the secretary is lower in status than the President, so uses an honorific in addressing or referring to him. Moments later, however, the same secretary is canceling one of the President’s external appointments for him, and says to the other executive’s secretary, “Our ‘humble’ President can’t keep his appointment because he is ill today,” using the (normally first-person-associated) humiliative marker. This is because, in canceling the appointment, she is acting on behalf of the President and his organization, and is thus identified with them; while in inquiring about his health, she is a family outsider and subordinate talking to an insider.

Collectively, the linguistic data suggest that we should be thinking in general of viewpoint as an intersubjective phenomenon (in Verhagen’s [2005] sense), rather than as a unitary first-person phenomenon – that is, addressees’ and others’ viewpoints are always relevant, along with the speaker’s own viewpoint, in contributing to the speaker’s choice of linguistically expressed viewpoint. We know that this is true with respect to how we negotiate spatial viewpoint and affordances; you would not ask someone to pass the salt unless it were *both* accessible to him and inaccessible to you. So we should not be surprised to find that similar joint determination applies to communicative and linguistic viewpoint structure.

Viewpoint across modalities

Examination of some of the areas where perspective is known to be central – such as linguistic deixis (Hanks 1990) or gestural pointing (Kita 2003) – has helped to motivate researchers to examine the much broader presence of

perspectival phenomena in language and bimodal communication. Signed language forms necessarily embody portrayed viewpoints in ways that spoken language does not (Liddell 1998, 2003; Dudis 2004; Janzen 2005, 2008; and Schaffer's and Janzen's chapters in this volume). Examining these phenomena has helped spoken language researchers to elaborate their understandings of the complexities of viewpoint. Data from all the relevant sources thus come together to push us towards a theory that includes mutual awareness of viewpoints between participants – and incorporation of addressees' and spectators' and readers' perceived viewpoints into the cognitive perspectives of speakers, narrators, and writers. Intersubjective viewpoint construction is important at every level of communication, from co-speech gesture to literary narrative.

First, let us briefly examine gestural viewpoint. Gesture, like language, shows deictic centers, displacement phenomena and blended deictic structure. Our bodies are the most flexible and powerful **material anchors** (Hutchins 1995; Fauconnier and Turner 2002) for representing and expressing viewpoint. This is because, as is obvious but rarely stated, there is no more powerful icon for a bodily viewpoint than an actual body with an actual inherent viewpoint. Gesture therefore expresses spatial indexicality via embodied spatial indexicality – for humans, a pointing gesture directs joint attention to a particular actual location relative to the actual pointing body. This in itself is a cognitive achievement; as Tomasello (1999, 2008) points out, great apes do not share human abilities to interpret points, though some dogs do. As Sizemore and Sweetser (2008) show, social deixis is also represented in gestural structure; speakers gesture in different parts of gesture space to adjust interaction (claim the floor, resist interruption), to depict content, and to elaborate spatial structure.

But of course, just like linguistic deixis, gestural deixis is not that simple. Displacement or deferral of gestural deictic viewpoint is pervasive. Haviland (2000) notes what he refers to as *transposed* points in his Mayan data; these are made relative to some imagined location rather than the gesturer's real current location. As an English example, I could say, *As you enter my office, the light switch is here*, moving my left hand as if to turn a switch on a wall to my left. Since English is a relative spatial language, my fellow English speakers will understand that the light switch is located to the left of a person standing in the doorway and looking inwards – no matter which direction I am currently facing, and no matter which direction my (non-present) office door faces. Crucially, they will not even think of looking for the light switch in the actual physical direction of my physical point in Real Space. Their understanding of this deferred point is a blend of the gesturer's current Real Space with the imagined space of my office door and its surroundings. If I were speaking an absolute spatial language, such as Guugu Ymithirr (Levinson 1996a, 1996b, 2003), interlocutors would instead probably create a different blended space, where my current absolute *direction* of stance (e.g. northeast)

was part of the input to the interpretation of the deferred point – that is, they might assume that someone entering my office would be facing northeast.

In general, gesture allows Real Space blends of some or part of the gesturer's body with a character's body (Parrill and Sweetser 2004). For example, describing a character climbing a ladder, a gesturer may use her hands in "clambering" gestures, as if going up an imaginary ladder, so that her hands represent the hands of the described climber. Alternatively, she may move one hand upwards in representation of the whole character's bodily upwards motion. This contrast represents two rather different viewpoints on the same scene, the first a Character viewpoint (the gesturer is "being" the character in the blend) and the other a Narrator (or Global) viewpoint (the gesturer is more of an observer of the character, as represented by her hand), as described by Parrill's chapter in this volume.

Gesture also uses Real Space more abstractly. It can exploit conventional Real Space blends to give abstract meaning to spatial locations – for example, conventional mappings of Time onto Space allow forwards-directed pointing gestures to have future time reference for English speakers. This is complex and language-specific, but apparently cognitively pervasive for speaker-gesturers (Boroditsky and Ramscar 2002; Matlock *et al.* 2005; Núñez and Sweetser 2006). Gesturers can also establish *loci*, locations in Real Space, which correspond to referents in content being discussed, without making necessary any mapping of spatial relations onto another imagined physical space – and even without the referents being physical entities. For example, *loci* can be entities being contrasted (my family and your family, logic and linguistics) or topics under discussion.

Signed languages famously not only represent the speaker's physical viewpoint, but use the visual-gestural modality to carry out linguistic deixis and reference. It is a question for research what similarities there may be between the use of referent *loci* in signed languages and the use of *loci* in gesture. But it is clear that in signed languages, pronouns consist of systematic pointing to (or moving signs towards or away from) such *loci*, with grammatically determined hand shapes. Or, as some might say (Liddell 1998, 2003; Taub 2001), interaction with spatial *loci* does much of the same referential work for signed languages that spoken pronouns or person-agreement affixes do in spoken languages. And referring to objects and people who are actually present is normally done by exploiting (pointing to, signing towards, etc.) their actual locations.

Signed languages can also exploit blended viewpoint structures of the kind seen in gesture, but in a far more systematic and grammaticalized way, as Shaffer and Janzen each discuss in their chapters in this volume. American Sign Language grammars in particular lay out a system of *role shifts*, where a signer wishing to take on the role of a character rotates her body noticeably

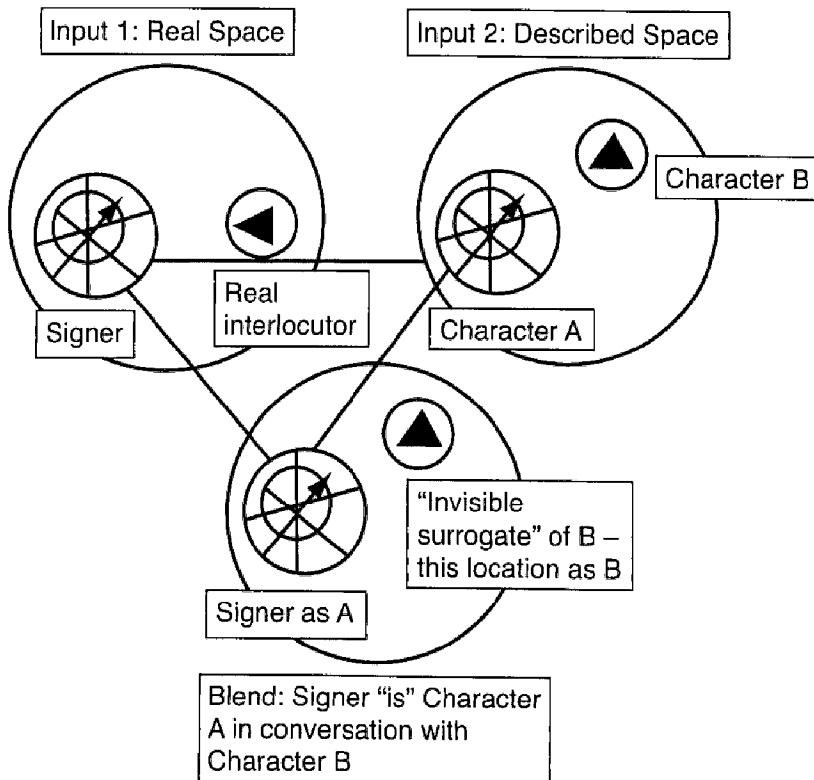


Figure 0.2

(perhaps 20 degrees or so) to one side or the other, and ends eye contact with the interlocutor (since the represented character is naturally not addressing the interlocutor). The signer's signed utterances, facial expressions, and gestural enactments while in role shift can be understood to be those of the character. The character's discourse may involve a different imagined physical space blended with the signer's actual Real Space (if he drives a car, the signer will be gripping an imagined steering wheel), and also different referential *loci*, depending on what the character needs to refer to. That is, the signer's Real Space is simply re-mapped to the character's spaces, rather than to the signer's own mental spaces and physical surroundings. Return to facing the interlocutor, with eye contact, indicates an end to the shifted role and a return to the signer's own viewpoint. The resulting role shift blends are *grounded blends* (one of whose inputs is the Real Space) – as are many gestural blends.

Figure 0.2 shows the blend that results from role shifts. Here the signer is taking on the role of a described Character A; she turns slightly towards her left, thereby aligning her own Ego-centered spatial coordinates towards those of the imagined Character A. (Crucially, she is no longer facing or signing directly towards her real-world interlocutor in Real Space – this marks the role shift.) In this role-shifted posture, the signer signs; we understand that this represents

the character as signing to another character, who has probably been identified in the discourse. So in the described world, we construct Character B, facing towards Character A and conversing with her – but there is no Real Space representation of B, only a location towards which the Signer faces. In the Real Space, there is thus no Character B, only the Real-Space interlocutor(s) of the signer. In the described space and in the Blend, the Real Space interlocutors do not exist, but Character B is present. A moment later, the signer may turn towards her right, creating a new blend in which she represents Character B, who is imagined to be facing rightwards, towards (and having a conversation with) Character A. Her spatial coordinates will then be blended with those of the second character. Thus, although just a moment ago she was looking to her left at an imagined interlocutor, a moment later (as that interlocutor) she may be looking to her right, at the same character she was enacting a moment earlier.

In this volume, Janzen further outlines another common ASL narrative viewpoint-changing strategy, not previously attended to by linguistic analysts, which involves *rotation* of the imagined space mapped onto the signing space, instead of rotation of the signer's body. If a signer is representing two characters who are facing each other and have opposite viewpoints, the signer may reassign the signing space alternately to each of the two characters. At least some of the same imagined physical space (the surroundings of the two characters) will be represented in the same signing space (the Real Space in front of the signer's body) – but the mappings will be rotated to fit each character's perspective. This is extremely different from role shift, but seems common and natural in ASL narratives.

In both role shifts and sign-space rotations, as in gestural blended “character viewpoints” (but unlike spoken-language viewpoint), a viewpointed body iconically represents another viewpointed body. This has immense consequences for representation of viewpoint, which often needs no active expression beyond offering a bodily posture. As Sweetser (2009) has pointed out, English speakers normally gesture about ongoing processes (solving a problem, getting a Ph.D.) with hand motions outwards from the body, rather than inwards towards the body – that is, the start of the activity is mapped onto points closer to the body, and later stages are mapped onto more distant points. (The exception to this generalization is reference to physical motion towards the speaker, where actual physical deictic structure dominates and a speaker will gesture towards herself as she says, *Come here*.) This is probably because we do experience our body as a source of actions – when we grasp something, we stretch out our hand to do so, and we stop stretching it out when our goal is attained. There is no such deictic structure necessarily built into spoken-language linguistic descriptions of getting a Ph.D. or solving a problem, confirming the instinctive impression that the visual/gestural modality is even more pervasively viewpointed than the spoken one.

Conclusion

The evidence is overwhelming that human thought is essentially embodied – literally, in the sense that it consists of neural patterns, and also in the broader sense that its character necessarily reflects embodied perception and reasoning (see Lakoff and Johnson 1999; Gibbs 2006, for overviews of this issue). Embodied human cognition is inherently pervaded by viewpoint, and therefore communicative manifestations of cognition are too, in a remarkably diverse range of ways. This makes perfect sense if we assume that linguistic forms are (Fauconnier 1994 [1985], 1997) prompts for mental space building, and that the process of space building involves mental *simulation* of the situations and events referred to. *Simulation-based* semantics – that is, an analysis of language based on the assumption that meaning involves activated neural simulation of described events – is now a theoretical framework being developed by a large community of researchers (Barsalou *et al.* 2005; Bergen and Chang 2005; Feldman 2006; Bergen 2007; Barsalou 2010). Those researchers are taking seriously the idea that even for a sentence such as *Joe walked into the café*, which is in itself neutral as to whether it is describing the event as viewed from inside or outside the café, the simulations prompted in listeners or readers cannot be equally neutral. Better understanding of all the varied manifestations of viewpoint in communication is therefore crucial to developing a more cognitively realistic understanding of processing – even of processing sentences that do not seem to prompt a particular viewpoint. And, as has been observed again and again (McNeill 1992, 2000, 2005; Cienki 1998; Goldin-Meadow 2003; Narayan's chapter in this volume), multimodal data will give evidence about communicators' cognitive processing that is not accessible from a single modality of output.

We also feel that, just as some theoretical frameworks are shown by the chapters in this volume to be increasingly helpful in examining viewpoint, other conceptual baggage is being shed as the field moves forwards. Readers will note that, perhaps surprisingly in a book on viewpoint, this introduction has not discussed the role of context in determining viewpoint. This is not because we, or our authors, are inattentive to contextual factors in communicative situations – on the contrary. Rather, we feel that the cover term *context* is too broad to be meaningful, when the range of studies is as varied as this. For analysts of literary texts, *context* crucially means the surrounding text. For analysts of face-to-face recorded communication, although the immediate linguistic “record” is of course relevant to production and interpretation at any time, the Real Space is also relevant, and can become relevant in unforeseen ways. When someone describes a person pounding on a table, having a surface to gesturally pound during the description suddenly becomes relevant context – although the Real Space surface may not have been very important to the communicative exchange until that point. In all communicative situations, context might also be

said to include shared models of how communication works (spoken, written, or signed) – so Sanders *et al.*'s (2009) Basic Communicative Space Network is context, as well as the actual circumstances of the exchange. Also ever-relevant to every communication are shared beliefs and frames that can be recruited to populate spaces – whether local or cultural frames. And the list could continue. For such a varied set of studies, we feel that it is more important to focus on the actually relevant aspects of context – different kinds of contextual affordances are present in different modalities and situations.

What this volume is doing, therefore, is bringing together researchers who work on widely differing kinds of data (experimental data, face-to-face recordings, corpora, literary texts) and in different modalities (signed and spoken languages as well as co-speech gesture) to bring a new overall picture of viewpoint. In some cases, our authors are building on long traditions, but significantly elaborating them (Vandelanotte) or reinterpreting them to offer new generalizations in the context of current grammatical and cognitive theories (Nikiforidou, Tobin and Israel). In other cases, the relevant phenomena are either underdescribed (Shaffer, Parrill) or previously undocumented (Janzen, Narayan). What fascinated the editors, when we initially gathered this group of contributors, was the degree to which the participants' presentations converged on a shared understanding – and also the degree to which that surprised them. We had initially thought that interaction between them was sure to be productive. But when we realized that this collaboration had helped the central participants to become more aware of the degree to which their work was addressing related issues, we were further convinced that the rest of the linguistic and cognitive science communities could also find this volume a useful contribution.

References

- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston: Routledge and Kegan Paul.
- Barsalou, Lawrence W. 2010. Grounded cognition: past, present and future. *Topics in Cognitive Science* 2, 716–24.
- Barsalou, Lawrence W., Diane Pecher, Rene Zeelenberg, W. Kyle Simmons, Woo-kyoung Ahn, and Stephan B. Hamann. 2005. Multi-modal simulation in conceptual processing. In Woo-kyoung Ahn, Robert L. Goldstone, Bradley G. Love, Arthur B. Markman, and Phillip Wolff (eds.). *Categorization Inside and Outside the Lab: Essays in Honor of Douglas L. Medin*. Washington, DC: American Psychological Association, 249–70.
- Bergen, Benjamin and Nancy Chang. 2005. Embodied construction grammar in simulation-based language understanding. In Jan-Ola Östman and Miriam Fried (eds.). *Construction Grammars: Cognitive Grounding and Theoretical Extensions*. Amsterdam: John Benjamins, 147–90.

- Bergen, Benjamin. 2007. Experimental methods for simulation semantics. In Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, and Michael J. Spivey (eds.). *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, 277–301.
- Bergen, Benjamin, Avis Lau, Shweta Narayan, Diana Stojanovic, and Kathryn Wheeler. 2010. Body part representations in verbal semantics. *Memory and Cognition* 38:7, 969–81.
- Boroditsky, Lera and Michael Ramscar. 2002. The roles of body and mind in abstract thought. *Psychological Science* 13:2, 185–8.
- Buccino, G., L. Riggio, G. Melli, F. Binkofski, V. Gallese, and G. Rizzdatti. 2005. Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research* 24, 355–63.
- Chafe, Wallace and Johanna Nichols (eds.). 1986. *Evidentiality: The Linguistic Encoding of Epistemology*. Norwood NJ: Ablex.
- Cienki, Alan. 1998. Metaphoric gestures and some of their relations to verbal metaphoric counterparts. In Jean-Pierre Koenig (ed.). *Discourse and Cognition: Bridging the Gap*. Stanford CA: Center for the Study of Language and Information, 189–205.
- Cutrer, Michelle. 1994. Time and tense in narratives and everyday language. Ph.D. dissertation, University of California, San Diego.
- Dancygier, Barbara. 1998. *Conditionals and Prediction*. Cambridge University Press.
- Dancygier, Barbara and Eve Sweetser. 2000. Constructions with *if*, *since* and *because*: causality, epistemic stance and clause order. In Elizabeth Couper-Kuhlen and Bernd Kortmann (eds.). *Cause, Condition, Concession, Contrast*. Berlin: Mouton de Gruyter, 111–42.
2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Dasher, Richard. 1995. Grammaticalization in the system of Japanese predicate honorifics. Ph.D. dissertation, Stanford University.
- Dudis, Paul. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15:2, 223–38.
- Emanatian, Michele. 1992. Chagga ‘come’ and ‘go’: metaphor and the development of tense-aspect. *Studies in Language* 16, 1–33.
- Fauconnier, Gilles. 1994 [1985]. *Mental Spaces*. 2nd edn. Cambridge University Press.
1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think. Conceptual Blending and the Mind’s Hidden Complexities*. New York: Basic Books.
- Feldman, Jerome A. 2006. *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge MA: MIT Press.
- Fillmore, Charles J. 1982. Frame semantics. In Linguistic Society of Korea (ed.). *Linguistics in the Morning Calm*. Seoul: Hanshin, 111–38.
- 1990a. Epistemic stance and grammatical form in conditional sentences. *CLS* 26, 137–62.
- 1990b. The contribution of linguistics to language understanding. In Aura Bocaz (ed.). *Proceedings of the 1st Symposium on Cognition, Language, and Culture*. Universidad de Chile, 109–28.
- 1997 [1971]. *Lectures on Deixis*. Stanford CA: Center for the Study of Language and Information. (Originally published 1971 by the Indiana University Linguistics Club as *The Santa Cruz Lectures on Deixis*.)

1986. Varieties of conditional sentences. *ESCOL* 3, 179–95.
1985. Frames and the semantics of understanding. *Quaderni di semantica* 6:2, 222–54.
- Fleischman, Suzanne. 1989. Temporal distance: a basic linguistic metaphor. *Studies in Language* 13; 1–50.
- Gibbs, Raymond W. 2006. *Embodiment and Cognitive Science*. Cambridge University Press.
- Goldin-Meadow, Susan. 2003. *Hearing Gesture: How Our Hands Help Us Think*. Cambridge MA: Harvard University Press.
- Grady, Joseph. 1997a. THEORIES ARE BUILDINGS revisited. *Cognitive Linguistics* 8, 269–90.
- 1997b. Foundations of meaning: primary metaphors and primary scenes. Ph.D. dissertation, University of California, Berkeley.
- Hanks, William F. 1990. *Referential Practice: Language and Lived Space in a Maya Community*. University of Chicago Press.
1996. Language form and communicative practices. In John J. Gumperz and Stephen C. Levinson (eds.), *Rethinking Linguistic Relativity*. Cambridge University Press, 232–70.
- Hauk, Olaf, Ingrid Johnsrude, and Friedemann Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41: 2, 301–7.
- Haviland, John. 2000. Pointing, gesture spaces and mental maps. In David McNeill (ed.). *Language and Gesture*. Cambridge University Press, 13–46.
- Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge MA: MIT Press.
- Janzen, Terry. 2005. *Perspective Shift Reflected in the Signer's Use of Space*. CDS/CLCS Monograph 1. Center for Deaf Studies, University of Dublin Trinity College.
2008. Perspective shifts in ASL narratives: the problem of clause structure. In Andrea Tyler, Yiyoung Kim, and Mari Takada (eds.). *Language in the Context of Use: Usage-Based Approaches to Language and Language Learning*. Berlin/New York: Mouton de Gruyter, 129–54.
- Johnson, Christopher. 1996. Learnability in the acquisition of multiple senses: SOURCE reconsidered. *Proceedings of the 22nd Annual Meeting of the Berkeley Linguistics Society*, 469–81.
- 1999a. Constructional grounding: the role of interpretational overlap in lexical and constructional acquisition. Ph.D. dissertation, University of California, Berkeley.
- 1999b. Metaphor vs. conflation in the acquisition of polysemy: the case of SEE. In Masako Hiraga, Chris Sinha, and Sherman Wilcox (eds.). *Cultural, Psychological and Typological Issues in Cognitive Linguistics*. Amsterdam: John Benjamins, 155–71.
- Kita, Sotaro (ed.). 2003. *Pointing: Where Language, Culture and Cognition Meet*. Mahwah NJ: Lawrence Erlbaum.
- Lakoff, George and Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. 2 vols. Stanford University Press, vol. I.
1990. Subjectification. *Cognitive Linguistics* 1, 5–38.
1991. *Foundations of Cognitive Grammar*. 2 vols. Stanford University Press, vol. II.

- Levinson, Stephen C. 1996a. Language and space. *Annual Review of Anthropology* 25, 353–82.
- 1996b. Relativity in spatial conception and description. In Gumperz and Levinson, 177–202.
2003. *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge University Press.
- Liddell, Scott. 1998. Grounded blends, gestures and conceptual shifts. *Cognitive Linguistics* 9:3, 283–314.
2003. *Grammar, gesture and meaning in American Sign Language*. Cambridge University Press.
- Matlock, Tenie, Michael Ramscar, and Lera Boroditsky. 2005. The experiential link between spatial and temporal language. *Cognitive Science* 29, 655–64.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.
- (ed.). 2000. *Language and Gesture*. Cambridge University Press.
2005. *Gesture and Thought*. University of Chicago Press.
- Núñez, Rafael and Eve Sweetser. 2006. With the future behind them: convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30, 401–50.
- Parrill, Fey and Eve Sweetser. 2004. What we mean by meaning: conceptual integration in gesture analysis and transcription. *Gesture* 4; 197–219.
- Pecher, Diane, Rene Zeelenberg, and Lawrence W. Barsalou. 2003. Verifying properties from different modalities for concepts produces switching costs. *Psychological Science* 14, 119–24.
2004. Sensorimotor simulations underlie conceptual representations: modality-specific effects of prior activation. *Psychonomic Bulletin & Review* 11, 164–7.
- Pulvermüller, Friedemann, Max Haerle, and Friedhelm Hummel. 2001. Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language* 78, 143–68.
- Richardson, Daniel, Michael Spivey, Kateri McRae, and Lawrence Barsalou. 2003. Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27, 767–80.
- Rizzolatti, Giacomo and Michael Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21:5, 188–94.
- Rizzolatti, Giacomo, Leonardo Fogassi, and Vittorio Gallese. 2001. Neuropsychological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience Reviews* 2, 661–70.
- Rubba, Johanna. 1996. Alternate grounds in the interpretation of deictic expressions. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 227–61.
- Sanders, José and Gisela Redeker. 1996. Perspective and representation in speech and thought in narrative discourse. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 290–317.
- Sanders, José, Ted Sanders, and Eve Sweetser. 2009. Causality, cognition and communication: a mental space analysis of subjectivity in causal connectives. In Ted Sanders and Eve Sweetser (eds.). *Causal Categories in Discourse and Cognition*. Berlin: Mouton de Gruyter, 19–59.

- Sizemore, Marisa and Eve Sweetser. 2008. Personal and interpersonal gesture space: functional contrasts in language and gesture. In Andrea Tyler, Yiyoung Kim, and Mari Takada (eds.), *Language in the Context of Use: Cognitive and Discourse Approaches to Language and Language Learning*. Berlin: Mouton de Gruyter, 25–51.
- Stein, Dieter and Susan Wright (eds.). 1995. *Subjectivity and Subjektivisation*. Cambridge University Press.
- Sweetser, Eve. 2009. What does it mean to compare language and gesture? Modalities and contrasts. In Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, and Şeyda Özçalışkan (eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Traditions of Dan Isaac Slobin*. New York: Psychology Press, 357–66.
- Stalnaker, Robert C. 1974. Pragmatic presuppositions. In Milton K. Munitz and Peter K. Unger (eds.), *Semantics and Philosophy*. New York University Press, 197–214.
- Stanfield, Robert A. and Rolf A. Zwaan. 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12, 153–6.
- Taub, Sarah F. 2001. *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge University Press.
- Tomasello, Michael. 1999. *The Cultural Origins of Human Cognition*. Cambridge MA: Harvard University Press.
2008. *Origins of Human Communication*. Cambridge MA: MIT Press.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann and Yakov Malkiel (eds.), *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, 245–71.
1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 57, 33–65.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge University Press.
- Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford University Press.
- Young, Katharine G. 2002. The dream body in somatic psychology: the kinaesthetics of gesture. *Gesture* 2:1, 47–50.
- Zunshine, Lisa. 2006. *Why We Read Fiction: Theory of Mind and the Novel*. Columbus OH: Ohio State University Press.

Part I

Intersubjectivity and subjectification



1 Irony as a viewpoint phenomenon

Vera Tobin and Michael Israel

1.1 A curious example

In *Jesus Is Magic*, a concert film of her off-Broadway stand-up comedy routine, Sarah Silverman (2005) tells a joke:

- (1) “Everybody blames the Jews for killing Christ, and then the Jews try to pass it off on the Romans. I’m one of the few people that believe it was the blacks.”

The audience seems to have no trouble identifying this utterance as a kind of ironic joke, in which the comedian is presenting an absurd or offensive position in order to mock it, rather than simply asserting it. In order to appreciate the joke, her audience must understand it as performing more than one speech act at once. At the most superficial level, there is the act performed by Silverman’s cheerfully narrow-minded, unflappable onstage persona, who is presenting her novel theory about the death of Christ. This ostensible act alludes to at least two other (sets of) speech acts: first, the centuries-old anti-Semitic trope that “Jews killed Jesus,” and second, Lenny Bruce’s famous jokes skewering that claim (catalogued in Bruce 1963: 155). For example:

- (2) “Yes, we did it. I did it. My family. I found a note in my basement: ‘We killed him – signed, Morty.’”

Where the joke in (2) is a fairly straightforward example of verbal irony, however, (1) seems to involve something more complicated and perhaps even problematic. This joke is not quite like its predecessor. Something about it leads commentators to wonder if laughing makes the audience complicit in “the cheap thrill of public racism” (Anderson 2005), even though it is also clear that Silverman does not really believe that “the blacks” killed Christ.

While the structure and context of the joke invite the audience to join the comedian in the contemplation of something from an ironic distance, the actual object of Silverman’s ironizing is unclear. The victims of Bruce’s irony are real anti-Semites, but there is no such real-life bigot who believes that “the blacks” killed Jesus. The absence of any obvious viewpoint one could share

with the speaker, or any obvious way of figuring out where her viewpoint might really be, generates interpretive tension. Part of Silverman's edgy appeal rests on the difficulty of decoding her ironic intentions. Is the joke on racists, on the audience, on political comedy?

This off-kilter experience seems to us to be exactly the point of the joke. It is a concise example of what Wayne Booth (1974) called "unstable irony": an irony that offers no final interpretation that is not subject to the prospect of further ironic undermining. We would like to explicate the processes of meaning construction that make this joke both ironic and unsettling. While pretense is involved in this performance (cf Clark and Gerrig 1984; Kreuz and Glucksberg 1989; Clark 1996), and so is a sort of echoic mention (cf Sperber and Wilson 1981) of the original anti-Semitic remark, neither of these factors is self-evidently the source of the unstable ironic effect itself.

1.2 Isn't it ironic?

Irony is a puzzling thing. It has been a source of wonder for scholars in many traditions, from German Romantics to psycholinguists and Alanis Morrisette (the popular singer who famously asked, "Isn't it ironic?"). But since the beginning of the Modern era, the variety of phenomena called "irony" by rhetoricians, literary scholars, and the public in general has proliferated (Knox 1961), so it is worth considering whether all these phenomena really have anything of substance in common at all.

Within the cognitive sciences, it is common (Sperber and Wilson 1981, 1998; Clark and Gerrig 1984; Kreuz and Glucksberg 1989) to restrict studies of irony to "verbal irony," of which sarcasm is the paradigm case. These accounts tend to focus on the problem of how these kinds of ironic utterances are recognized; the basic goal is to identify the necessary features for an ironic utterance and the cognitive mechanisms that enable hearers to identify and interpret such utterances. Literary studies tend to come at irony from the opposite direction. Rather than considering readers' "successful" interpretations of any given ironic statement, literary accounts often seek to tease out more and more ironies surrounding a text, and to point out how these ironies make it difficult or impossible to pin down stable meanings (Empson 1947; Colebrook 2004).

We suggest that these concerns are in fact complementary, and that literary and linguistic theories of irony have much to gain from one another. There are two major issues that we feel have been somewhat neglected in linguistic theories of irony. First, "irony" is the name of not one thing, but a whole range of phenomena. Our account builds on theories that treat irony as a form of echoic mention (Sperber and Wilson 1981, 1998) or pretense (Clark and Gerrig 1984; Kreuz and Glucksberg 1989), but applies to a broader range of literary and cultural phenomena, including classical cases like Swift's "Modest Proposal";

cosmic and dramatic ironies; multilayered ironies of the sort found in Borges; irony as a kind of sensibility, as in Romantic irony and camp; and the existence of entire ironic genres, such as the pseudo-scholarly articles produced by fans of Arthur Conan Doyle, which operate under the conceit that Sherlock Holmes was a genuine historical figure.

It is possible that these different senses of irony are only related by a chain of historical associations, and do not form a natural type. Still, almost as long as there have been formal discussions of irony, it has been treated as a phenomenon with many guises. Even Quintilian, whose definition of irony as a trope in which one says the opposite of what one means is often presented as the canonical, overly simplistic “classical view” of the form, in fact described irony as something that could be expressed over the course of extended, discursive “figures of thought,” as well as through simple anaphrasis (Butler 1921).

Rather than presenting an account of verbal irony alone, then, theorists should consider whether verbal irony and other phenomena sometimes called “ironic” do, in fact, have anything significant in common. The production and interpretation of sarcastic utterances may well rely on similar cognitive mechanisms to those that underlie the performance and appreciation of dramatic, situational, and Socratic irony.

Intuitively, what seems to unite the various sorts of irony is the existence of some kind of complex viewpoint on a single situation, the quality that Fowler (1926) described as the “double audience” that distinguishes irony from mere incongruity. Our account suggests that this intuitive connection reflects genuine, shared, underlying conceptual structure.

The second issue is an apparent paradox. On the one hand, irony is difficult. The ability to understand it comes relatively late in cognitive development, and even adults frequently misinterpret it, so much so that the potential for misunderstanding appears to be a defining feature of the figure. But irony is also ubiquitous. It is a commonplace of cultural criticism that certain strata of Western culture can no longer be sincere, only “post-ironic.” Literature, of course, has long depended on irony for tragic or comedic effect, but irony also fills the emails we send (with newly invented typographic effects to signal one’s lack of sincerity), the music we listen to, even the clothes we wear and the food we eat – think of urbane adults who wear My Little Pony shirts or serve Moon Pies at their weddings. We are incapable, it seems, of resisting the ironic urge.

In light of these facts, a theory of irony ideally ought to explain how verbal irony relates to other kinds of irony, as well as why irony is both sometimes very hard to understand and sometimes very hard to control, allowing for apparently endless layering in certain contexts.

Drawing on work in Mental Spaces Theory (Fauconnier 1985, 1997) and Cognitive Grammar (Langacker 1999), we argue that irony is fundamentally a viewpoint effect in which a conceptualization is simultaneously accessed

from multiple perspectives. Acts of ironic understanding in general, including verbal, dramatic, and situational ironies, involve a type of dynamic reconstrual in which attention “zooms out” from the focused content of a mental space to a higher viewpoint from which the original Viewpoint Space is reassessed. In this interpretive process, a meaning is accessed from one viewpoint (the ironized) and then, simultaneously or a little later, re-accessed from a higher viewpoint (the ironic).

1.3 Various views of verbal irony

Figurative language in general poses a fundamental problem for a theory of utterance interpretation – how is it that a speaker can say one thing, mean something else, and yet hope to be understood?

The unifying quality of so-called “verbal ironies” (as opposed, for instance, to dramatic ironies) is that typically they can be “decoded” by understanding that the speaker’s actual position and the speaker’s sarcastically adopted position differ in crucial ways. Sarcasm is the paradigm case of concise verbal irony. Swift’s *Modest Proposal* is the classic extended example. Swift’s narrator proposes that Irish babies should be bred and slaughtered as meat for human consumption. The successful interpreter understands that the implied Swift himself proposes no such thing; instead he is presenting a savage and satirical criticism of the cruelty of the English landlord class.

We consider traditional distinctions between different rhetorical figures theoretically justified so long as they involve distinct cognitive strategies. In this light, what actually counts as an instance of verbal irony depends in part on how one understands the phenomenon to work.

A classical view, going back at least to Cicero and Quintilian, is that an ironic utterance is one that means the opposite of what it says: as Johnson put it in his dictionary, irony is “a mode of speech in which the meaning is contrary to the words” (1755: 1134). Neither Johnson nor Quintilian is fully explicit as to just what sorts of meanings should count as contrary here, but the most common assumption seems to be that an ironic meaning should be the polar contrary of the expressed meaning – for example, saying *He’s a fine friend* to mean “He’s a lousy friend,” or *Rotten meat! How delightful!* to mean “Rotten meat! How disgusting!”

The problem with this account is that there are many utterances that intuitively count as ironic, but in which the speaker does not mean anything like the opposite of what he says. Thus Gibbs (1986) cites the example of a disgruntled driver who exclaims, “I love people who signal,” after another driver turns into his lane without signaling. Clearly, the speaker here does not mean that he hates people who signal, nor that he loves people who do not signal. The proposition that the speaker has expressed is in fact just what the speaker believes – the

irony here is not in what the speaker says, but rather in the fact that he should choose this occasion to say it.

Approaches based on echoic mention, such as Sperber and Wilson (1981), assimilate verbal irony to the broader phenomenon of the use/mention distinction. The basic idea is that a speaker, in producing an ironic utterance, mentions an expressed proposition rather than using it. This theory is often misunderstood as implying that irony always involves some sort of literal echo or quotation of a previous utterance. However, the claim is only that the speaker presents an expressed proposition as the kind of thing one might say. The thought being echoed may not have been expressed in an utterance; it may not be attributable to any specific person, but merely to a type of person, or people in general; it may be merely a cultural aspiration or norm (Wilson and Sperber 1992: 60). Echoed thoughts may be a reflection of actual utterances, or of hopes, desires, attributed thoughts, or cultural norms.

One advantage of this approach is that it offers a neat explanation of why verbal irony is easier when it takes the form of a positive comment on a manifestly negative situation (e.g. *Brilliant!* as a comment on a boneheaded action, or *Lovely weather!* as a comment on a sudden downpour) than in the opposite case, when it takes the form of a negative comment on a positive situation (*What a jerk!* of someone who has been very helpful, *What an idiot!* of a Nobel Prize winner, or *What foul weather!* said on a sunny day). According to the echoic mention account, the former examples work in almost any context, because they echo positive cultural norms; the latter only work where there is some accessible prior utterance or expectation that they can echo:

- (3) Peter: The weather is going to turn foul. I have a nasty feeling about that picnic.

[Peter and Mary go to the picnic planned with their friends. The sun shines.]

Mary: Pretty foul weather, all right!

While the theory has a number of virtues, and does seem to describe an important variety of ironic utterances, it remains open to criticism. It is not clear that echoic mention is always a necessary condition for irony. It is difficult to see how the irony in a satire like *A Modest Proposal*, for example, is predicated sheerly on echoic mention: there is no cultural norm, prior utterance, or expectation for eating Irish babies being echoed here. Nor is echoic mention by itself sufficient to explain irony. Giora (1995: 248) points out that utterances like (4b) are both echoic and disparaging, but not ironic, while utterances like (5b) are indeed ironic.

- (4a) Dina: I missed the last news broadcast. What did the Prime Minister say about the Palestinians?

- (4b) Mira (with ridiculing aversion): That we should deport them.

- (5a) Dina: I missed the last news broadcast. What did the Prime Minister say about the Palestinians?
- (5b) Mira: That we should host them in 5-star hotels in Lebanon.

Alternate approaches based on pretense (Clark and Gerrig 1984; Clark 1996; also Kreuz and Glucksberg 1989; Kumon-Nakamura *et al.* 1995) can handle both of these cases. These accounts propose that verbal irony occurs when the speaker pretends to some attitude that she does not really feel and expects her audience to recognize that it is a pretense. Swift's *Modest Proposal* thus can be explained as follows: the author is pretending an attitude, that of sincerely proposing that Irish children be slaughtered and sold as meat for human consumption. The pretense theory can also explain the lack of irony in (4), by pointing out that Mira is in no way pretending to hold a view or to perform a role that is not her own. Mira really does believe that the Prime Minister said that the Palestinians should be deported.

However, just as there are many kinds of mention that are not ironic, there are many forms of non-ironic pretense. Further, as pointed out by Sperber (1984), there is something unsatisfying, at the very least, about the pretense account when it comes to ironic utterances that are manifestly self-contradictory, as in (6), which is a perfectly well-formed piece of sarcasm, but not a very coherent act of pretense.

- (6) Oh, yes, how right you are; this disgusting state of affairs is just delightful.

Since the 1980s, a number of hybrid accounts have been proposed. Giora (1995) suggests a modification of the Gricean pragmatic analysis of irony, in which she argues that irony is a kind of *indirect negation* produced by an apparent violation of the cooperative requirements for discourse coherence, specifically produced by a clash between the most salient meaning of an utterance and the degree of informativeness that is appropriate for that utterance, given its discourse context. An ironic utterance thus highlights the difference between an actual state of affairs and some implicated message about a more desirable state of affairs. Attardo (2000) also takes a neo-Gricean view of irony, explaining it as a case of *relevant inappropriateness*: if an utterance is both inappropriate and relevant to the context, it will count as ironic.

However, some examples that are handled fairly straightforwardly by other approaches turn out to be problematic for these theories. The condition of relevant inappropriateness covers many cases that do not seem to qualify as irony, such as polite understatement. Meanwhile, so-called "purely echoic" cases, like the ironic "Pretty foul weather, all right!" in (3), do not conjure up a more desirable state of affairs, as predicted by the indirect negation account, and indeed Giora (2003) argues that these cases represent a fundamentally different phenomenon than the kinds of sarcasm explained by her theory. This

is a reasonable line to take on the issue, but since our project in this chapter is to account for the features that verbal irony does share with other kinds of irony, it means that these theories will not be sufficient for our purposes.

1.4 Irony as a viewpoint phenomenon

The most common way of thinking about irony is as an operation on a focused proposition. We suggest that irony is instead an operation on the way a focused proposition is accessed and viewed: it is a way of *construing* an expressed proposition or an observed scene. That is, ironic utterances, like ironic situations, are distinguished by the sort of interpretive process they evoke. In particular, we claim that the interpretation of irony involves three key elements:

1. a layered configuration of mental spaces;
2. a shift in attention from an inner to an outer layer – a “zooming-out”;
3. a dynamic blended construal of an event from two distinct viewpoints.

We see these elements as essential not only to the appreciation of verbal irony, but also to the experience of situational ironies, ironic sensibilities, and structural ironies that are built up over the course of an extended narrative.

In speaking here of irony as an interpretive process, we seek to highlight the dynamic and unsettling nature of the ironic experience, but we do not wish thereby to suggest that there is only one way this experience can be achieved. Canonically, perhaps, the elements of an ironic interpretation are built online – the interpreter starting with a view that proves somehow inadequate (the ironized view), and then adjusting to a new, more satisfying (ironic) viewpoint. In practice, however, readers and interlocutors may approach the act of interpretation with an ironic attitude right from the start, deploying an ironic mental space configuration as a default mode of understanding, as, for example, in that peculiarly sophisticated attitude that takes pleasure in the enjoyment of camp (cf Sontag 1964), or in the simultaneous appreciation of several mutually exclusive explanations of the world that Schlegel described as the ironic sensibility of Romanticism.

A precise articulation of this proposal will depend on a few technical details of Mental Spaces Theory. First, every mental space configuration canonically includes a Base, a Viewpoint, an Event, and a Focus, although a single space can serve as more than one of these at the same time (Cutrer 1994; Fauconnier 1997). The Base Space serves as the subjectively construed Ground of interpretation. The Viewpoint Space is the space from which conceptual content is accessed. The space in Focus is the space on which attention is concentrated. The Event Space is the one in which an event takes place.

Mental spaces also have *status* with relation to other spaces. These relative statuses can be hierarchical: mental spaces can be embedded within other mental spaces. They can be temporal: a space can be figured as past with respect to

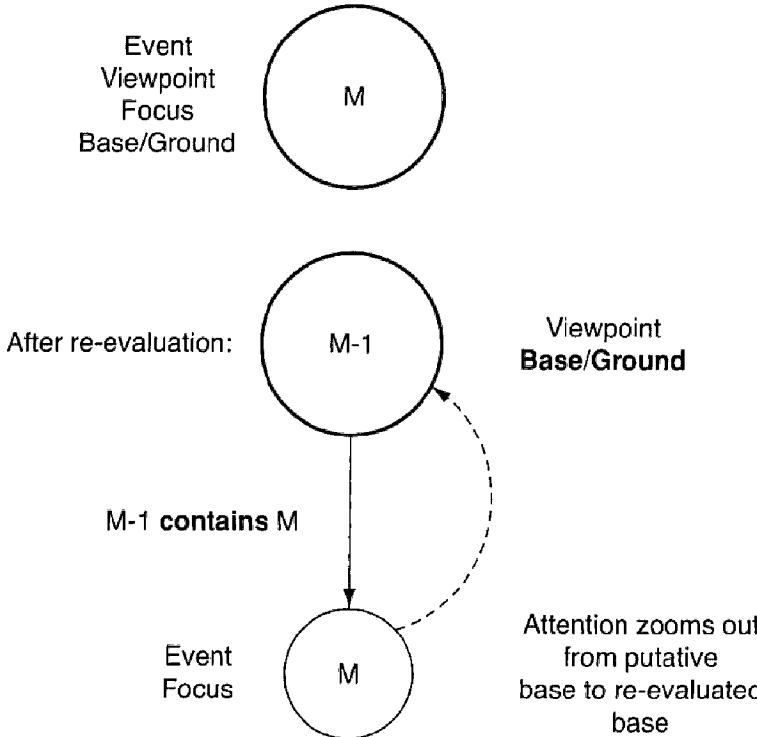


Figure 1.1

one space and present with respect to another. They can also be epistemic: one space can have the status of fact with respect to another space, for example, or prediction.

Finally, mental spaces are also potential objects of joint attention, as speakers and hearers try to coordinate their mental representations and share attention to various aspects of those representations. Conceptualizers can shift their attention within mental space configurations, moving their viewpoint from space to space.

In these terms, then, we propose that irony is a figure of attention flow consisting of three minimal steps:

1. The presentation of a proposition p in a Focus Space F from a Viewpoint Space V (where F and V may, but need not be identical).
2. The assessment of some conflict or incongruity between p and some set of assumptions that are accessible in the context of p .
3. The reconstrual of F , V , and p from a higher Viewpoint Space, V' , in a way that resolves any inconsistencies.

The effect of this process is that an ironic utterance presents a proposition almost simultaneously from at least two distinct points of view: an ironized viewpoint (V) and an ironic viewpoint (V').

The basic ironic configuration, illustrated very schematically in Figure 1.1, involves a perceived incompatibility between a profiled Event and some set of

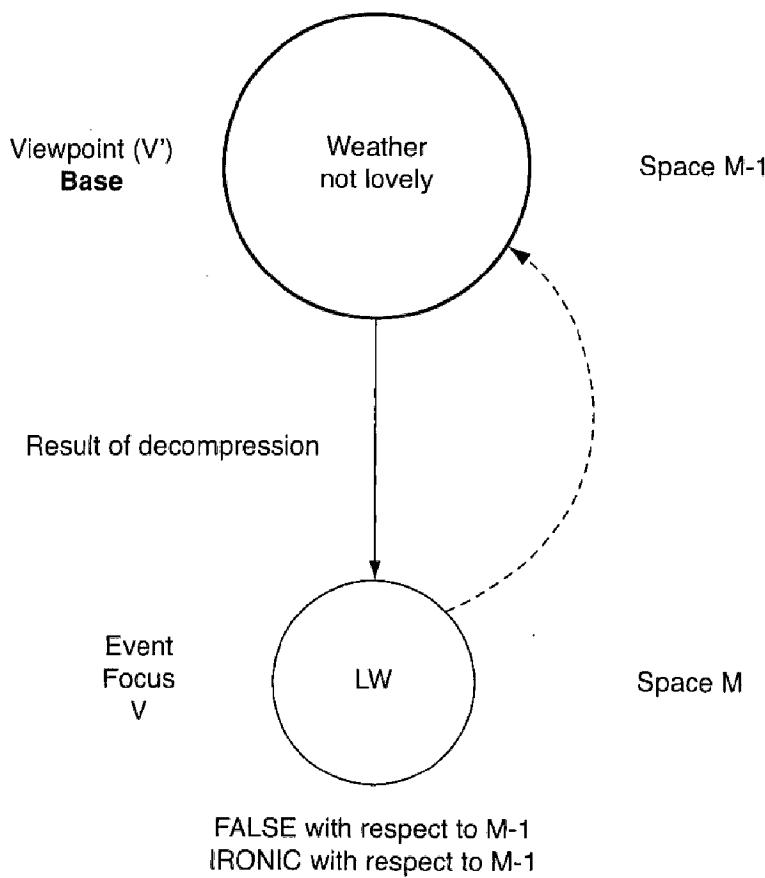


Figure 1.2

tacit assumptions about the common ground. Verbal irony involves a mental space configuration in which what is said (i.e. what is manifestly communicated in a speech act) is somehow incongruent with the conditions of its own utterance. In order to resolve this incongruity, the interpreter must effectively reconceptualize the context of utterance, thus prompting the construction of a new viewpoint from a new, re-evaluated Base (M-1), which contains the original Event Space (M).

Figure 1.2 illustrates how this works with a relatively straightforward sarcastic utterance: "Lovely weather," said in response to a sudden downpour. Here, the utterance *Lovely weather* sets up space M, containing the proposition that the weather (presumably the weather at hand, of which speaker and hearer share awareness in their common ground) is lovely.

Normally, space M might be understood as a factual belief space, representing the speaker's current view of reality, but in this case the reality of the situation is plainly at odds with the expressed proposition. The ironic interpretation arises when a hearer both recognizes this incongruity and thereby recognizes that the expressed viewpoint of the utterance is not in fact that of the speaker. This

recognition may be achieved by pragmatic inferencing alone, in which case a hearer may first consider and then reject a literal interpretation, or it may be facilitated by paralinguistic cues involving facial expressions or tone of voice, in which case it may be virtually immediate.¹

Either way, the result is a mental space configuration in which the expressed viewpoint and the speaker viewpoint are somehow disentangled. This disentangling requires a second Base Space, M-1, which serves as a kind of higher ground from which space M can be reconsidered. In Figure 1.2, the solid arrow marks the status of M as subordinate to M-1, while the dashed arrow represents the flow of attention zooming out from M to the new higher ground. Because the irony here is intentional and is recognized as such, speaker and hearer share this conception of the entire ironic configuration. M retains an associated viewpoint, V, which speaker and hearer “look down on” from their shared ironic viewpoint, V’. The zoom-out effect of irony is a form of alienation from this lower-level viewpoint and from those who hold it; this is the reason that irony often has a victim.

Again, it is worth emphasizing that in our view irony does not require an interpreter to first entertain and then reject a literal interpretation, but may come more or less instantaneously in a pre-compiled, complex space configuration. But whether an irony is built online or pre-compiled as part of an ironic sensibility, the experience consists of the simultaneous apprehension of two incompatible viewpoints, one of which is rejected and in effect looked down on. The difference between these two ways of experiencing irony is roughly analogous to the distinction drawn by Langacker (1987: 144–5) between the two ways of construing an event that unfolds in time, either by scanning the stages of the event sequentially in processing time or by imagining the event all at once, with a summary scanning of its subparts.

This account is compatible with the “distancing” viewpoint configuration described by Vandelanotte (this volume), in which certain examples of indirect speech and thought serve to report or present another person’s discourse, while keeping the deictic center of person, place, and time all firmly with the current speaker. The claim is that this kind of speech and thought representation involves mental space evocation (Dancygier and Sweetser 2000, 2005) rather than embedding – that is, it requires the interpreter to look for an appropriate space, accessible within the current discourse configuration, rather than having to create an all-new, embedded space. Such cases are thus echoic, in the sense of Sperber and Wilson (1981), but they are not always, or even usually, ironic.

They can be, though. Vandelanotte observes that the effect of these kinds of utterances is distancing, because of the unexpected confluence of contextual signals that indicate that there are two speech situations in play – both the current and the reported – without any explicit linguistic indication of this complexity. Moreover, different versions of this configuration can invoke more or

less associative or dissociative attitudes between current and reported speaker. Vandelanotte connects more dissociative configurations with the expression of literary irony and sarcasm, which is just the way we see it as well.

In the last few years, other researchers have proposed accounts of irony within Mental Spaces Theory. These theories have generally focused on verbal irony, and have claimed that the underlying structure of sarcasm depends crucially on counterfactual thinking. Coulson (2005: 136) argues that in “sarcastic language, the listener is confronted with a blend that she must unpack into two input spaces: an *expected reaction space* and a *counterfactual trigger space*.” Kihara (2005: 236) similarly suggests that “ironical remarks have their effects by referring to a counterfactual mental space of expectation without any distinct space builders.” We find these accounts appealing in many ways, but note that there are good reasons not to pin an account of irony too tightly to counterfactuality. Kumon-Nakamura *et al.* (1995), among others, point out that there are many kinds of insincerity that are not counterfactual, but are perceived as ironic, such as statements like, “You sure know a lot,” directed to someone who is indeed knowledgeable, but being an obnoxious show-off about it.

In our account, irony involves a special kind of viewing arrangement – a view of a Viewpoint in a complex mental space configuration. Like Coulson, we thus see irony as involving a kind of a blend; however, in our account, the blending takes place not in the Focus Space where a proposition is expressed, but rather in the Viewpoint from which an expressed proposition is accessed. The experience involves the same propositional content being accessed simultaneously from two incompatible viewpoints, one of which encompasses the other. This is what makes irony different from the experience of simply being in two minds about something. The ironic effect itself appears to arise from the way a whole construal (including an expressed proposition in focus from some viewpoint) itself becomes an object of construal. This is similar to what happens in the cases of mental space alignment that Vandelanotte (this volume) describes, in which “the represented speaker’s discourse ends up submerged in that of the current speaker.”

What happens with the experience of irony, then, is an adjustment from this blend to a new Viewpoint (zooming out) that is construed as both separate and superordinate: in Haiman’s (1998: 80) terms, distinguishing “the difference between a behaving and a scrutinizing self.” This viewpoint adjustment is a kind of decompression (Fauconnier and Turner 2002).

This account has several merits. It generalizes insights of previous theories of verbal irony to handle other kinds of ironic effects, and it allows for a more articulated treatment of layered ironies and ironies that seem to violate the logic of stacked levels. Irony can arise wherever a discourse structure provides a multilevel network of mental spaces. Where verbal irony involves a mismatch between what a speaker says and some set of mutually manifest

assumptions in the common ground, cosmic irony involves a mismatch between facts and expectations at the level of an event itself, and dramatic irony involves a mismatch between facts at the event level and beliefs at a higher narrative level.

It also suggests that there is a reason why some genres seem particularly to lend themselves to irony. The ironizing viewpoint constitutes a new common ground between the interpreter and some implicit or explicit interlocutor, or fellow-observer. This higher ground may be constructed in the process of interpretation, but it may also be already established by the discourse situation in which the irony occurs. Narrative fiction, for example, comes with a (usually overt) narrator, who is understood as distinct from, if sometimes closely aligned with, an implied author.

1.5 A stable structural irony: Huck Finn reverts to his wicked ways

As we have seen, of the phenomena commonly described as “ironic,” verbal irony is by far the most frequently discussed in the cognitive science and linguistics literature. Verbal ironies are intentional, and understanding them relies at least in part on recognizing a difference between what is said and the proposition and attitude that the speaker intends to convey.

Cosmic or situational irony, by contrast, is perpetrated by the universe, rather than by a speaker. It arises from twists of fate in which hopes and expectations are overturned in some fundamental way. To die of thirst surrounded by water, or to lose the thing you love best through the very actions that you take in order to preserve it, is to be the victim of a cosmic irony. To be subject to such an irony is perhaps poignant, but also absurd, or at least faintly ridiculous. There is a sense that destiny has conspired to play a joke on the irony’s unhappy target, as if fate, the universe, or some other omniscient agent were in some way the author of the irony.

The complex viewpoint involved in appreciating cosmic irony arises not from a conflict between an expressed proposition and the real communicative intentions of its utterer, but between something like the apparent “intentions” of the universe and the futile original intentions or expectations of the irony’s victim, so that the former make a mockery of the latter. Appreciating the irony in such a circumstance requires a certain amount of detachment: again, it calls for the interpreter to take a particular view of a viewpoint. In taking one’s own circumstances to be ironic, one must momentarily step outside oneself, to indulge in a perhaps rueful or bitter chuckle at one’s own expense.

A similar process underlies the interpretation of stable dramatic and structural ironies – the kinds of irony that play a central role in *Tristram Shandy*, *Oedipus Rex*, *Mansfield Park*, or *The Adventures of Huckleberry Finn*. A stable irony (Booth 1974) is one that can be grasped in one go, without the prospect of

an infinite regress of further ironic undermining. In these cases, the particular disparity is between what a narrating character or persona *takes herself to mean* and the deeper or higher significance that the implied author seems to intend the reader to understand. As in verbal irony, these texts say one thing and mean another, but here the double meaning arises from the presentation of a character whose account of events is clearly unreliable in some way. She is untrustworthy or naive; some failing that the implied author of the text recognizes and does not share impairs her judgment: prejudice, perhaps, or limited perspicacity, or personal interest.

In interpreting these ironies, the reader can make use of an already established detached or superior Viewpoint, or decompress a blend, to make an ironizing Viewpoint/Ground newly available. For example, the moment in *The Adventures of Huckleberry Finn* when Huck decides to help Jim, even though he believes that doing so is a sin, involves a sustained clash between the focalizing viewpoint and what the reader takes to be the case, prompting the reader to zoom out to the higher ground associated with the implied author. Here is the passage in question (Twain 2008 [1884]: 143):

I felt good and all washed clean of sin for the first time I had ever felt so in my life, and I knowed I could pray now. But I didn't do it straight off, but laid the paper down and set there thinking – thinking how good it was all this happened so, and how near I come to being lost and going to hell. And went on thinking. And got to thinking over our trip down the river; and I see Jim before me all the time: in the day and in the night-time, sometimes moonlight, sometimes storms, and we a-floating along, talking and singing and laughing. But somehow I couldn't seem to strike no places to harden me against him, but only the other kind. I'd see him standing my watch on top of his'n, 'stead of calling me, so I could go on sleeping; and see him how glad he was when I come back out of the fog; and when I come to him again in the swamp, up there where the feud was; and such-like times; and would always call me honey, and pet me and do everything he could think of for me, and how good he always was; and at last I struck the time I saved him by telling the men we had small-pox aboard, and he was so grateful, and said I was the best friend old Jim ever had in the world, and the *only* one he's got now; and then I happened to look around and see that paper.

It was a close place. I took it up, and held it in my hand. I was a-trembling, because I'd got to decide, forever, betwixt two things, and I knowed it. I studied a minute, sort of holding my breath, and then says to myself:

"All right, then, I'll go to hell" – and tore it up.

It was awful thoughts and awful words, but they was said. And I let them stay said; and never thought no more about reforming. I shoved the whole thing out of my head, and said I would take up wickedness again, which was in my line, being brought up to it, and the other warn't. And for a starter I would go to work and steal Jim out of slavery again; and if I could think up anything worse, I would do that, too; because as long as I was in, and in for good, I might as well go the whole hog.

Huck is a victim of this irony, although not in the same way as the object of a satire is its victim. Here, the effect is sympathetic, even poignant; Huck is a victim of his own ironic circumstances, and both reader and implied author see that irony, crediting Huck with additional virtue thereby. Huck's conviction that his decision is damnable makes its true laudability all the more evident.

The ostensible viewpoint in this passage lies squarely with the narrator, Huck. But the reader knows, through information built up over the course of the novel, that the perspective implicated in statements like *how near I come to being lost and going to hell* is more complex. Here, Huck is re-enacting discourse and professing beliefs that properly belong to someone else – to the pious and small-minded Miss Douglas, to the prevailing views of white Southern society that surround him – and the distance between Huck's true convictions and the beliefs that he echoes is made increasingly explicit over the course of this passage. So far, this state of affairs looks strikingly like the examples of dissociative, ironic DIST discussed in Vandelanotte (this volume). However, in this case, the putative speaker is patently unaware of the distance between the two positions. This is not a verbal irony that can be ascribed to the speaker himself. Something else is going on.

This kind of structural irony involves much the same kind of interpretive work as verbal irony, but with an additional layer. Something is said. The reader who enjoys the irony then appreciates that an attitude is being conveyed that differs from what is being said, but that the intention behind this double articulation cannot be ascribed to the putative speaker. To resolve the clash of perspectives, the reader must take recourse to a higher level of the discourse situation, as illustrated in Figure 1.3. The normally latent perspective of the implied author is recruited to provide a viewpoint from which the subjective narrating viewpoint can be newly construed as an object of conceptualization. The "I" of the higher Viewpoint Space looks upon, rather than participating in, the vexed viewpoint represented by the "I" of Huck's Ground.

Ordinarily, readers are not continually consciously aware of the implied author as they proceed through a text, and the more successful a piece of fiction, the more fully immersed readers are in the deictic frame associated with the current speaker. That is, much of the time when we are reading a narrated fiction, we are inhabiting a blend in which the narrator *is* the (one and only) speaker of the narrative.

In the blend, the narrator partakes of all the aspects of the speaker role distinguished by Goffman (1981) – he is the *animator*, the person who does the uttering, the *author* or composer of the utterance, and the *principal* whose position is established by the words that are uttered. Appreciating a structural irony requires a decompression of this blend. The new construal involves a fresh awareness that both narrative viewpoint and the role of *current speaker*

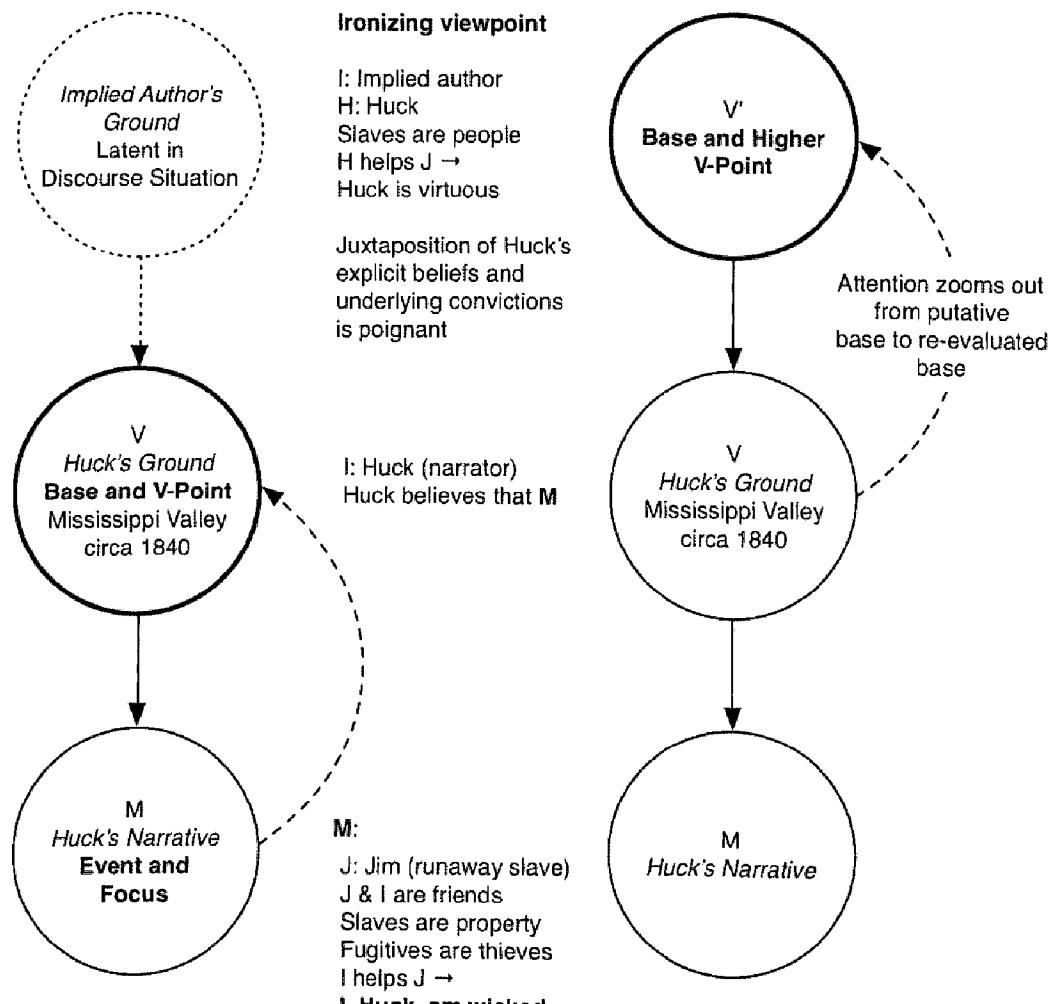


Figure 1.3

are complex. This decompression of the conflated authorial and narrating viewpoints activates, or reintroduces, a higher-level Base Space identified with the discourse situation and viewpoint of the implied author.

From this perspective, the ironizing reader can appreciate that what Huck takes himself to mean – indeed, what he takes himself to believe – is at odds with the true significance of his actions and motivations, and that this inconsistency is in itself significant and intentional. The discrepancy between Huck's presentation of his own situation, including his own discourse situation, and what the reader is given to understand the true situation to be generates a structural irony. In recognizing the source of this discrepancy, readers shift Viewpoint and Ground to a more distant perspective, creating a sense of both ironic distance and complicity with the implied author.

1.6 Instabilities

Just as not all ironies are verbal, neither are all ironies resolvable. The viewpoint account of irony can also handle unstable ironies, “in which the truth asserted or implied is that no stable reconstruction can be made out of the ruins revealed through the irony” (Booth 1974: 240); “ironies that will turn into infinities if pursued” (*ibid.*: 246). Sarah Silverman’s routine, described in the introduction and recapitulated in (7), induces one kind of unstable irony.

- (7) “Everybody blames the Jews for killing Christ, and then the Jews try to pass it off on the Romans. I’m one of the few people that believe it was the blacks.”

Unstable ironies set up zoom-out configurations in which the potential views of viewpoints threaten to proliferate uncontrollably. Silverman’s unsettling humor invites the interpreter to vacillate among these proliferating viewpoints, while also recognizing that the author of the irony intended this unsteady view: the interpreter must take a view of the fluctuating view of a viewpoint. This interpretive process involves a mental space configuration like the one illustrated in Figure 1.4.

The joke sets up a speaker’s reality space, S1, and three belief spaces: M1, M2, and M3. The punchline, *I’m one of the few people that believe it was the blacks*, provokes a clash that seems at first blush to be a classic verbal irony. Just as with the “Lovely weather” example, an assertion that would ordinarily count as an expressed belief of the speaker appears under utterance conditions that conflict with this interpretation.

The surprise makes the audience laugh, and the recognition of the inappropriateness prompts hearers to decompress the blend of expressed viewpoint and speaker viewpoint, zooming out to a new Base Space. This is S2, the “Real Sarah Silverman” space. S1 retains its associated viewpoint, that of the Sarah Silverman persona, which the comedian and the audience can look down on together from a shared ironic viewpoint. So far, so sarcastic.

However, while the irony here may be what Booth (1974) calls “locally” stable, its status with respect to the larger discourse is less clear. When someone looks at the rain and says “Lovely weather!” in disgusted tones, there is no worry that she will turn out to have really meant that the weather was indeed lovely.

Similarly, nothing in the discourse context or our background knowledge suggests that we need to worry that it is actually the case that Sarah Silverman truly believes that black people were responsible for the death of Jesus. At the same time, her intended message does not seem to be straightforwardly sarcastic. The point is not to covertly express something like, “As if black people killed Jesus! Can you believe the kind of bozo who would think such a thing?” There is no such bozo.

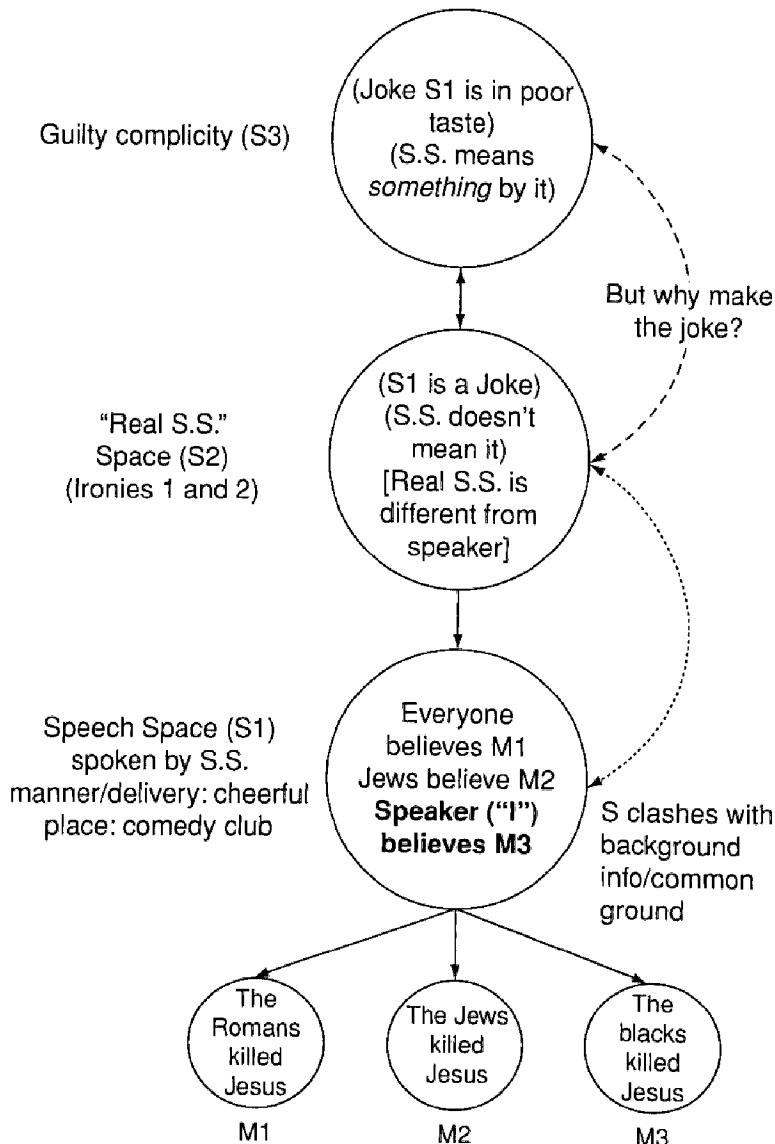


Figure 1.4

The intended message, indeed, is not clear. If irony typically involves something like relevant inappropriateness, one might well wonder what the relevance of this inappropriate statement might be. In other words: why make the joke? Why use the offensively marked noun phrase “the blacks”? And why should the audience take pleasure in it? This doubt about the motivations of a speaker and hearer who partake of the viewpoint in S2 motivates a fresh zooming out to a new critical view of a viewpoint in S3. But there is no way to be sure of a correct final interpretation.

If we take the instability to be intentional, speaker and hearer share this conception of the entire ironic configuration, and Silverman counts as an

especially masterful ironizer. If we do not take it to be intentional, we end up with a similar mental space configuration, but a very different opinion of Silverman, because the zoom-out effect is a form of alienation from the lower-level viewpoint and from those who hold it. In that case, the reconstrual is not one of complicity but of estrangement. At this point, we have the option of stabilizing the irony by rejecting the joke. But as long as we allow the possibility that the performance is ironic and that the joke is funny, audience and speaker alike are implicated in the viewpoints of both S2 and S3, and the ultimate stance remains indeterminate, as indicated by the double-headed dashed line in Figure 1.4.

Ambiguous ironic intentions like Silverman's are not the only possible source of ironic instability, and we would like, finally, to discuss the structurally unstable irony generated by the short story "Borges and I" (Borges 1967 [1964]). This story presents a situation in which, by virtue of their expression, the thoughts and characteristics of Borges become those of the public persona "Borges." Thus everything the narrator tells us about the relationship between Borges and "Borges" accrues not to Borges but to "Borges," and the more Borges tries to express this irony, the worse it gets:

The other one, the one called Borges, is the one things happen to . . . I like hourglasses, maps, eighteenth-century typography, the taste of coffee and the prose of Stevenson; he shares these preferences, but in a vain way that turns them into the attributes of an actor. It would be an exaggeration to say that ours is a hostile relationship; I live, let myself go on living, so that Borges may contrive his literature, and this literature justifies me.

... Years ago I tried to free myself from him and went from the mythologies of the suburbs to games with time and infinity, but now those games belong to Borges and I will have to think up something else. Thus is my life a flight and I lose everything and everything belongs to oblivion, or to him.

I do not know which of us has written this page.

If ever there was an irony that would "turn into infinities if pursued," this is it. The basic situational irony takes as its ingredients an experiencer (A), a situation, and an expectation. In the Focus/Event Space, A acts with intention to bring about some result p . In the higher Viewpoint Space, A's action brings about $\sim p$. "Borges and I" invokes a similar causal structure and combines it with a potentially infinite regress of ironic reconstruals.

The action that gives rise to Borges' dilemma – presenting himself in writing – and the articulation of that dilemma are one and the same. As a result, the ordinary zoom-out construal involved in appreciating a situational irony gives rise each time to a new dilemma and a new need to construct a higher-level Viewpoint/Ground. The irony can never be resolved. This

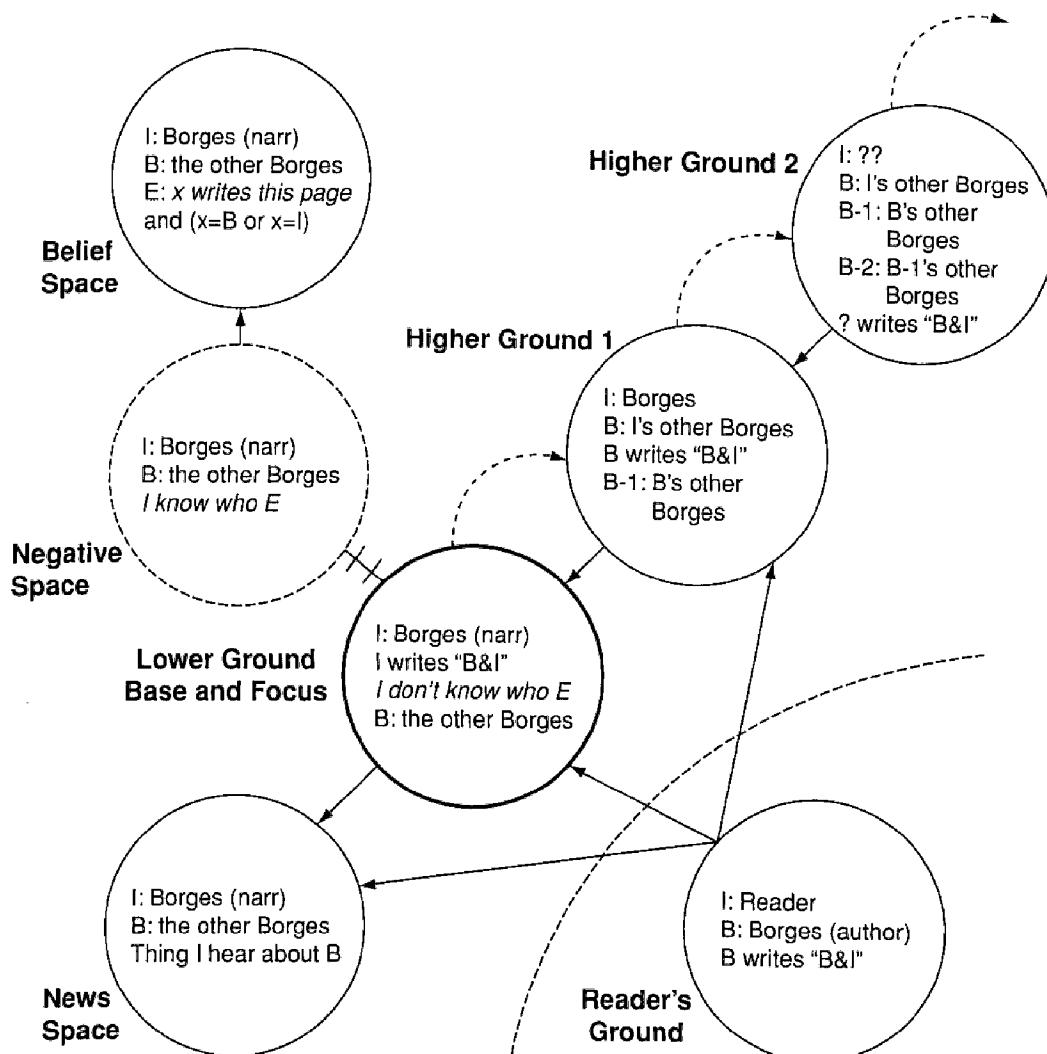


Figure 1.5

viewpoint configuration generates the *mise en abyme*, or infinite regress, that for many postmodern scholars characterizes the very height of irony: irony as a potent, rapidly proliferating, perhaps even uncontrollable, means by which stable meaning is undermined and made permanently uncertain.

We would note, as well, that constructing this kind of ironic construal unavoidably involves active and conscious interpretive labor. Interpreting a canonical sarcastic utterance in a context where conventional discourse goals support that interpretation is relatively rapid and automatic for adults (see, for example, Kreuz and Link 2002). In novel and complex cases like these, however, noticeable introspection and explicit puzzling over the speaker's "genuine" intentions are required before the full unsettling instability emerges.

1.7 Conclusion

Irony in all its forms is a figure of subjectivity. More precisely, it is a figure of *desubjectification*: the process whereby conceptual contents that are first construed subjectively are reconstrued as an object of conceptualization. The possibility for irony is, in effect, a natural consequence of the narrative mind: irony arises from the fact that any situation we encounter is subject to interpretation both as something that happens and as something that is represented. However, irony's operation is constrained by the high costs it puts on processing and a consequent need for highly ritualized discourse contexts (cf Haiman 1998).

Our approach complements existing theories of irony by viewing it as a variety of interpretive experience, and by focusing on the close – and, we argue, natural – relations between different sorts of verbal, situational, and structural ironies, to show why irony is difficult, why it is unsettling, why it typically has a victim, and why it is subject to proliferation in certain discourse contexts.

References

- Anderson, Sam. 2005. Irony maiden: how Sarah Silverman is raping American comedy. *Slate*. www.slate.com/id/2130006 (accessed October 13, 2011).
- Attardo, Salvatore. 2000. Irony as relevant inappropriateness. *Journal of Pragmatics* 32, 793–826.
- Booth, Wayne C. 1974. *The Rhetoric of Irony*. Chicago and London: University of Chicago Press.
- Borges, Jorge Luis. 1967 [1964]. Borges and I. Translated from German by James E. Irby. In Anthony Kerrigan (ed.). *A Personal Anthology*. New York: Grove Press, 200–1.
- Bruce, Lenny. 1963. *How to Talk Dirty and Influence People*. New York: Fireside.
- Butler, Harold E. 1921. *The Institutio Oratoria of Quintilian*. 4 vols. London: William Heinemann Ltd, vol. III.
- Clark, Herbert H. 1996. *Using Language*. Cambridge University Press.
- Clark, Herbert H. and Richard J. Gerrig. 1984. On the pretense theory of irony. *Journal of Experimental Psychology: General* 113, 121–6.
- Colebrook, Claire. 2004. *Irony: The New Critical Idiom*. London and New York: Routledge.
- Coulson, Seana. 2005. Sarcasm and the space structuring model. In Seana Coulson and Barbara Lewandowska-Tomaszczyk (eds.). *The Literal and Nonliteral in Language and Thought*. New York: Peter Lang, 129–44.
- Cutrer, Michelle. 1994. Time and tense in narrative and in everyday language. Ph.D. dissertation, University of California, San Diego.
- Dancygier, Barbara and Eve Sweetser. 2000. Constructions with *if*, *since* and *because*: causality, epistemic stance, and clause order. In Elizabeth Couper-Kuhlen and

- Bernd Kortmann (eds.). *Cause – Condition – Concession – Contrast*. Berlin: Mouton de Gruyter, 111–42.
2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Empson, William. 1947. *Seven Types of Ambiguity*. 2nd edn. New York: New Directions.
- Fauconnier, Gilles. 1985. *Mental Spaces*. Cambridge University Press.
1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think*. New York: Basic Books.
- Fowler, Henry Watson. 1926. *A Dictionary of Modern English Usage*. 1st edn. Oxford University Press.
- Gibbs, Raymond W. 1986. On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General* 115, 3–15.
- Giora, Rachel. 1995. On irony and negation. *Discourse Processes* 19, 239–64.
2003. *On Our Mind: Salience, Context, and Figurative Language*. Oxford University Press.
- Goffman, Erving. 1981. *Forms of Talk*. University of Pennsylvania Press.
- Haiman, John. 1998. *Talk Is Cheap: Sarcasm, Alienation, and the Evolution of Language*. Oxford University Press.
- Israel, Michael. 2006. Saying less and meaning less. In Betty Birner and Gregory Ward (eds.). *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*. Amsterdam: John Benjamins, 143–62.
- Johnson, Samuel. 1755. *Dictionary of the English Language*. 2nd edition. 2 vols. London: printed by W. Strahan, for J. and P. Knapton; T. and T. Longman; C. Hitch and L. Hawes; A. Millar; and R. and J. Dodsley, vol. I.
- Kihara, Yoshikio. 2005. The mental space structure of verbal irony. *Cognitive Linguistics* 16:3, 513–30.
- Knox, Norman. 1961. *The Word Irony and Its Context, 1500–1755*. Durham NC: Duke University Press.
- Kreuz, Roger T. and Sam Glucksberg. 1989. How to be sarcastic: the echoic reminder theory of verbal irony. *Journal of Experimental Psychology, General* 118, 374–86.
- Kreuz, Roger J. and Kristen E. Link. 2002. Asymmetries in the perception of verbal irony. *Journal of Language and Social Psychology* 21, 127–43.
- Kumon-Nakamura, Sachi, Sam Glucksberg, and Mary Brown. 1995. How about another piece of pie: the allusional pretense theory of discourse irony. *Journal of Experimental Psychology, General* 124, 3–21.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford University Press.
1999. *Grammar and Conceptualization*. Berlin and New York: Mouton de Gruyter.
- Silverman, Sarah. 2005. *Jesus is Magic* [film]. Black Gold Films.
- Sontag, Susan. 1964. Notes on camp. In Susan Sontag. *Against Interpretation and Other Essays*. New York: Farrar, Straus & Giroux.
- Sperber, Dan. 1984. Verbal irony: pretense or echoic mention? *Journal of Experimental Psychology, General* 113, 130–6.
- Sperber, Dan and Deirdre Wilson. 1981. Irony and the use-mention distinction. In Peter Cole (ed.), *Radical Pragmatics*. New York: Academic Press, 295–318.

1998. Irony and relevance: a reply to Seto, Hamamoto and Yamanashi. In Robin Carston and Seiji Uchida (eds.), *Relevance Theory: Applications and Implications*. Amsterdam: John Benjamins, 283–94.
- Twain, Mark. 2008 [1884]. *The Adventures of Huckleberry Finn*. Repr. edn. London: Puffin Books.
- Wilson, Deirdre and Dan Sperber. 1992. On verbal irony. *Lingua* 87, 53–76.

2 Subjectivity and upwards projection in mental space structure

Lilian Ferrari and Eve Sweetser

2.1 A Mental Space approach to subjectivity

This chapter proposes an analysis of historical processes of meaning *subjectification*, in terms of viewpoint relations in a dynamic network of mental spaces. We argue that defining subjectification in terms of mental space structure allows added precision both in identifying subjective aspects of meaning and in assessing *degrees* of subjectivity – and hence in accurately describing directions of meaning change. We shall begin by giving our definition, in the context of the extensive scholarship on subjectification, and continue with examination of particular cases.

Linguists' understanding of viewpoint has been advanced by, among other things, almost thirty years of research on subjectivity and subjectification. Traugott (1982, 1989, 1995), Langacker (1987, 1990, 1991), Hopper and Traugott (1993), Traugott and Dasher (2002), and others have in different ways used these words to define the relationship between the referential meaning of a word and the understanding of the speech setting – the speaker's and hearer's physical setting, beliefs, and interaction. Research also points (Sweetser 1990; Dancygier and Sweetser 2005; Sanders *et al.* 2009) to a cline of subjectivity between these aspects of the speech setting: the physical setting is more "objective" than the speaker's and hearer's mental states and interaction. Traugott has proposed the generalization that meanings can move towards greater subjectivity, but not towards decreased subjectivity. Relatively high subjectivity is present in grammatical meaning domains such as tense (time relative to the Speaker and Addressee's Now), epistemic modality (reference to the Speaker's reasoning processes), and (in)definiteness marking (reference to informational accessibility by Addressee, as assessed by the Speaker in context).

Langacker has defined subjectivity as *implicit* (or relatively *unprofiled*) reference to the Speaker, Hearer, and generally to the Ground (S, H, and their physical and temporal discourse setting); Traugott's examples of subjectification also clearly involve added meaning of this implicit kind. A meaning such as tense would be subjective by Langacker's definition, since the primary

reference of the clause (perhaps some past situation) need not be about the Speaker's Here and Now, in order for past tense reference to be computed relative to the Here and Now. Similarly, establishment of a definite reference primarily profiles the entity so designated, not the Speaker and Addressee relative to whose mental access the definiteness is calculated: *the cat* "refers" to the relevant feline, not to S and H's assumption that they can both successfully identify it, though the assumption is necessary for the reference to work. Even pronouns such as *I*, *you*, *she* can be said ultimately to *refer* to the entities designated, and only secondarily to the implicit speech situation via which such referential interpretation occurs. Crucially, as Langacker says, in all such cases viewpoint is present: your *I* is my *you*, and vice versa, while tense reference is relative to some Ego's current Now. No utterance is referential without Grounding; without a Now, tense marking fails to locate a described event in time, and without a shared discourse record or physical context including a cat, *the cat* fails to locate a specific cat. Although at opposite ends of the cline, there is a contrast between very minimal profiling of Ground elements (every English sentence requires tense marking) and full overt profiling (mention of the Speaker or of *now*), there is also every possible degree of relative profiling of Ground and Content in between these extremes.

Mental Spaces Theory is particularly well adapted to formalize such concepts; the general notion of Viewpoint Space (Fauconnier 1985, 1997; Cutrer 1994; Sweetser and Fauconnier 1996) is already formalized as meaning the space via which information about the Focus (referentially profiled) Space is accessed. Past tense marking in *Chris drove to San Francisco* thus indicates that while the Focus Space is the Past Content Space where the profiled events and situations are being elaborated, the access to that Focus Space is via a separate Viewpoint Space, a Present Space wherein Chris's drive is no longer taking place. The temporal reference of both of these Content Spaces makes further implicit reference, however, to the Ground of the Speaker–Hearer interaction. Although S and H and their location and setting are not mentioned, the temporal reference of the Content Space network is most likely calculated based on S and H's Now (though of course it could also be an imagined Narrator's Now).

Sanders *et al.* (2009) have formalized the concept of a Basic Communicative Space Network (henceforth BCSN), which includes the construal of a Content Space network relative to a Ground network. Following past work by Sweetser (1990), Dancygier (1998), and Dancygier and Sweetser (2005), the Ground network necessarily involves:

- (1) a Ground Base Space, or Real Space (cf Liddell 1998, 2003): the real S and H in their spatiotemporal setting;
- (2) one or more Epistemic Spaces: understandings of S's and H's beliefs and reasoning processes, which are also background to the communication;

Basic Communicative Space Network

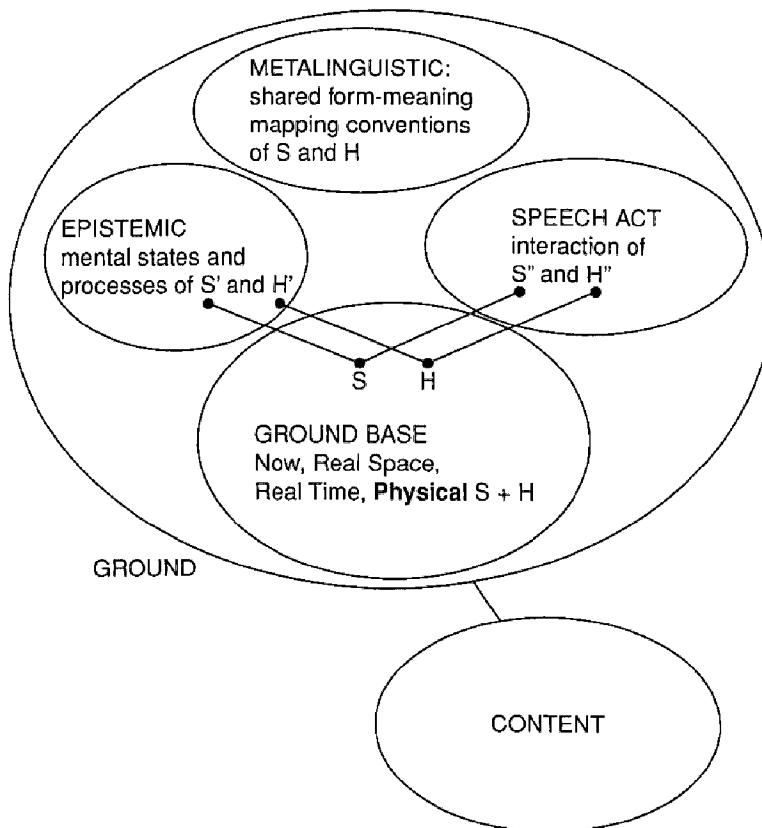


Figure 2.1

- (3) a Speech Act Space of performed speech interactions (e.g. this is a statement or a question);
- (4) a Metalinguistic Space of shared linguistic forms that can be referenced if chosen usages are brought into consciousness or disputed.

This extended BCSN complex is what allows the referential interpretation of the Content Spaces – that is, even though we may not be talking about the Ground, we inevitably need it (in Langacker's terms) to referentially “ground” our talk about other things. This space network, as argued in the work cited above, is therefore built up automatically and implicitly by any communicative use of language. One need not refer overtly to S, H, their mental processes, and so on, in order to make these entities available either for direct reference or as the Grounding for reference. No form can be used communicatively without some grounding of the Content with respect to a Ground; and as Langacker points out, there is a complex cline between primary profiling of Content (as in *Chris drove to San Francisco*), and primary profiling of Ground and interaction (as in *Hello*). Figure 2.1 shows a simplified version of Sanders *et al.*'s BCSN model.

It is general usage in the Mental Spaces framework to refer to embedded spaces as being *lower* in the network than the spaces in which they are embedded; hence the Content Space Network is lower in the BCSN than the Ground part of the network. A past Content Space is lower than the present one in which it is embedded; and in *Susie says that the world is flat*, the Content Space wherein *the world is flat* holds is embedded with respect to the higher Content Space wherein Susie is speaking.

Thus far we are not aware of a technical definition or model of subjectivity and subjectification in terms of mental space structure. We suggest that the model formulated below can bring together Traugott's and Langacker's distinct but evidently overlapping definitions, and will also help us to disentangle the concept of grammaticalization from that of subjectification. If subjectification is *increased dependence of the conventional meaning of a form on the Ground*, in Mental Space terms this can happen in at least two ways:

- (1) **Adding conventional semantic structure to the Grounding Spaces while the Content Space remains Focus, or moving Focus (basic reference) entirely from the Content to the Ground.** In the former case, the change involves adding more information about the still-implicit Ground to the semantics: an example of this, to be discussed below, would be the development from GOING TO INFINITIVE to the GONNA future, since the latter marks not only description of a future event, but also the Speaker's epistemic assessment of the present situation as giving rise to it. In the latter case, the shift makes the Ground the primary area of reference as opposed to some represented Content: an example of this would be the development of *God be with you* to *Good-bye*, where *Good-bye* no longer has any descriptive content meaning (hence no "truth conditions" in truth-conditional semantic models), but is meaningful only as a parting marker in the Speech Act Space. In both these cases, we might say that meaning migrates "upwards" in the space network.
- (2) **Increasing the distance between Grounding Spaces and Content Focus, while Focus remains lower in the network** – that is, increasing the calculation needed to locate the Focus Space with respect to the still-implicit Ground. We shall be examining cases of this as well – for instance, the development of the root or deontic meaning of the modal MUST to the epistemic meaning.

Our model of the Ground is situated within the BCSN model presented in Figure 2.1. Besides the Ground network (including the Base, the Speech Act Space, and the Epistemic Space(s)), and the Content Space network, a BCSN includes a metalinguistic space of putatively shared form-meaning mappings, and a metatextual space of shared discourse structure and history (cf Dancygier 1993, 1998; Dancygier and Sweetser 2005). These are parts of the Ground network, though not discussed in detail here. Placement within this extended BCSN

complex is necessary for the interpretation of any Content Space network – full meaning interpretation cannot happen without some relationship to the Ground.

Within the Ground, we take the Ground Base Space (or Real Space) to be less subjective, and the Speech Act and Epistemic Spaces to be more subjective. This is because the Real Space (defined by Liddell as a participant's conceptualization of their immediate shared spatiotemporal surroundings) is more intersubjectively verifiable in experience than the more abstract spaces of construals of mental states or speech interaction. Epistemic and Speech Act Spaces are inferentially dependent on the presence of an S and an H in the Real Space, to whom mental spaces and speech acts can be attributed. We never directly access someone else's mental representations; the connection between Focus Space content and S's epistemic states is one that is mediated via Real-Space interaction. More subjectivity is therefore involved in referencing these higher spaces, which are further from the Focus Space. Meaning that migrates upwards to or depends on these spaces is thus even more subjective than meaning that resides in or depends on the Ground Base Space.

As an example, let us take the development of definite articles from deictic modifiers with meanings such as THIS and THAT. Such developments are common cross-linguistically, and the English definite article is one instance of this natural grammaticalization and subjectification process (Traugott 1982, 1989; Hopper and Traugott 1993). It is important to note the level of subjectivity already present in deictic markers, although this is not focused on by Traugott. THIS and THAT clearly have reference to the Ground built into their meaning; in order to access the referent object (*this book* or *that tree*), they implicitly refer to the speaker's and/or the hearer's locations relative to the referent object, or to their manual or visual access to it. However, Traugott seems quite right to say that there is increasing subjectivity in the development of a deictic to a definite article, as the new meaning depends more on the mental processes of S and H, rather than on the physical Ground or Real Space of spatial relations.

In Figure 2.2, the mental space viewpoint structure of a THIS/THAT deictic is represented. The Content Space is understood to be primarily in Focus, including reference to certain objects. The upwards arrow indicates that access to those referents is achieved via reference to the unmentioned structure of the Real Space Ground of spatial relations – S's assessment of these relations as mutually accessible to S and H in their environment – even though the primary target of reference is not this presumed and unmentioned network of spatial relations, but the entity pointed out within it (a sentence such as *This book is red* does not profile the Ground).¹

In Figure 2.3, we diagram the relationships involved in interpretation of definiteness markers, which are frequently historically derived from deictic forms (as in the case of English *the*) and are one of Traugott's (1982, 1989) clear

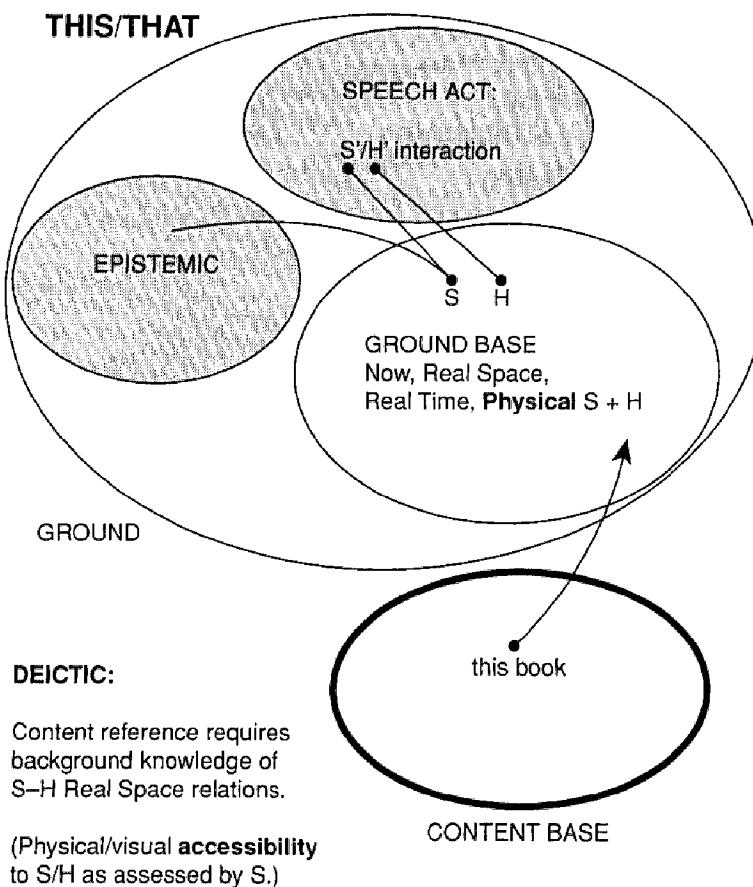


Figure 2.2

examples of subjectification of meaning. To frame or decode the reference of *the book*, unlike *that book*, it cannot be enough to know the Base Ground Space relationships involved in the Ground; the Speaker is communicating an assessment of the Hearer's cognitive access to relevant context, which may include speech act context (and discourse record) as well as Real Space context. As Figure 2.3 shows, we consider these spaces to be higher in the mental space network than the Base Ground Space, although necessarily linked to it (inasmuch as people in all Real Spaces have epistemic processes and interactional records). The Base Ground Space is more "objective" than these other spaces, because it is more intersubjectively verifiable and "shared." A definite article is thus more subjective than a deictic, because it involves implicit reference not just to the immediate presence of S and H in Real Space, but to the less accessible attendant spaces of their cognitive and discourse structures.

This kind of meaning structure is not in fact unique to grammatical markers. An example of a lexical item that simultaneously has meaning at such multiple levels is the verb *know*. *Ann knows that the world is round* in principle refers to (focuses on) Ann's cognitive state; *know* is a mental space builder

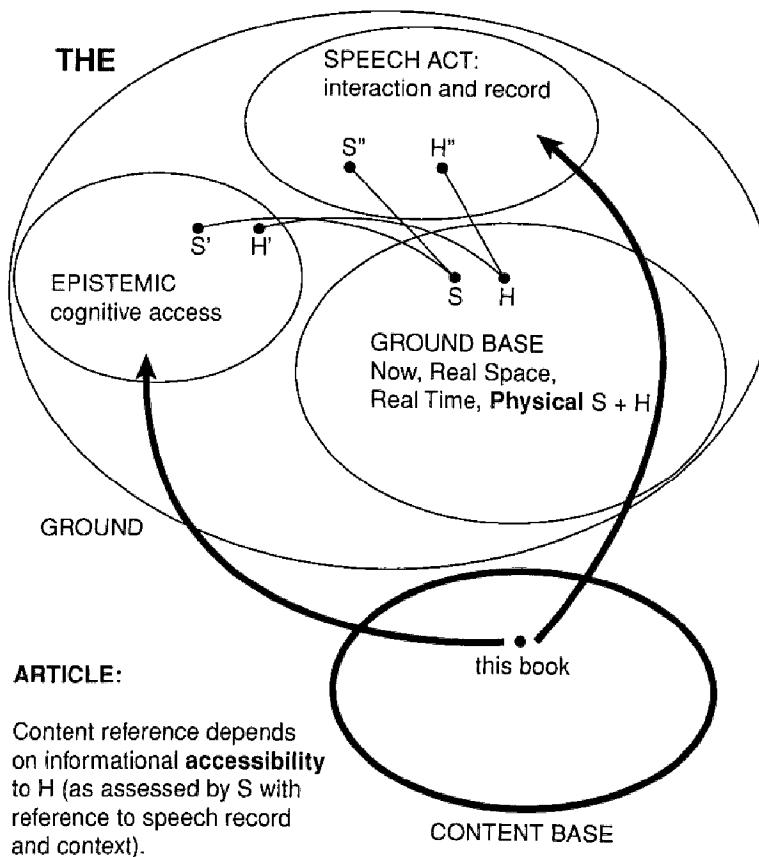


Figure 2.3

that opens up such a Focus Space of *stance* (see Dancygier, this volume). But it also necessarily makes reference to S's epistemic state, as shown in Figure 2.4. If S did not *also* believe that the world is round, then S would have to say something like *Ann thinks* or *Ann believes*, rather than *Ann knows*. *Know* and *the* both involve presuppositional structures; definite marking presupposes accessibility of the referent, which may in turn involve presupposing existence in some space. Conventional presuppositional items are ones that involve this kind of lexicalized structure in the non-focused Ground, as well as in the focused Content Space; reference to *stopping* in a Content Space (even as a query) marks S as believing in the presence of some ongoing situation that could stop, although it may not mark S as believing in the (explicitly referred to) stopping.

KNOW and STOP are good examples of structures involving subjective presuppositional meaning structure that we need not attribute to historical subjectification; many languages find these meaning packages useful, and express them, with no evidence that they derive historically from less subjective senses. Further, this presuppositional meaning is intersubjective in Verhagen's (2005)

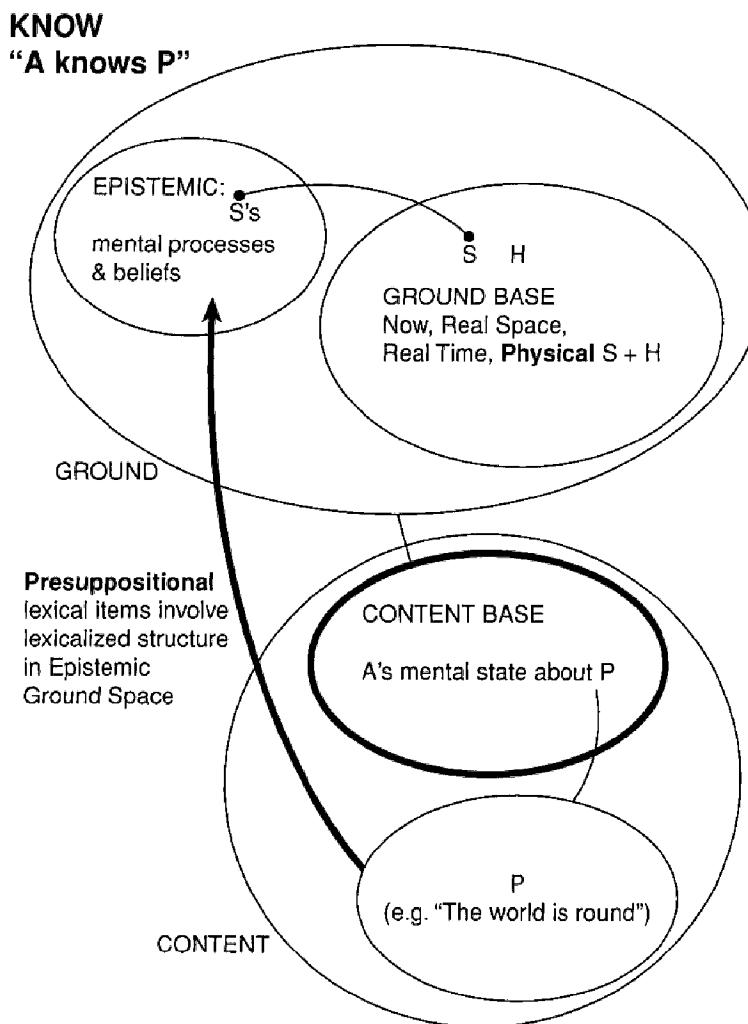


Figure 2.4

terms, rather than simply subjective. At least the Hearer's Epistemic Space is involved, not just the Speaker's: consider the difference between *I know P* and *I think P*; here the former pulls in an assumed shared “knowledge base” in a way that the latter does not. (These differences are further discussed by Dancygier (this volume).)

At this point it is important to mention the fact that it is oversimplifying to refer to “a” Viewpoint Space with respect to which a Focus Space is being accessed at any given time. As Fauconnier (1997) says, and Cutrer (1994) lays out in detail, any complex tense structure involves (analogously to the Reichenbach 1947 system) layers of such access: a pluperfect such as *had walked* demands placement of the Focus Space prior to some (reference or landmark) temporal space, which in turn is prior to the Real Time of the

Ground Space network. Further, referential structures are frequently mixed (cf Nikiforidou's and Vandelanotte's chapters in this volume). Imagine a speaker who is describing an earlier conversational encounter, and says, *Joe said I was now an experienced enough driver to borrow his car*. What Joe might have said was something like, *You're now an experienced enough driver to borrow my car*. The quoting speaker has retained Joe's original *now*, but changed the pronouns to her own viewpoint (*I* and *his* replace *you* and *my*), and imprinted temporal access from the speech Ground on the content by using the past tense *was*. In such cases there is more than one Viewpoint Space for the utterance as a whole. What is special about the forms discussed above is that the *same morpheme* is building space structure simultaneously at separate Focus and Viewpoint levels. Subjectification, then, involves conventionalizing new relationships between Viewpoint and Focus Spaces (KNOW = "belief that P in the Focus Space, while also belief that P in the Speaker's Epistemic Space").

Subjectification is thus closely related to our existing list of ways in which information contextually migrates upwards in a space network. Fauconnier (1985, 1997) has discussed Inheritance, whereby information present in a higher space such as the Base is allowed to move downwards and fill in relevant slots in lower Focus Spaces. For example, if I write about a fictional world where space flight is possible and humans are telepathic, readers will assume (unless told otherwise) that gravity works as it does in the real world (the Base), that human sensory perception works as it does in the real world, and so on. The inverse tendency is Presupposition Float, whereby information introduced into a lower space may be allowed to "float" upwards into the Base. Fauconnier gives the example of a speaker and addressee who pass one of the speaker's colleagues in a hallway. The colleague looks depressed, and the speaker says, *Maybe his children are worrying him*. Although the worrying is posited as taking place only in the *maybe* space, the existence of the children may well be assumed by the hearer to hold in the Base Space as well, given the speaker's likely knowledge about the immediate family structures of colleagues.

Subjectification is the historical result of such processes of information flow between spaces. When there is a frequently accessed pattern of inferential flow between spaces in a network, the "floated" inferences can eventually become a *conventional* part of a higher space. Further, as we shall see, the information transferred may be analogically or inferentially related to the Focus Space information, rather than being identical to it. The Ground, into which information moves upwards from the content, is not an unanalyzable whole; as we have said, it consists of a network of spaces, a sub-network of the BCSN diagrammed in Figure 2.1. The next section of the chapter lays out different ways in which historical meaning change involves shifting of meaning aspects between areas of the BCSN.

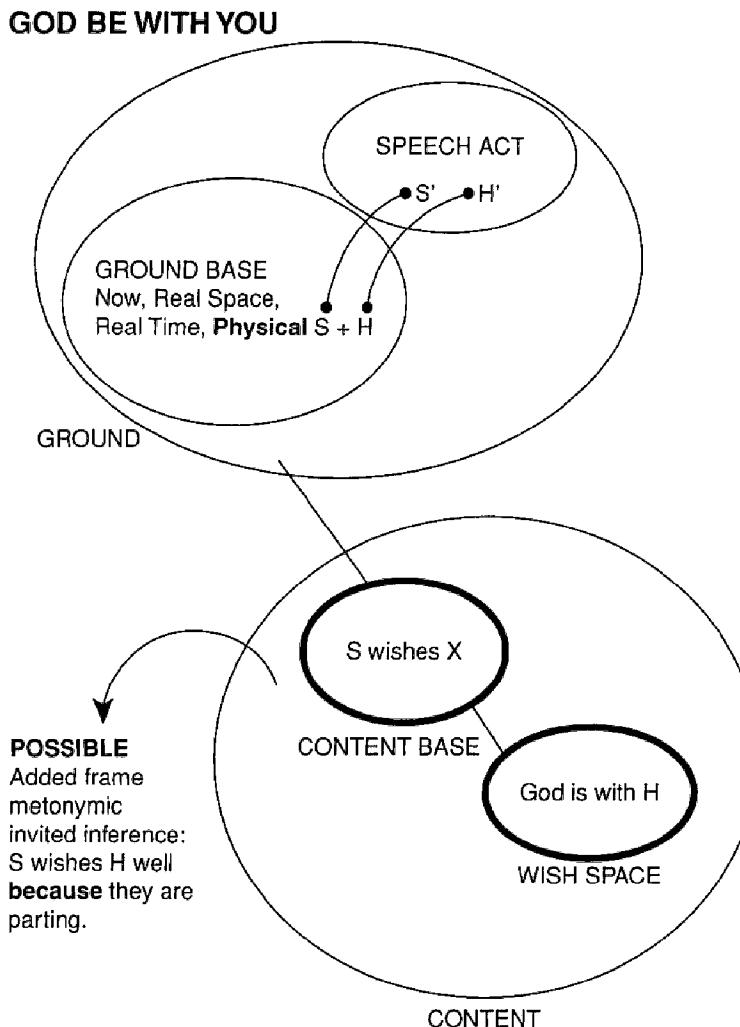
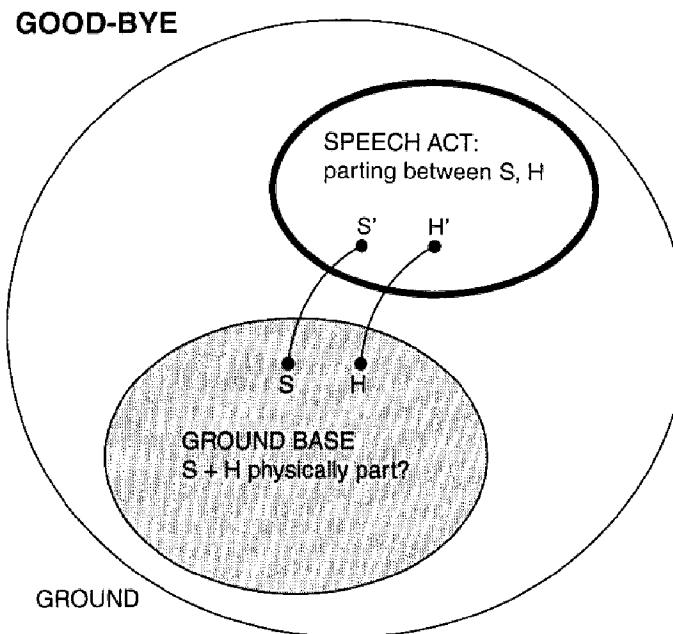


Figure 2.5

2.2 Subjectification and inferential transfer

In our typology of subjectification, one major sub-case is where the form completely loses Content Space referential meaning, and moves one-time inferences into the Ground sub-network to create a new conventional semantics. In Figures 2.5 and 2.6, we compare the semantics of *God be with you* and of its descendant *good-bye*, which are an example of this kind of subjectification.

Originally, *God be with you* was a positive wish from S to H; it set up a Focus Wish Space in the Content sub-network, wherein God is with H (presumably watching over and protecting H). In the Ground network, there was some situation involving S and H, and a discourse frame. One conventional cultural reason for wishing God to take care of H was an imminent parting between S and H, so the Ground Base frequently (though not always) contained a parting



No lower content spaces of wishing or wish content.

The focus is now in the Speech Act Space, regulating Ground Base interaction. The meaning is now entirely interactional – though without **overt** reference to S or H.

Figure 2.6

situation when this wish was uttered. Linguistic forms, we know, can either label extant situational frames or evoke the frames they normally label; therefore, if someone said *God be with you*, expressing a Wish Space with content of divine care for H, their hearers might well infer that in the Ground, parting was about to happen. (A similar story could be told for the more etymologically transparent *farewell*.) They might further infer that the speech act of uttering *God be with you* was a speech act intended to negotiate that parting. Once such inferences become conventionalized, *good-bye* or *farewell* is not only a good wish, but a speech act marking parting. And alongside the conventionalization of these Ground Space parting markers, their good-wish Content became backgrounded and gradually disappeared.

Good-bye does not today constitute a blessing or a good wish; it is just a social marker of parting. The meaning is now all interactional; yet the Ground is still not fully explicit, in that S and H are not overtly mentioned ("I the speaker hereby part from you the hearer"), but rather presupposed as present, the focus being on the negotiation of their speech act interaction. *Good-bye* and *Hello* are among words claimed by some semanticists to have "pragmatics but no semantics" – by which is apparently meant that they have no truth conditions,

MUST deontic > epistemic

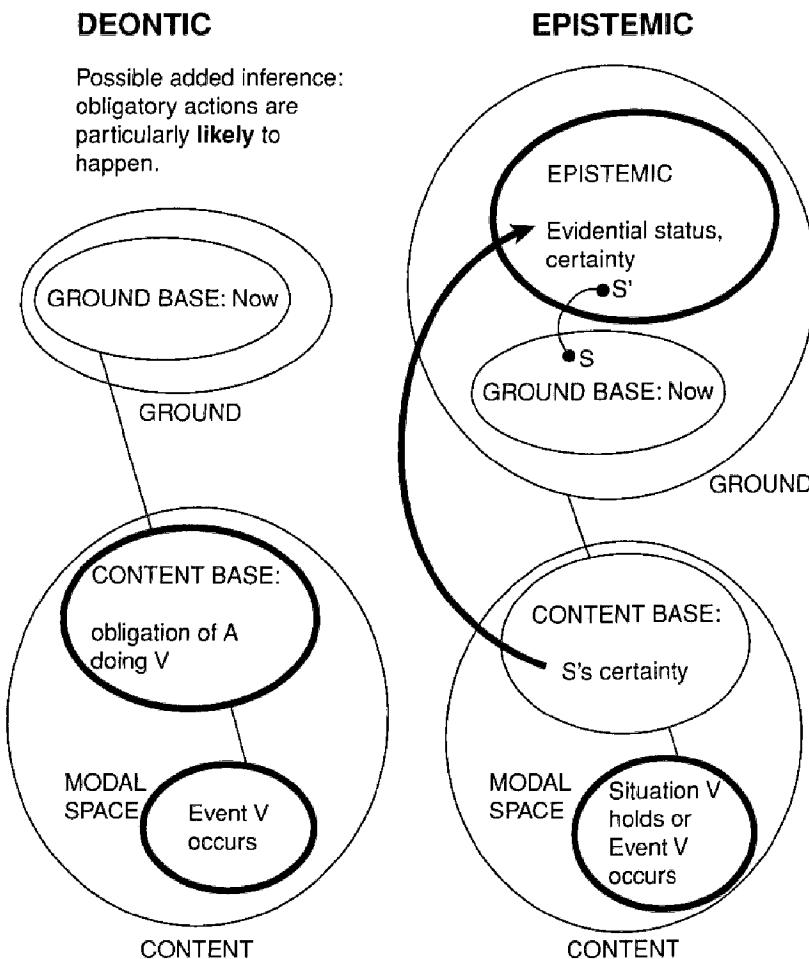


Figure 2.7

only pragmatic felicity conditions and pragmatic effects. We prefer to say that their conventional meaning resides in the Speech Act Space rather than in some lower content space – and yet, as just said, they assume the Ground and act on it, rather than overtly naming it.

Perhaps one of the most discussed examples of subjectification is the development of epistemic modal meanings from deontic or root modal meanings (Traugott 1982, 1989; Bybee and Pagliuca 1985; Sweetser 1990; Bybee *et al.* 1991, 1994; Hopper and Traugott 1993; Bybee and Fleischman 1995). In Figure 2.7, this shift is depicted for English *must*. Following Traugott's inferential transfer analysis, a statement of strong obligation on the part of an agent (in the Content or Focus Space) might naturally give rise to contextual inferences about the speaker's level of certainty that the strongly obliged action or event

would happen. That inference, however, is information not about the described event in the Focus Space, but about S's Epistemic Space. Once conventionalized into the word's semantics, it constitutes a new sense of the modal *must* – and a more subjective one. Not only have we added information upwards in the mental space tree, and lost information lower in the network – but *must* still refers implicitly, rather than overtly, to S and to the assumed Ground. In the sentence *He must be home by now*, meaning "I have good reason to think he is home by now," the broader clausal informational focus is on the content (someone being home), and the meaning of epistemic *must* focuses on the relationship *between* that Ground and the clausal Focus content (certainty), but not on the structure of the Ground itself.

Once again, in this case, as with deictics and definite markers, even the source domain of root modality is subjective and grammatical. Deontic modal meaning depends on some subject's modal space being accessed; it could be the Speaker's own space, if the Speaker is performatively imposing modality (*you must come home right away, because I say so*), or it could be some other subject's space, if the Speaker is merely describing the modality. But an epistemic modal requires us to go higher in the mental space structure, implicitly accessing information specifically about the cognitive processes of the Speaker, rather than of some lower subject.

Since tense and aspect markers have figured as important examples in the literature on subjectification and grammaticalization (cf Bybee and Pagliuca 1985; Heine *et al.* 1991; Hopper and Traugott 1993; Bybee *et al.* 1994; Bybee 1995), let us examine the grammaticalization of English *going to* to *gonna*, diagrammed in Figure 2.8. *She is going to Paris* presumably describes a content situation. There is no added information about any mental space other than the focal one where the described travel happens. From information about deictic motion, however, further inferences may be drawn about later locations of the moving entity: if someone is going to Paris at Time 1, then at some later Time N she may reach it. But, as Fleischman (1982) has pointed out, *She is gonna write that paper* conveys not only a description of a Future Mental Space, with content clearly distinct from the Base (present) Content Space, but also some information about the speaker's mental space, namely that the prediction about the future is based on some *assessment* (by default, the speaker's assessment) of the *current situation*. (This is distinct from the *will* future, where a prediction is made in the sense of speaker commitment to a statement about the future, but there is no such semantics of *present assessment*; cf Fleischman 1982; Dancygier 1998; Dancygier and Sweetser 2005.) The English *gonna* future thus involves two added spaces not required to represent the GO + INFINITIVE construction: a Future Focus Space embedded under the Present Content Space, and a speaker's Epistemic Assessment Space, present above the Real Space Ground. Again, the Focus is below, while the reference to

GO + INFINITIVE > GONNA

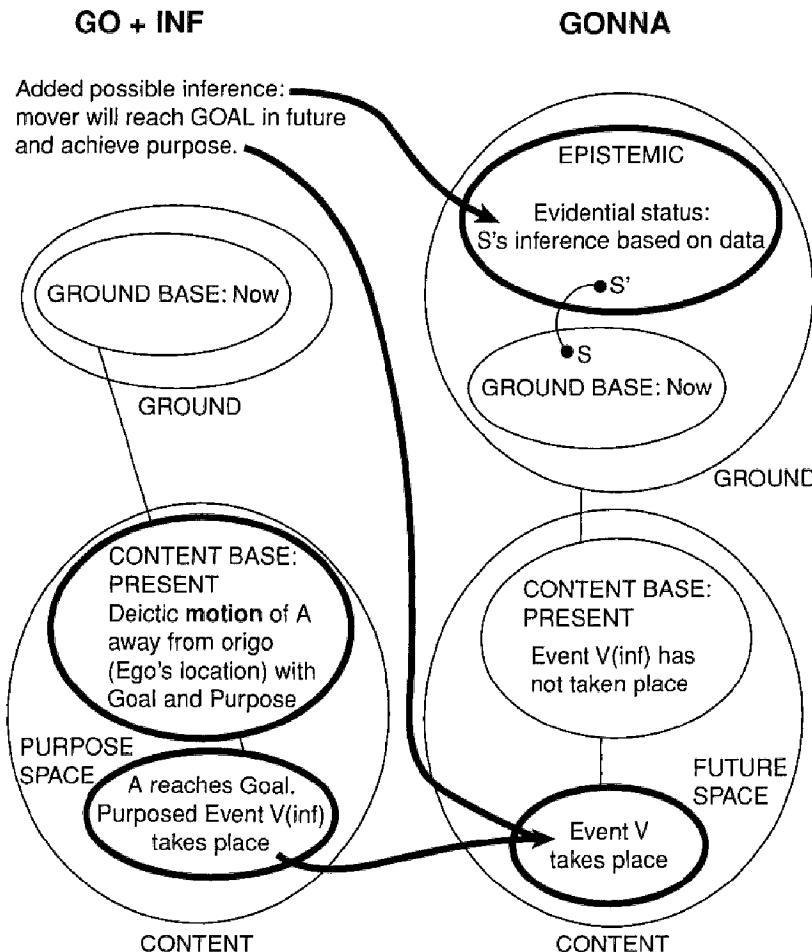


Figure 2.8

the higher Epistemic Space is implicit. Here, note specifically that the distance covered by the connections within the space network has increased as well. Where GOING + INF originally only involved one Content Base Present Space, with only the usual temporal grounding to the Ground Base, GONNA involves connections between an embedded Content Future Space and a higher Ground Space, the Epistemic Space of assessment.

In Figure 2.8, the left-to-right arrows mark correspondences between information in the GOING TO space network and information in the GONNA network.

The WILL future (Figure 2.9) has a similar story. The verb *will* once meant “want,” or “desire.” It thus set up a Desire Space wherein the situation described in the complement took place. In the highest-level Content Space, a Wanter wants P to happen; within the Want Space, Event P happens. By inference, Event

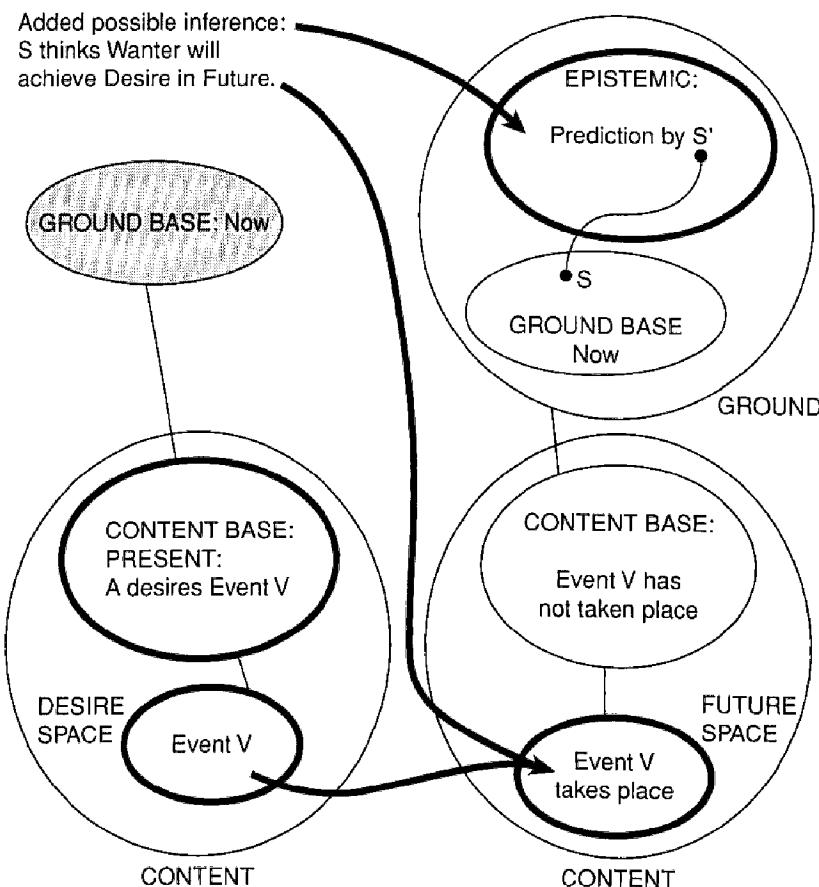
WANT (+INF) > WILL**WANT (+COMPLEMENT)****WILL V**

Figure 2.9

P does not hold in the Wanter's Base Space. However, Desire being a future-oriented intentional state, it is possible that S thinks that the Wanter will actually achieve his or her desired situation in the Future. By conventionalization of this inference, we get the switch from a Want Space to a Future Space in the meaning of WILL V constructions. Also, the Epistemic Space now contains semantics of *prediction* (commitment to a statement about a future situation), which is a conventionalization of the possibility that when S said *NP WILL INFINITIVE*, S was involved in prompting or sharing inferences about the future attained goal, not just the present desire state.

Note that in the GONNA and WILL futures, the literal meanings of WANT and GO are replaced by deictic tense meanings (building Future Spaces), which require negotiation of a larger temporal space network to relate the Focus Space

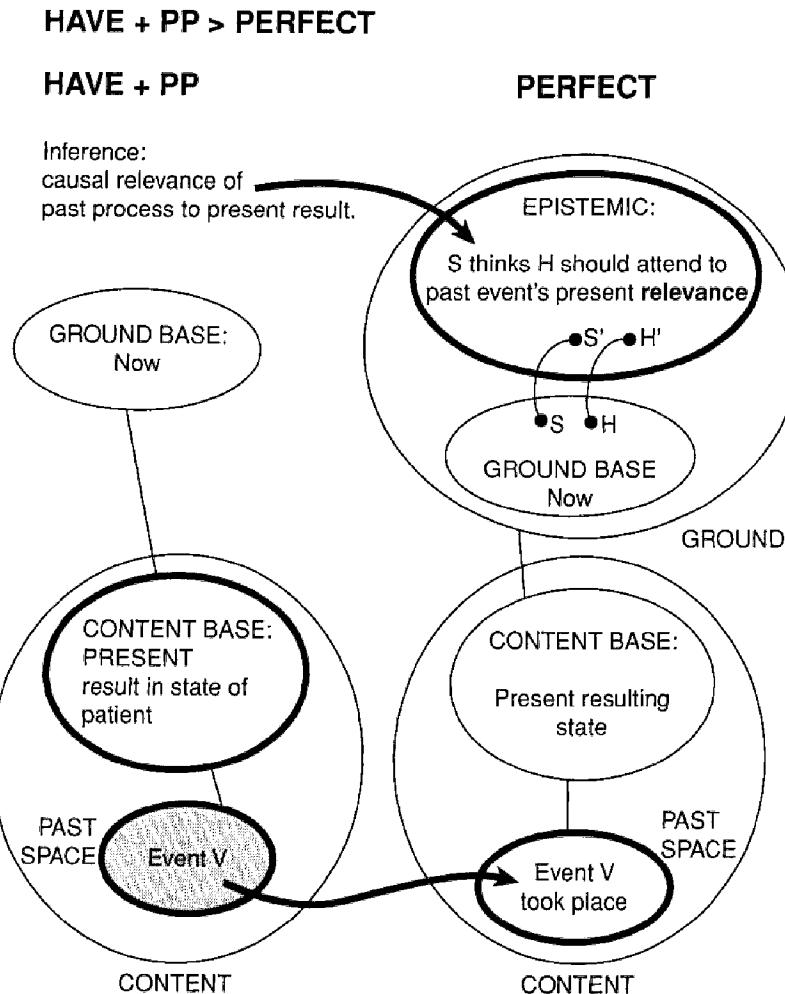


Figure 2.10

to the Ground's Here and Now than a present statement about deictic motion or wanting. Speaker activities of assessment and prediction are conventionally referred to by these grammatical forms as well, and this information is located higher in the space network than any of the content-level conventional meaning structures of lexical *go* or *want*.

Finally, let us consider two other semantic shifts, the HAVE + PP > Perfect shift (Figure 2.10), which happened in English and various other Indo-European languages, and the cross-linguistically common Perfect > Preterite shift (Figure 2.11). In the HAVE + PP > Perfect shift, we see something like the story we told for the English future morphemes (on perfects, see Fleischman 1983; Carey 1995).

We start out with HAVE + PP, involving a focal Content Space wherein the result of some process is tangibly present in the PRESENT state of affairs. This

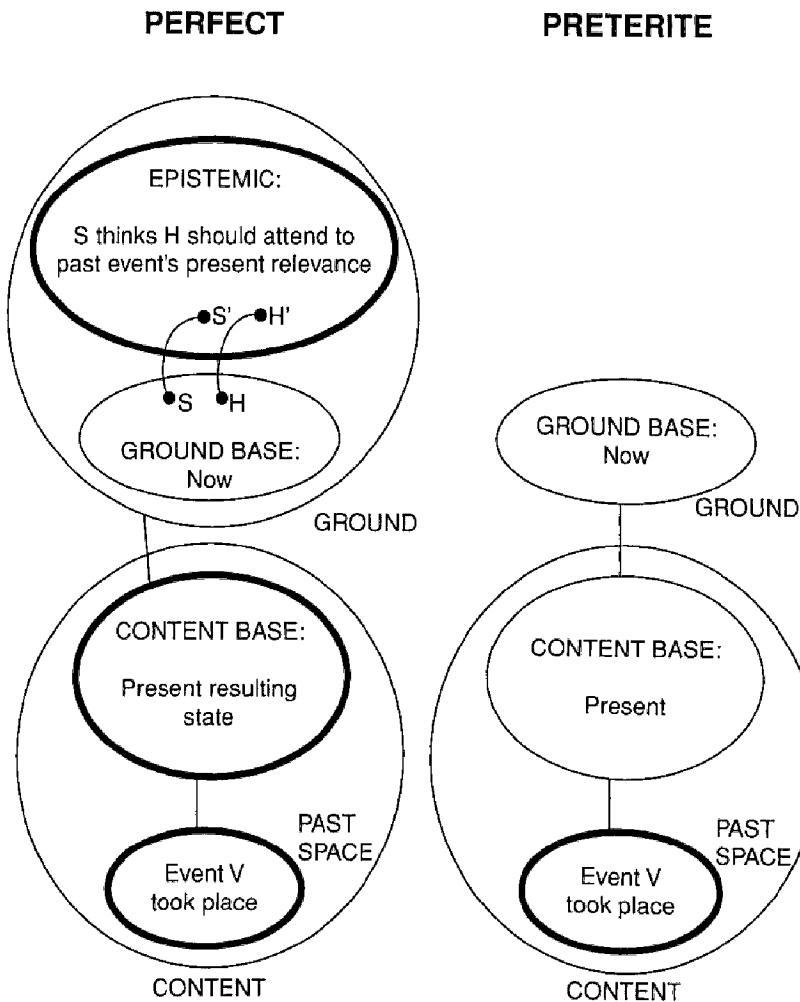


Figure 2.11

information allows inferences, in particular that the process in question really took place in the Past – but the past event itself is not profiled or in focus. And the Perfect indeed has a complex viewpoint structure, referring to two levels of Focal Spaces: *I have read the book* refers primarily to the present situation, where the results of reading are present and relevant, and secondarily to the past reading event, which gave rise to those results. Further, the use of the English Perfect conventionally presents the speaker as noting the relevance of the past event to the present situation. The simple past, *I read the book*, which refers only to the past, does not profile in the same way the relevance of past reading to my present state. This special PERFECT past–present relationship is a conventionalization of the earlier contextually accessible inference that the present state of things is related to past processes in which they were involved; as such, it requires conventional representation of aspects of the speaker's Epistemic Space, as well as of Present and Past Content Spaces. This can again

be seen as *implicit* reference to a higher aspect of the Ground Space network; there is no more *explicit* reference to the speaker's epistemic processes in a Perfect than in a HAVE PP sentence or a Past, but there is more need to refer to those processes. Again, with this process of change, we see both addition of new meaning in the upper more subjective areas of the Ground network, but also a longer distance between the parts of the space network that must be connected for the Content Past Space – for example – to be grounded.

And finally, PERFECT > PRETERITE (Figure 2.11) seems rather a puzzler, if we assume that meaning must always get more subjective. This semantic shift has been rather frequently documented; a much-discussed example is the French *passé composé*, which apparently was more of a present-relevance Perfect and is now simply an aspectually perfective Preterite (in contrast with an imperfective past tense, the Imperfect or *imparfait*). On the one hand, the Preterite has simple focus, since only the Past Space is profiled; the focus is no longer spread over Present and Past Spaces, as in the Perfect. This would mean that slightly more distance has to be traveled from the embedded Focus Space to the Ground – one of our suggested definitions of subjectification. And on the other hand, the information about S's Epistemic Space is gone: a preterite refers to a past event, not to its present relevance (although contextually that relevance may be inferable). Many analysts (cf Fleischman 1983) describe the French Perfect > Preterite shift as a loss of emphasis or immediacy of relevance; basically, the conventional special meaning of relevance is worn away until all we have is the basic level of relevance demanded of any speech act. But is this really increasing subjectivity, or for that matter increasing grammaticalization? How can we tell? Surely both meanings are highly grammaticized, perhaps equally so. And both involve the basic subjectivity of deictic tense reference, which relates the content to the Ground Here and Now. But it does seem that the reference to Ground meaning has diminished rather than increased, no matter how we compute it.

2.3 Conclusions

A mental space model involving a BCSN is, we find, helpful in modeling viewpoint, subjectivity, and subjectification. Added internal elaboration of the Ground into a space network is also helpful, since it allows us to notice movement of information to higher or lower locations within the Ground network. Traugott's and Langacker's models are readily brought together in this framework. Traugott's inferential transfers are now specifically transfers of information between spaces. Langacker's Ground is given a more elaborate mental space instantiation as a BCSN. And the crucial role of implicitness in subjectivity is recognized by stipulating that when meaning is subjective, the Focus is not in the Ground, although Viewpoint is at least partly there.

Grammaticalization is precisely and simply development of full conventionality in Bybee's or Langacker's terms, while subjectification is change along a particular semantic dimension. Pulling these apart, we see that the model in principle allows for a construction to remain equally grammatical, while possibly losing some degree of subjectivity (PERFECT > PRETERITE?).

Much of morphological and syntactic "grammar" is involved in Grounding, as Langacker says. A verb with no ultimate tense grounding, or a noun with no ultimate referential grounding (however inherited), would not be a usable construction. And Grounding is deeply subjective. But not all grammatical marking is equally subjective. Do we want to say that plural or genitive marking is more subjective (in the relevant sense) than content semantics of nouns or verbs? Everything in language is subjective to some degree, and viewpointed, in that language always represents a particular cognitive construal rather than objective portrayal. But some content-level meaning structures (plurality could be one of these) just need constant reference in discourse, and get grammaticalized, without necessarily having the extra, implicit and construction-specific, conventional meaning at a higher level of the space network. Similarly, some content-level meaning structures (*know*, *stop*) refer to presuppositional structures, and are thus more subjective than others, despite not being grammaticalized.

We have now developed a concrete account of subjectification in the mental spaces framework, using, in particular, the BCSN expansion of the framework, which allows us to understand the structure of the implicit Ground in more detail. We have also shown how this account brings forward new questions about the degrees of subjectivity in meaning and the relationship between subjective meaning and grammar.

What are the broader lessons for our developing understanding of viewpoint? One is that implicit and explicit viewpoints are not just separate kinds of things; this is why implicit viewpoint structure can develop into explicit viewpoint structure historically over time, in subjectification processes. The distinction between profiling the Ground and profiling the expressed content is fluid, or a matter of degree – as we would predict from Langacker's (1987, 1990, 1991) framing of the issues. Another is that we need to attend to the internal structure of the complex Ground spatial network, since newly explicit structure may end up being represented specifically in one place: *know* and other presuppositional forms elaborate the representation specifically of the Epistemic Space of the Speaker, for example, while *good-bye* adjusts the Speaker–Addressee social relationship.

A question that we have not addressed here, though it is addressed by Sanders *et al.* (2009) and by a number of other chapters in this volume, is the relationship of Ground participants' viewpoint to character viewpoint; free indirect style, for example, as a construction, expresses a particular blend of narrator and

character viewpoints. Historical work could be done on the development of these larger discourse-level constructions as well, and Construction Grammar and Mental Spaces Theory provide potentially useful tools for such work.

And last but not least, this work reinforces a point that Fauconnier and Turner (2002) make very clearly, but which seems insufficiently profiled in some work on mental spaces; “meaning” does not reside in one space, but in a spatial network. Tobin and Israel (this volume) make this clear with respect to irony – particular content may be in one space, but ironic interpretation of that content resides in the network to which it is linked. Similarly, with respect to the interpretation of tense and viewpoint in narrative (for example, chapters in this volume by Nikiforidou, Vandelanotte, and Janzen), how to interpret a past tense form, or a signer’s bodily position, depends on what spatial network is available, not just on the form itself. In our examples, we have seen that subjectification generally involves added complexity in the distribution of information across the BCSN. The Mental Spaces model allows us to talk about this explicitly, and to describe the network as a whole rather than just the Content Space. We end, therefore, with the hope that cognitive linguistic models will go on to facilitate understanding of subjectification and subjectivity in new domains of language.

References

- Bybee, Joan L. 1995. The development of past tense modals in English. In Joan L. Bybee and Suzanne Fleischman (eds.). *Modality in Grammar and Discourse*. Amsterdam: John Benjamins 503–17.
- Bybee, Joan and Suzanne Fleischman (eds.). 1995. *Modality in Grammar and Discourse*. Amsterdam: John Benjamins.
- Bybee, Joan and William Pagliuca. 1985. Crosslinguistics comparison and the development of grammatical meaning. In Jacek Fisiak (ed.). *Historical Semantics and Historical Word Formation*. Berlin: Mouton, 59–83.
- Bybee, Joan, Revere D. Perkins, and William Pagliuca. 1991. Back to the future. In Elizabeth Closs Traugott and Bernd Heine (eds.). *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 17–58.
1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. University of Chicago Press.
- Carey, Kathleen. 1995. Subjectification and the development of the English Perfect. In Dieter Stein and Susan Wright (eds.). *Subjectivity and Subjectivisation*. Cambridge University Press, 83–102.
- Cutrer, Michele. 1994. Time and tense in narratives and everyday language. Ph.D. dissertation, University of California, San Diego.
- Dancygier, Barbara. 1993. Interpreting conditionals: time, knowledge and causation. *Journal of Pragmatics* 19, 403–34.
1998. *Conditionals and Prediction*. Cambridge University Press.

- Dancygier, Barbara and Eve Sweetser. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Fauconnier, Gilles. 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge MA: MIT Press. (Reprinted 1994, Cambridge University Press.)
1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fauconnier, Gilles and Eve Sweetser (eds.). 1996. *Spaces, Worlds and Grammar*. University of Chicago Press.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Fleischman, Suzanne. 1982. *The Future in Thought and Language*. Cambridge University Press.
1983. From pragmatics to grammar: diachronic reflections on complex pasts and futures in Romance. *Lingua* 40, 183–214.
- Hanks, William F. 1990. *Referential Practice: Language and Lived Space Among the Maya*. University of Chicago Press.
- Heine, Bernd, Ulrike Claudi, and Frederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. University of Chicago Press.
- Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge University Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. 2 vols. Stanford University Press, vol. I.
1990. Subjectification. *Cognitive Linguistics* 1, 5–38.
1991. *Foundations of Cognitive Grammar*. 2 vols. Stanford University Press, vol. II.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge University Press.
- Liddell, Scott. 1998. Grounded blends, gestures and conceptual shifts. *Cognitive Linguistics* 9:3, 283–314.
2003. *Grammar, Gesture and Meaning in American Sign Language*. Cambridge University Press.
- Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. London: Collier-Macmillan. (Reprinted New York: Free Press, 1966.)
- Rubba, Johanna. 1996. Alternate grounds in the interpretation of deictic expressions. In Fauconnier and Sweetser, 227–61.
- Sanders, José, Ted Sanders, and Eve Sweetser. 2009. Causality, cognition and communication: a mental space analysis of subjectivity in causal connectives. In Ted Sanders and Eve Sweetser (eds.). *Causal Categories in Discourse and Cognition*. Berlin: Mouton de Gruyter, 19–59.
- Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press.
- Sweetser, Eve and Gilles Fauconnier. 1996. Cognitive Links and domains: basic aspects of Mental Space Theory. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds and Grammar*. University of Chicago Press.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann and Yakov Malkiel (eds.). *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, 245–71.

1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 57, 33–65.
1995. Subjectification in grammaticalization. In Dieter Stein and Susan Wright (eds.), *Subjectivity and Subjectivisation*. Cambridge University Press, 31–55.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge University Press.
- Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax and Cognition*. Oxford University Press.

3 Negation, stance verbs, and intersubjectivity

Barbara Dancygier

This chapter reconsiders some widely held views on negation, in the context of the mental spaces framework. Specifically, I develop the idea of inherent “alternativity” of negation to show its function as a viewpoint device, and, more specifically, as a stance-marking form. Also, I consider some constructions involving negation to explain its intersubjective role and the way it is used to negotiate various viewpoints available in the specific context. I also show how this approach can clarify the confusion surrounding the interpretation of constructional forms such as Neg-Raising or metalinguistic negation.

3.1 Negation, mental spaces, and intersubjectivity

In the mental spaces framework, negation has come to be treated as the primary example of alternativity (Fauconnier 1994 [1985], 1997; Sweetser 2006). The negative particle *not* is thus said to set up two alternative spaces, rather than just one: the negative space described in the sentence and its positive alternative. Indeed, the main reason for using linguistic negation seems to be to counter a contextually available claim or belief, or to present that claim or belief as potentially assertable. In now-classic examples such as (1):

- (1) There is no milk in the fridge. (Fauconnier 1994 [1985])

the mental space described (no milk) typically makes sense in the context where the presence of the milk is expected or cognitively accessible (hence the potential interpretation of [1] as a request or complaint). Whichever one of the possible interpretations is chosen, it will crucially rely on the mental space alternative to (1), such that the milk is (will, was, should be) in the fridge. This approach is essentially different from the truth-functional one in that it does not limit its interest to the speaker’s commitment to truth or falsehood of one of two incompatible instantiations of a proposition, focusing instead on the way in which the speaker exploits negative constructions to evoke and use their alternatives. In a sense, this approach also uses the contextually available information in specific ways, without necessarily relying on implicature or other

relevance-related phenomena, or treating all oppositions of meaning in a uniform way (cf Jeffries 2010).

In this chapter I argue that the alternativity of negation also makes it possible to see negation as a marker of epistemic stance, and that there are various contexts in which the type of stance is built into a specific mental space network. One could argue that a positive equivalent of (1), such as *There is milk in the fridge*, could also be a marker of stance when the speaker is questioning the hearer's assumption expressed earlier or marking surprise. But when negation is used there is no doubt – partly with respect to the alternative set-up evoked – that there are at least two options on the table, and the speaker as well as other participants may be variously epistemically aligned with the alternatives. Furthermore, negation can prompt complex intersubjective networks involving stances available in the Ground. But even the simplest case such as (1) involves a level of discourse where the speaker and the hearer both have access to the positive alternative.

The alternativity approach to negation also offers an explanation of the use of grammatical devices relying on the availability of the alternative space, as in Example (2),

- (2) I didn't buy a car. There was no room for **it** in the garage.

where the pronoun *it* refers to an instance of the category *car* inhabiting the negated space where the purchase has actually been made. As this chapter argues, negative forms participate in various grammatical constructions relying on much more than truth and falsehood. There is a variety of such constructions, but I will focus specifically on those related to stance expressions and intersubjectivity (for other examples of negative constructions, see Oakley 2005).

Possibly the best instantiation of the space evocation power of negation is the consistent use of negation as a narrative device in Joseph Heller's *Catch-22* (1955). A full analysis of negation in the novel is beyond the goals of this chapter, but just one quote will do for now (in reading, it is important to keep in mind that the club has not been mentioned in the text until now).

- (3) [Nately] was crazy, too, and had gone every free day to work on the officers' club that Yossarian had *not* helped build. Actually there were many officers' clubs that Yossarian had *not* helped build, but he was proudest of the one in Pianosa. It was a sturdy and complex monument to his powers of determination. (p. 27)

The space this fragment sets up is the shared effort of all officers in Pianosa in building a club for themselves. One might expect that Yossarian, as an officer, would participate in the work and be proud of the result, but he is proud precisely because all was done without his help. Yossarian's consistent refusal

to be ‘positive’ about any aspect of the war period is thus underscored by his alignment with a space that negates every other officer’s experience.

Alternativity is sometimes also expressed lexically, as in Wisława Szymborska’s poem *Atlantis* (1997), which starts with *They were or they weren’t, On an island or not, An ocean or not an ocean swallowed them up or it didn’t*. The use of *or* makes the consideration of the two alternative spaces equally explicit, and the construction further underlines not just the contrast, but also the epistemic stance, whereby either one of the alternatives may hold. The explicit reliance on alternativity of negation in this case, as in possible colloquial expressions of this kind, such as *He’ll marry her or maybe he won’t*, is a form of epistemic statement where both options are available, and thus truth and falsehood are backgrounded even further. In fact, the poem, which is a list of parallel alternative statements, can be naturally read as a comment on our lack of epistemic grounds to speak of mythical places such as Atlantis (cf Dancygier 2010); similarly, the ‘marriage’ example may be understood as a complaint about the man’s unwillingness to make a commitment. Examples like this suggest that negation itself can be seen as an epistemic stance device, one that may favor one of the alternatives, or reject one, or not mark any of them with positive stance.

The alternativity of negation naturally explains most uses of negation in available literature (e.g. Horn 1985, 1989). In fact, it also explains the uses in which negation moderates the commitment of the speaker, rather than the facts in question. In the example considered by Sweetser (2006), repeated here as (4), the speaker/narrator is not prepared to judge the actual appreciation of the listeners, but notes both the effort and its positive result.

- (4) I will not say that they all appreciated the music which they heard, but they were intent on appearing to do so... and they were not unsuccessful. (Trollope, *The Warden*)

The irony and caution of the narrator’s words result from the use of negation in ways that play on the alternative evaluations, and not on facts. Crucially, in such contexts the use of negation does not rely on truth and falsehood, but has a moderating and distancing effect instead. It is interesting to note that alternativity of negation can be exploited in this way in view of Davies’ proposal (1979) that negation suggests two polarized values, such as “0” and “1,” with a cline of modal stances in between. It appears that explicitly juxtaposing the alternatives has the effect of opening up the entire range of modal possibilities and degrees of epistemic commitment.

Alternativity, though of a somewhat different kind, also lies at the root of the use of negation termed “metalinguistic negation” (see Horn 1985, 1989):

- (5) Grandma isn’t feeling lousy, Johnny, she is just a tad indisposed.

Horn argues that cases like (5) are not aligned with ordinary negation, because no proposition is in fact negated. The objection in the first clause (*Grandma isn't feeling lousy*) is not directed at the actual health of the grandmother, but at the choice of words, which does not show sufficient respect for Grandma. Horn's term, *metalinguistic negation*, has been useful in turning the semanticists' attention to overt comments on the linguistic form of utterances, but it avoids the question of why negation would be used differently in the propositional versus non-propositional cases. In a sense, the distinction is not the result of the actual differences in the kind of negation, but an attempt to maintain the assumption that negation is primarily propositional. It also fails to acknowledge the fact that metalinguistic negation occurs in a rather specific constructional format, where two alternative ways of presenting the "same" content are contrasted (*feeling lousy/a tad indisposed*). In other words, a sentence such as *Grandma isn't feeling lousy* would not carry a metalinguistic force on its own, since the construction typically profiles both the offending expression and its accepted alternative. In her discussion of negation and conditionals, Dancygier (1992, 1998) observes that the so-called metalinguistic uses of negation are crucially dependent on constructions where the questionable (and thus negated) expression is immediately followed by a proposed repair, and further, that the same constructional format is available to propositional negation, as in (6).

- (6) John wasn't born in Boston, but in Philadelphia.

In other words, this specific use of negation is not strictly metalinguistic, but, more broadly, metatextual, and requires that the offending expression has occurred in preceding discourse and is both quoted and negated, so that a repair can be offered.

What is of particular importance in this case is that the construction relies on its context in a rather specific way. The negated part of the utterance (*Grandma is feeling lousy, John was born in Boston*) is echoic – that is, repeated after being made available in preceding discourse. This suggests that there are two different ways in which the alternativity of negation uses the spaces available in the context. In a case like *There is no milk in the fridge*, the hearer's attention is drawn to desired or past presence of milk in the fridge, so that the alternative is prompted by the negation, while in a case like *John wasn't born in Boston*, the alternative is available in the immediate context and negated. As I will try to show below, the difference is important in terms of constructional features, and also in the epistemic consequences. Furthermore, a constructional approach to echoic mention and related phenomena, such as distancing or irony, can also clarify the many different ways in which echoic discourse is used.

In more recent work, examples like (4) and (6) were discussed by Verhagen (2005, 2006) as constructions relying on intersubjective mental space set-ups.

This approach clearly highlights the constructional nature of such usage and its primary goal of referring to utterances available in the immediate context. Verhagen is discussing such cases in terms of “intersubjectivity” to highlight the fact that the utterance may have different discourse sources, as in the cases discussed above.

The concepts of subjectivity and intersubjectivity are rather contentious ones, and they are both directly pertinent to the topic of this volume – viewpoint. There seem to be several dimensions that delineate different positions researchers take, but the one most important for this chapter is the issue of correlation of subjective meaning with specific linguistic forms. Langacker’s view (1985, 1987, 1990, 2006) is that subjectivity is a matter of offstage, implied viewing subject, situated in the context of the Ground – the broadly conceived situation of utterance, including the participants, their conversation record, location, and so on. This is also the sense that Coulson and Oakley (2005) use in their theoretical attempt to include discourse grounding in the analysis of mental spaces and blending. While Langacker’s view acknowledges the presence of both the speaker and the hearer, he is not interested in the way in which the negotiation of meaning across these subjectivities is coded in the utterance, and this is what Verhagen (2005, 2006) attempts to add to the picture. Thus Verhagen’s concept of intersubjectivity seeks to explain how specific constructions are used when one of the participants is arguing for a construal or reporting a construal proposed by another participant.

There is also the approach, represented primarily by Traugott (1989, 1995, 2003) and Traugott and Dasher (2005), whereby both concepts (subjectivity and intersubjectivity) are relevant in the explanation of the meanings of specific kinds of expressions. Subjectivity may be useful in accounting for the meaning of modals or deictic expressions, as shown in Ferrari and Sweetser (this volume), while intersubjectivity requires overt attention to the speaker and the hearer as discourse participants, not as conceptualizers (hence it is applicable in accounting for politeness, hedges, etc.). Consequently, Traugott’s view is different from both Langacker’s focus on offstage conceptualizations and Verhagen’s interest in argumentation. Furthermore, Traugott uses the same approach in accounting for semantic change, so that subjectified or intersubjectified meanings of specific forms may emerge through their historical development. Indeed, some combination of Traugott’s and Verhagen’s approaches is needed to explain the use of negation and its description as a stance device.

Verhagen focuses on the constructions that engage the speaker’s and the hearer’s viewpoint and are used to *argue* for a specific construal and negotiate it across different subjectivities. Among other things, Verhagen’s approach explains the crucial difference between sentential negation and morphological negation in (7) and (8).

- (7) Mary is not happy. On the contrary, she is feeling really depressed.
- (8) #Mary is a bit sad/unhappy. On the contrary, she is feeling really depressed. (Verhagen 2005)

The expression *on the contrary* is possible in the intersubjective case, where the speaker uses negation to oppose a positive opinion about Mary's mood that is available in the context; for comparison, the expressions in (8) can only be read as the speaker's own evaluations, and thus cannot be subject to further negotiation. Let us also note that intonation can further highlight the questioned expression. Clearly, there are many constructional features to rely on in distinguishing various types of negative constructions.

The alternativity of negation is the source of its functioning as an argumentative or stance device. There are, in fact, other similar examples of correlations between alternativity and stance, as in Dancygier and Sweetser's (1997, 2006) discussion of alternativity as the mental space configuration underlying the so-called predictive conditionals, such as *If you chill the cranberries after cooking them, the sauce will gel*. The alternativity of the set-up lies in the simultaneous evocation of two scenarios: chilling, followed by gelling, which is the scenario described in the sentence, or keeping warm without gelling (the alternative suggested by a predictive construction). As Dancygier and Sweetser argue, alternativity also opens the construction to the marking of stance, as in *If you had chilled the cranberries after cooking them, the sauce would have gelled*, where the past tense forms prompt an interpretation whereby the speaker believes that the cranberries have been kept warm. The stance thus marked could further extend over the following discourse, as in *and we could have had it with the turkey*. Dancygier and Sweetser also suggest that the possibility of marking stance in a conditional depends directly on the alternative set-up – if there are alternative scenarios, they are open to different evaluations in terms of predictability, probability, and so on. Since negation has been described as the ultimate prompt for alternative spaces, it is worth considering how the resulting set-up is then used for stance marking. This is one of the questions this chapter attempts to explore.

3.2 Stance verbs

The concept of stance is clearly related to subjectivity, in that stance expressions are talked about as representations of the speaker's specific evaluation of assertability, built into specific expressions. It is possible to argue that 'stance' represents the kind of subjectivity where the speaker's viewpoint is directly coded in the expression, whether as a result of subjectification or another semantic change process. In this sense, expressions such as *I think*, modals, and past tense of modals (*may* versus *might*), are naturally treated as stance

expressions, but pronominal deictics may not be, though they are viewpoint expressions (cf Ferrari and Sweetser [this volume]; consider the difference between *this* and *that* or *I* and *you*). One recurring reason for analysts to talk about stance is to distinguish different degrees of stance commitment (though there are surely stance expressions that are not naturally graded).

The most-studied kind of stance is probably epistemic stance. In much of the work within cognitive linguistics, starting with Fillmore's discussion of conditionals as displaying neutral, negative, and positive epistemic stance (Fillmore 1990a, 1990b), through Sweetser's (1996) discussion of stance embedding, to Dancygier and Sweetser's (2005) analysis of conditionals, stance has been associated primarily with the choice of verb forms. In particular, past tense forms have often been described as marking "attitude," "distance," or "negative stance" when they signal doubt, politeness, counterfactuality, and related stance concepts. There is also some work on lexical expression of stance, as in Biber and Finegan's (1988) discussion of adverbial stance types or Kärkkäinen's (2003) discussion of the verb *think*. In fact, the verb *think* seems to be an ideal candidate for a stance marker, since it explicitly invokes mental attitudes not necessarily quantifiable in terms of truth and falsehood. At the same time, however, *think* is one of the most common verbs used in the constructions of represented speech and thought, and the contrasts between different constructions often downplay issues related to stance. *Think* has thus been discussed primarily with respect to the contrast between its *representational* uses, where someone's beliefs are being represented and can then be judged as true or false (*He thinks the earth is flat*), and the interpersonal or *subjectified* uses, which are grammaticalized expressions of stance (*I think, I guess . . .*) (cf Traugott 1989, 1995; Thompson and Mulac 1991; Vandelanotte 2009). Recent work by Wierzbicka (2006) and Brinton (2008) reviews a number of similar expressions (*I reckon, I believe, I find*, and many others), arguing for the fine stance differences they represent in English (in fact, Wierzbicka also observes that other languages express fewer varieties of stance). Weighing the degree of subjectification in such stance expressions against the degree to which they remain representational is often difficult, and relies at least to some degree on the choice of the pronoun. Expressions using the first person pronoun are more readily seen as subjectified, while the ones with third person pronouns are more naturally interpreted as representational.

However, as Verhagen proposes, sentences with stance that represent beliefs and epistemic attitudes of participants referred to via third person expressions are not free of subjective meanings. Specifically, Verhagen argues that constructions of represented speech and thought (STR) such as *X said/knew/hought that Y . . .* are intersubjective, and serve as instructions to entertain Y in the way X entertains it. However, the "reporting" function alone does not necessarily evoke subjectivity, and there are other ways to represent contextually available

stretches of discourse that are not like STR in attributing beliefs to participants. For example, recent work (Dancygier and Vandelanotte 2009) identifies usage referred to as “distanced discourse,” where specific discourse expressions are incorporated into the speaker’s discourse (as in specific forms of STR such as DIST, see Vandelanotte 2004a, 2004b, 2009, this volume) and co-opted into the speaker’s deictic viewpoint, while maintaining their discourse alignment with another speaker. It should be noted, then, that subjectivity or intersubjectivity may emerge independently of constructional patterns, while similar constructions may not be equally subjective or intersubjective.

The degree of incorporation into the speaker’s discourse is an important criterion in considering the meaning of “intersubjectivity,” especially in application to negation, as well as verbs like *think*. The “intersubjective” use of negation exemplified in (7) relies on the contextual availability of an evaluation of Mary’s current disposition as “happy”; similarly, reporting someone’s beliefs, as in *Jeremy thinks/says/believes the earth is flat*, ostensibly recalls those beliefs or reported words from the available context. While in both cases the speaker is relaying another utterance or opinion to the hearer, the degree of incorporation and the hearer involvement are different. While in the case of negating Mary being happy there is a clear argumentative goal of at least questioning an available belief, the reporting verb has no such built-in argumentative function (though it may acquire it, if the subject *Jeremy* is pronounced with contrastive intonation, or if *Jeremy* happens to have expressed a very unfavorable opinion of the hearer, as in *Jeremy says you are a liar*). Furthermore, in the clear-cut metalinguistic cases like *Grandma isn’t feeling lousy*, the hearer is presumably the exact person whose words are being echoed and questioned. Also, while reporting constructions have their own constructional features, they are not argumentative in that they do not require a repair, while the uses in (5), (6), and (7) do.

It might be suggested, then, that while “intersubjectivity” is a broad term covering a number of very different constructions, engaging subjective viewpoints of participants other than the speaker (including, but not limited to, the hearer), specific features of constructions further determine the degree to which an intersubjective set-up is also argumentative. More specifically, the way in which the hearer’s viewpoint is either foregrounded (as in [5], [6], or [7]) or practically absent (as in *He thinks the earth is flat* example) determines the argumentative aspects of the construction. In the same way, constructions that use negation as a stance device rely in very specific ways on the hearer’s subjectivity and discourse participation.

There is then the issue of reporting verbs as stance expressions. Many verbs (*think*, *guess*, *believe*, etc.) have naturally been treated that way based on their meaning (see Kärkkäinen 2003), but can the category be extended over all verbs that can be used as matrix verbs in speech and thought representation,

including *announce*, *declare*, *claim*, *wonder*, *order*, or *mutter*? The answer will, naturally, depend on the definition of stance, but if it is defined as an expression of the represented speaker's viewpoint – epistemic, emotional, or interpersonal – then most of the cases fall into one broad category. Thus verbs such as *announce* or *declare* highlight the interpersonal or speech act aspect of the utterance, while *order* is much more than a performative verb, because it also profiles the represented speaker in an interpersonal relationship where the speaker has authority to issue commands. Moreover, while *mutter* has no influence on how the content is understood, it also represents the evoked speaker as reluctant or unwilling to speak out. This approach would present reporting verbs as participating in a variety of constructions that report the content, while being specific as to some stance taken by the reporting subject, but the “intersubjectivity” of these constructions does not rely crucially on negotiation or argumentation – that is, such constructions may be used to make content originating elsewhere available to the hearer in the course of the exchange, but the hearer's subjectivity and discourse participation are not affected. Unlike in the cases of intersubjective/metalinguistic negation, the hearer's viewpoint is not involved, and the speaker makes no attempt to signal her stance towards the expression quoted. It is then a case of representing multiple viewpoints and the represented speaker's stance, rather than argumentation. However, the use of negation with stance verbs may prompt argumentative meanings.

As should be clear from the discussion so far, subjectified uses of *think* or *guess* need not be seen as essentially different from the representational ones, at least with respect to the stance involved (I would, however, exclude *think about* from the discussion, as a verb of mental activity, not stance). The fact that there are effects of bleaching and pragmatic strengthening, and that the subjectified first person use appears in markedly specific constructional formats, is of course important, but it seems possible to assume that the stance made available by all such uses of *think* is similar enough to allow for some unified discussion, especially with respect to the role of negation. In the remainder of this section, I will consider some more common stance verbs and their interaction with negation.

The syntactic format and interpretation of all complementation constructions with stance verbs relies on the use of the stance verb in the matrix clause, or in the mental space that is higher in the network, and on a description of some state of affairs in the complement clause. The verbs vary in the nature and strength of their commitment, and they also exemplify various combinations of the strength of epistemic stance and emotional stance. At the same time, verbs expressing epistemic stance interact with negation as an independent stance device.

First, it seems useful to further distinguish epistemic stance, as in *think*, from assertive stance, as in *know*. *Know* expresses a stronger stance than that of

think, one that assumes some degree of verifiability or some available evidence, and the epistemic commitment of the speaker is stronger. However, the specific degree or source of certainty required to claim that one “knows” may be a subject of legal dispute, as in the court exchange quoted in *Harper’s Magazine* (November 2007, 26–29). The case concerns an Islamic charity that sued the US government for illegal surveillance, after a document indicating this was accidentally released. The Justice Department argued that the evidence of the document was inconclusive, and that the plaintiffs “think they know, but they don’t,” because only an official confirmation of surveillance could have given them the knowledge, and such a confirmation cannot be given for national security reasons. The argument presented explains that *[w]hen plaintiffs explain what they mean when they say they, in quotes “know,”, they don’t know*. In the plaintiffs’ view, having a document indicating the wiretapping allows them to assert with certainty that surveillance happened, and they are thus relying on the general understanding of the stance marked by *know*. The lawyer argues that they only *think they know* – that is, they have a positive epistemic stance towards their belief, but there is no independent evidence of the belief being assertable as true. In other words, the lawyer presents “knowing” as a stance requiring evidence outside of the speaker’s conviction, but “thinking” as naturally representing the speaker’s mind alone. Stretched as his argument may be, it does rely on the sense of a difference between two stances.

Verbs such as *know*, *think*, *guess*, *doubt*, *hope*, or *wish* can thus be distinguished based on the type(s) of stance they represent (emotional, epistemic, or assertive) and the degree of commitment. As the examples below suggest, a stance verb can have a positive or a negative clause in its scope.

- (9) *know (I know she’s there/she’s not there)*
positive epistemic stance and positive assertive stance
- think (I think she’s there/she’s not there)*
positive epistemic or emotional stance and neutral assertive stance
- guess (I guess she’s there/she’s not there)*
neutral epistemic stance and neutral assertive stance
- doubt (I doubt that/if she’s there;?I doubt that she’s not there)*
negative epistemic stance
- hope (I hope she’s there/she’s not there)*
neutral epistemic stance and positive emotional stance
- wish (I wish she were there/she weren’t there)*
negative epistemic stance and positive emotional stance

While some of the verbs clearly express two stances, not all kinds of stances are marked by specific forms; for example, the emotional stance of *hope* and *wish* does not seem to require specific verb forms, but the negative epistemic stance is obligatorily marked through past verb forms in the embedded clause.

Dancygier and Sweetser (2005) thus point out that in examples such as (10) the layers of past morphology indicate the speaker's lack of expectation of the visit happening, while her positive emotional stance (expressed lexically in *wish*, along with epistemic distance) has no such grammatical indicators.

- (10) I wish I could visit London/had visited London.

Regardless of stance, complement clauses of constructions with stance verbs can contain negation (as the examples in [9] show), and thus suggest that the speaker has reasons to consider the positive counterpart space as a possibility. This is different from clauses that are positive, which are only subject to stance marking imposed by the stance verb. Because negation evokes a cognitively accessible positive scenario, the negative form naturally prompts for a construal where the stance expressed by the speaker emerges in response to an already communicated (and perhaps stance-free) positive construal. For example, *I hope she's there* is a straightforward case of stance marking with respect to the potential space of "her being there," but *I hope she isn't there* suggests that the positive scenario is present in the conversational background. In other words, stance constructions using negation in complementation spaces bring up alternativity in the same way as negative statements like (1) or (2). At the same time, the stance marked by the verb is projected down to the complement space; as a result, in cases like *doubt* (*I doubt that she isn't there*), negative stance of the stance verb "doubles" the negative part of the two-space construal in the complement space, and the stance of the construction overall may require more contextualization. Interestingly, this presents no problem with *wish*, where the primary stance is the emotional one, which does not directly clash with negation in the complement clause. However, negation not only gives rise to alternativity, but is also itself a stance marker, and thus interacts with stance verbs in interesting ways.

At the same time, stance constructions with negated complement clauses seem to be less sensitive to the source of alternativity. While examples like (6) and (7) were claimed to differ from (1) in terms of the direction in which alternativity is used in the construction (where negation either alerts the hearer to the alternative set-up or proposes a response to the hearer's original formulation), this distinction seems to be backgrounded in most of the negative examples in (9). In other words, *I wish she weren't there* may be said either because the speaker has some knowledge about "her" whereabouts, or because the hearer suggested it. However, in the case of *doubt*, it seems more likely that the negative complement clause relies on contextually available knowledge – possibly because of the rarity and/or markedness of the situation of expressing negative stance with respect to a negative belief. *Wish* does not pose such problems, because of its strong positive emotional stance, which overrides the negative epistemic stance.

The mental space set-up of stance constructions in general can thus be described as follows: the stance space is higher in the network, and it projects its stance into the complement space. The complement space can be split further into an alternative stance set-up when negation is present. Such a configuration seems to be closely correlated with the syntactic choices in the constructions – the stance verb is the matrix verb and the complement space is syntactically subordinate to it. However, there is also the question of the specific place of stance in a mental space network. Sweetser and Ferrari (this volume) discuss a number of cases of subjectified expressions in which stance has to be marked in the network, even though there may be no dedicated expression marking stance in the construction. This raises important questions about representation of stance in cases of advanced subjectification, and those where subjectification is less of an issue.

There seem to be several possibilities here. First, there are expressions like *Goodbye*, where grammaticalization processes have made the original constructional features invisible. Then there are cases where space builders such as *if* prompt the entire network of spaces that remains in the space builder's scope and thus is marked with stance. We can possibly extend this case to stance verbs like *think*, *doubt*, or *wish*, which also build their complement spaces and project stance to them. But there are also stance markers that appear on the clausal level, such as past tense in conditionals. For instance, *If I had a cat, I'd always feed him tuna* uses the past tense to strengthen the neutral stance of *if* into a negative stance, thus opening the sentence to an interpretation whereby the speaker does not have a cat. What such cases make clear is that stance is a composite phenomenon, and that a mental space network marked with stance by its space-builder such as *if* (which marks neutral rather than positive stance) is subject to further stance marking by clause-level forms (such as tense or negation). This opens a number of questions about the ways in which stance and subjectification are marked and how the markers are or are not constructionally salient. In this chapter, I am particularly interested in stance markers that also distinguish types of constructions.

In the cases where there is a stance-relevant space builder, as in the case of stance verbs, the constructional effects are visible in the embedded structure. For example, the influence of the higher stance space can be seen in the choice of the complementizer. In most cases, a zero-complementizer is acceptable, and then stance is not exerting any influence over the choice. But in the cases where the verb expresses negative or negated stance (like *doubt*, or *don't know*), *if* is the preferred complementizer, marking non-positive stance similarly to conditionals. As Dancygier (1998) suggests, *if* is a generic marker of unassertability, or neutral/negative stance, and it plays a similar stance-weakening role both in conditionals and in reported clauses (embedded questions and the like).

It also seems possible to argue that representational and subjectified uses of stance verbs like *think* both rely on the same set-up, and the difference between them is mostly visible in the first person versus third person usage. *I think he'll win* and *He thinks he'll win* may be perceived as different because the use of *I think* is more easily viewed as subjectified (as can be seen, for instance, in its near functional equivalence to a modal adverb such as *possibly* or *perhaps*),¹ and thus is primarily a stance expression. But they can also be viewed as quite similar in that epistemic stance is attributed in each case, but it is attributed to the speaker only in the first person usage. In the third person construction (as Verhagen shows), the speaker is not expressing her stance, but perhaps guessing, or perhaps reporting the stance attributed to another subject. The greater syntactic flexibility and grammaticalized status of *I think* seems to result from the simpler mental space configuration, with no need to use the speaker as the mediator between another speaker's stance and what is communicated to the hearer in the complement space. Furthermore, being more grammaticalized than other expressions with stance verbs, *I think* marks a range of positive stances, from epistemic evaluation, through degree of commitment (e.g. to a future action), to level of imposition (e.g. when a positive belief is toned down to a non-assertive stance).

3.3 Neg-Raising controversy

It has been argued repeatedly (also within Cognitive Grammar; cf Sumnicht 2004) that sentences such as *I think she's not there* and *I don't think she's there* mean (roughly) the same, as negation is “raised” from the embedded clause to the main clause. While the meaning similarity is easily recognizable, it is not clear why the direction of the so-called Neg-Raising should be up to the *think*-clause, and not down to the “being there” clause. It has also been noted that the versions differ in interpretation: *think that not X* expresses a lower level of uncertainty than *not think that X*; and, finally, it is not always the case that the pairs are indeed naturally interpreted in the same way. For example,

- (11) “I don't think I have anything to add.” (Trent Lott, commenting on Bush's Iraq plan)

is not readily interpretable as *I think I don't have anything to add*. Lott seems to be politely refusing to offer more comments, and thus the presence of *I don't think* is better explained as an added expression of subjectivity and not a matrix clause. Similarly, Example (12)

- (12) “I don't think I'm going to get into the hypotheticals, I'd rather wait and see.” (Donald Rumsfeld)

does not equal *I think I'm not going to get into the hypotheticals*. As Ferrari and Sweetser (this volume) point out, the expression *gonna* is itself a marker of subjectified meaning, so its interaction with *I think* is even more complicated in terms of stance, but Rumsfeld, like Lott, seems to be politely refusing something that reporters want him to do.

These remarks are further supported by Croft's (2001: 213) discussion of an attested example, *I don't think that ever before have the media played such a major role in a kidnapping*. As Croft points out, such examples cannot be explained via an implicature account advocated by Horn (1978), since an analysis such as "It is not the case that I think that ever before have the media played such a major role in a kidnapping" is not quite acceptable. I would like to add that the next step of the implicature account, whereby such a formulation leads to "I believe that never before have the media played such a major role in a kidnapping" is also not sufficiently accurate, since the epistemic stance of the *I don't think that X...* variant (even if it were acceptable) is much less positive than the *I believe not X* one. The assumption that there is little essential difference between *think* and *believe*, and that one can substitute for the other when negation is strategically placed to highlight the intended viewpoint, distorts a whole range of linguistic facts with regard to the concept of stance, and the numerous subtle varieties of levels and kinds of commitment. Finally, the implicature account does not add anything to the understanding of negation with stance verbs, except "translating" a somewhat outdated idea of a syntactic transformation into a pragmatic concept that has been in need of serious revamping since Relevance Theory. It seems that *I don't think* can be treated as a stance marker with respect to the entire syntactic structure in its scope, and thus any approach dissecting negation out of it distorts its actual usage.

The main issue in distinguishing differently negated versions of *I think* constructions is first establishing what negation is doing in the stance space and the complement space. In terms of alternative space set-ups, it plays a different role depending on the level of embedding in the network. *I don't think* negates the stance the verb expresses, and thus presents the speaker's stance as negative (rather than positive), in a situation where positive stance might be expected or desired (alternative space). In the cases of (11) and (12), the stance expressed is that of weakening the speaker's commitment to make a refusal sound polite, but neither expression can be naturally paraphrased with an expression of positive stance. In other words, negative stance towards doing X is not the same as a positive commitment towards not doing X, because in the latter case the alternativity is marked elsewhere. The difference is clearly seen in other instances of stance marking, such as the contrast between *You mustn't smoke here* and *You don't have to smoke here* (obligation to not smoke versus lack of obligation to smoke), and though it is less clear in the case of *I think*,

it is still a valid distinction. With the stance marked by *I don't think* projecting downwards into the complement space, it is natural to interpret the situation described there as negative (e.g. in [11], the speaker will presumably not say much more), but the alternativity is not between having something to add or not, but between wanting to continue one's comments or not, and not wanting to do so naturally implies not doing it. The use of negation described above further confirms that, while negation is a stance-related marker, it is also the primary marker of alternativity, and as such can add alternativity to expressions of stance, while also adding specificity to how stance is expressed.

Corpus search reveals that sentences with positive stance towards negative complement spaces (*I think I don't X*) are not common; also, those that occur often express another stance in X.

- (13) I think I don't have a problem with gangster movies per se, it's just that some of them are too gory (ramblings-n-more.blogspot.com/2005/09/couple-of-good-films.html)

Possibly the most common combination is something like *I think he doesn't want X*, where negation cannot really migrate upwards, because the wrong stance of the wrong subject would thus be negated.

To sum up, a Neg-Raising account does not explain the ways in which *think not X* versus *not think X* sentences differ. It also does not account for other uses of *think* with negation (see below); it does not extend naturally to include other stance verbs; and it does not explain the interaction between negation and stance.

3.4 Negation in the stance space

I will thus assume that expressions in which a stance verb appears with negation genuinely profile an alternative-stance space set-up, rather than merely projecting negative stance into the complement. That is, an expression such as *I don't know* sets up two alternative stance spaces, "knowing" and "not knowing," and the stance (positive or negative) is then projected into the complement space (*she is there*).

- (14) I don't know if she's there.
 I don't think she's there.
 ??I don't guess she's there.
 ?I don't hope she is there.
 I don't doubt she is there.

Some "special effects" need to be noticed here, especially the difficulty in finding natural contexts where *hope* and *guess* can be negated. While essentially any stance can be questioned, it is possible that expressions of neutral stance

are less likely to be negated, given that there is so little commitment involved in the stance verb itself. *Guess* is a less troublesome example here, since it seems to be used now as a somewhat less committed variant of *think*. It is not common in standard discourse, but it clearly emerges from the subjectified *I guess* – a cousin of *I think*.

- (15) I still live in a nice house, so **I don't guess** I'm broke. (www.scenedaily.com/news/articles/sprintcupseries/waltrip_i_dont_think_were_doing_anything_wrong.html)

However, it is worth noting how contextually available assumptions can affect the interpretation of *I don't think*.

- (16)
- Tucker, does the strategy seem to be working to you?
 - Yes. **I don't think** it's a foolish strategy at all... To this point Kerry's vision consists of getting the UN in there, which **I don't think** most serious people **think** is a solution.

In this fragment from a CNN interview, Tucker Carlson is gently prompted to criticize the US strategy towards assumed weapons of mass destruction in Iraq. He does not go along with it, and thus his first *I don't think* is a form of a rebuttal, something like “you suggest I might think it is a foolish strategy, but I don't.” It is thus negating a positive stance towards the evaluation of the strategy as foolish, but it is also responding to the interviewer's suggested attribution of the positive stance to Tucker. In a way, the wording is expressing a negative stance both subjectively and intersubjectively. This is different from his second *I don't think*, which is a straightforward and unprompted expression of negative stance (towards other people's stance). The “double” stance marking of the first *I don't think*, however, calls for an explanation. It seems that the alternativity of negation opens such a possibility. The negated stance is an expression of the speaker's subjectivity, but the alternative positive stance towards the same complement space is an alternative that has been suggested and is now overtly rejected. Negation in the stance space may thus be different from negation in the complement space in that it exploits alternativity for more complex stance attribution. As I suggested above, there are two ways in which alternativity of negation is used: either one uses negation to alert the hearer to the alternative (as in *There's no milk in the fridge*), or a contextually activated assumption prompts negation as a way to overtly state the alternative, possibly as an expression of the speaker's belief or stance. In the case of the stance space, which is higher in the network, the available contextual alternatives are not located lower in the same network, but are intersubjectively picked up from another network in the discourse. As a result, they are naturally read as refutations of another subject's stance or intended attribution of stance, depending on the direction of activation. Example (17) illustrates such cases further.

- (17) – What **do you hope** the mainstream audiences will learn about this artistic/bohemian culture . . .
 – **I don't hope** they learn anything. . . I don't know who goes to see films and why. (<http://romanticmovies.about.com/library/weekly/aa042602b.htm>)

The interviewer's question presents the speaker as "hoping," and asks for a specific elaboration of the complement space of *hope*. The speaker, however, rejects the stance as part of his intention in film-making. The expression *I don't hope* is thus formulating a negative alternative to the assumed hopeful stance suggested by the interviewer. As a result, the director refuses to answer the question as it was stated.

A particularly interesting case of intersubjective negation is presented in (18).

- (18) I had an abortion when I was 20 years old. I don't regret it at all. Sometimes I think about how my life would have been different had I not aborted, but **never in the sense that I wish my life were different** than it is now. **I don't wish** I had ended up on welfare, **I don't wish** I'd created an emotional and financial strain for my family and **I don't wish not to have** the wonderful partner I have now. (www.imnotsorry.net/2010/09/06/marnis_story)

The blogger comments on the ways in which having an abortion helped her avoid stressful consequences of having an unwanted child. When she refers to the alternative "abortion space," though, she does not frame it as her presumed opponents might (they might think she feels remorse and actually regrets having done it). On the contrary, she builds on the counterfactual meaning of *wish* by describing all the unwanted situations that she thinks she has avoided. While she uses *I don't wish* to reject the regretful thoughts attributed to her, she is in some sense mocking and refuting what the opponents of abortion might be saying to her. Consequently, her repeated negation of *wish* has a double argumentative function – rejecting regret, and also fleshing out her reasons why she still stands by her choice. Interestingly, in the final clause (*I don't wish not to have . . .*) the complement space is itself negative. I will discuss more such cases below.

The mechanism of refutation that relies on negating attributed stance is an interesting instance of a variety of language processes relying on the so-called "echoic mention" (Sperber and Wilson 1991 [1981], 1995 [1986]). While the literal understanding of echoic use as a repetition of an utterance that can easily be accessed in the context is often weakened to various kinds of discourse grounding and evocation (see Tobin and Israel and Vandelanotte in this volume; Dancygier 1998; Dancygier and Vandelanotte 2009), the case of

stance refutation may be even further removed from it in that often it is not in the basic sense a “recall” of a specific utterance, but may be a speaker’s formulation of what the speaker perceives as a stance attributed to her. Besides the pragmatic specificity of these uses, we should also note the role of salient constructional features, recurring in a range of constructions.

3.5 Negation in both spaces

I argued above that negation in the stance space has a different function from negation of the complement space. The differences are best seen in examples where both spaces are negated. Such examples are quite rare, and they all seem to suggest a similar set-up.

- (19) **I don't think I won't have slumps** like this again, but I will never fully give up on myself. (Blogger describes hardships of dieting; www.blurry.com/~lil_miss_chub/day/2003/11/30)
- (20) I don't think about nor am I concerned with political correctness.
I don't think I won't offend someone. (www.seleda.com/feb01/bawza.shtml)
- (21) **I don't think he can't write songs**, either. It's just that, (in my opinion, I never said everybody's) he should get somebody else to sing and play. (www.radioparadise.com/content.php?name=songinfo&song_id=2787)

All these examples are argumentative, responding to assumptions activated in the background. In (19), the writer responds to the assumption “in the air” that she might believe (perhaps naively) that she “won’t have slumps like that.” She rejects the attribution of that belief, knowing that slumps might happen in spite of her best intentions, but she is nevertheless optimistic about the final result.

Examples (20) and (21) are similar: someone may assume “I think he can’t write songs” or “I think I won’t offend,” but in fact this is not my assumed epistemic stance. Such constructions are thus intersubjective on both levels. The negative belief in the complement space is being attributed to the writer, in response to the previously activated positive alternative (you had a slump once, so you will have more; if you act this way, you expect to offend; people think he can write songs). The writer or blogger then rejects the attribution of the negative belief to her, by negating her stance to it. The overall interpretation might be in agreement with the general “two negatives make a positive” rule, but the complex construal offered does much more than this – it develops the argument from its inception all the way to the blogger’s standpoint.

Even more importantly, sentence (20) is using a positive polarity item *someone* in its negative embedded clause (*I won't offend someone*). This seems to result directly from the reading I propose – “Contrary to what is being suggested, I think I might offend someone” – and to confirm the fact that the negation in the clause does not have the predicate *offend someone* in its scope, and it marks stance only. In other words, the situation of “not offending anyone” is not under discussion at all. The blogger acknowledges that others may expect her to be concerned about potential offences, but she rejects that concern, and so the two uses of negation reject these two stances (the expectations others have and her rejection of these expectations as valid). It is thus possible for the positive polarity item to remain unaffected under negation in its clause.

What seems to be interesting here is that such examples of “double” negation (in the stance space and the complement space) are in fact available only in cases where the complement space is not a straightforward description of a state of affairs subject to stance evaluation, but counts as stance expression itself. In a sense, the intersubjective set-up described above (with the positive alternative activated in the discourse) is doubled in (19)–(21) – it appears once in the stance space, and then in the complement space. Such examples are extremely rare, but they are possible when discourse makes such an argumentative set-up possible. We should note further that with the argumentative role of negation in both spaces, stance marking is not only doubled, but also “stacked.” The negative stance expressed in the higher clause/space projects into the complement space, but the complement space is marked for stance as well, and thus it becomes possible to express a stance towards a stance. This is possible when the construction can dedicate a clausal level to each stance expression, and even if the only stance marking in the lower space is via negation. The interpretation then emerges through the interaction between alternative (negative) set-ups in the network and assumptions available elsewhere in the discourse – for example, in the hearer’s words or assumed beliefs of other participants. In other words, the alternativity of negation is the source of argumentative uses of stance, as it allows for projections of stance not only up and down the network, but also “sideways” – through the alternatives. When the alternatives are signaled in the discourse, the argumentative function of negation prevails.

3.6 Constructional compositionality and stance

As the above review of negation suggests, its uses, while invariably relying on alternativity, interact with stance, and also express stance. Negative stance may be signaled in different ways – past tense verb forms and lexical expressions prevail – but it may be suggested via alternativity as well, for the purposes of argumentation. Even more importantly, there are constructional aspects of

stance expression that need to be considered. In earlier work, I have argued for the constructional nature of so-called metalinguistic negation (refutation and repair), but there are many other constructional formats to be considered. For example, stance adverbs such as *admittedly*, *possibly*, or *undoubtedly* can appear sentence-internally or sentence-initially, and the sentence-initial ones, when separated by intonation, are often described as disjuncts – adverbials that provide the speaker's comment on the entirety of the sentence. The syntactic separation of disjuncts is often connected to their scope over the whole sentence, but it should be added that their constructional slot is also paired with the highest position in the mental space set-up, which often holds stance spaces. Other syntactic constructions using stance adjectives as predicates behave similarly, as can be seen in Verhagen's (2005) discussion of so-called *it*-extraposition ("It is possible that X" format instead of "That X is possible"), which suggests that the construction favors an argumentative set-up. The evidence provided by negation, disjuncts, or extraposition clearly supports the need for a broad discussion of constructional correlates of argumentative and stance-related meanings.

One of the questions raised above concerned the actual argumentative force of constructions that Verhagen labels "intersubjective." It seems clear from the review above that there is a cline of argumentative or inferential impact among such constructions. While *John thinks that X* does require that X be construed from the point of view of John, any further complication of the constructional allocation of stance (via negation in any of the spaces or both, choice of reporting verb and complementizer, the change of pronoun to *I*, which opens the construction to subjectified reading, etc.) adds a level of argumentative complexity. What I want to suggest is that the concept of intersubjectivity is best approached not only in terms of types of constructions that clearly involve the speaker's and the hearer's subjectivities, which is what Verhagen has proposed, but also in terms of constructional compositionality. The devices I have looked at here (very briefly in some cases) include various uses of negation, stance verbs, complementizers, clauses, and pronouns. What seems to emerge from the overview is the observation that argumentative complexity increases with constructional complexity. In other words, all these devices bring their own stance-related contributions to the expressions, but stance marking itself, while possibly intersubjective, is not necessarily argumentative. The argumentative power of a construction emerges out of the complexity of the constructional means used to represent different stances. Stance verbs themselves are not high on that scale, but when accompanied by negation in one of the spaces, they increase their argumentative power, and the constructions with negation in both spaces are clearly the ones negotiating stance across different mental spaces and subjectivities. These correlations seem to support the idea of constructional

compositionality (Dancygier and Sweetser 2005), in that the argumentative complexity of a construction emerges out of the lower-level constructional choices that prompt specific construals.

This view is supported by examples of constructions that increase stance-allocation complexities even further. These are specialized constructions that place the stance expression in a higher space and also in a higher clause – hence I will refer to them as *stance-stacking constructions*. The cases of disjuncts and *it*-extraposition mentioned above seem to have just that effect: making sure that the stance is located in the highest syntactic slot and projects over the utterance as a whole. The cases of negation discussed above are more complex, because their task is to make the scope of stance more specific, possibly applying to a phrase and not the whole clause. Classical examples of external negation are a case in point, as in *It is not the case that X* versus *not X*. In fact, such constructions can easily be used as a paraphrase of some of the examples discussed above, as in *It is not the case that/I'm not saying that/I'm not suggesting/It is not that I think that he can't write songs*. Furthermore, negation is the only stance expression that can be “extracted” out of a counterfactual conditional, which relies on negative stance. For example, a sentence such as *If you hadn't helped, I would have been fired* can be paraphrased as *If it hadn't been for your help, I would have been fired*, or *If not for you, I would have been fired*. Apart from focusing on a selected aspect of a conditional space, the construction *if not for X* (see Dancygier and Sweetser 2005) seems to have no other role but to highlight the negative stance by extracting it out of the scope of the subject of the *if*-clause and thus making it clear that the stance is expressed by the speaker, rather than being attributed to the subject.

What such constructions do is set up a new clause at the head of the construction, with verbs like *be*, *say*, or *suggest*, which can be construed as upholding negative stance to a possible interpretation of a discourse fragment without adding much lexical content. The additional negative marker of stance is then placed in the clause, to separate this level of stance marking from stances potentially represented lower in the construction. The fact that these constructions use negation in a higher clause specifically added for that purpose further reinforces the question of the correlation between the placement of stance in constructions and its participation in grammaticalized expressions such as those discussed by Ferrari and Sweetser (this volume).

There are clearly many ways in which a construction may represent stance, but there are still many unanswered questions about the possible correlation between constructionally available overt means of precise allocation of stance to various discourse participants, real or construed, and the subjectification processes that in fact “hide” stance from the constructional form. While the cases discussed here occupy the “constructionally overt” end of the spectrum,

the cases discussed by Sweetser and Ferrari address the other end, and the correlation emerging out of the comparison clearly suggests that the need for stance to be projected to spaces lower down in the network makes it natural to suppose that stance spaces should appear high in the network, whether there is a constructional correlate to this phenomenon or not. Interestingly, the case of *I think* is particularly helpful here. Distinguishing the representational uses of *I think he lied* from the subjectified ones is largely based on the meaning of the complement clause – the more interpersonal its impact, the more likely *I think* will be read as subjectified. However, there are also constructional correlates (see Thompson and Mulac 1991; Vandelanotte 2009), such as the possibility of the *I think* clause being moved to the end of the sentence (*He lied, I think*). What this constructional variant suggests is that the subjectified/grammaticalized *I think* may not require to be placed at the head of the construction, and it is still clear that the stance affects the matrix clause *He lied*, rather than the next sentence. This may in fact be a correlation that further research might test – the higher the level of grammaticalization, the fewer constructional correlates of subjectification. In other words, complex stance markers such as *I think/I don't think* remain “outside” the clausal syntax.

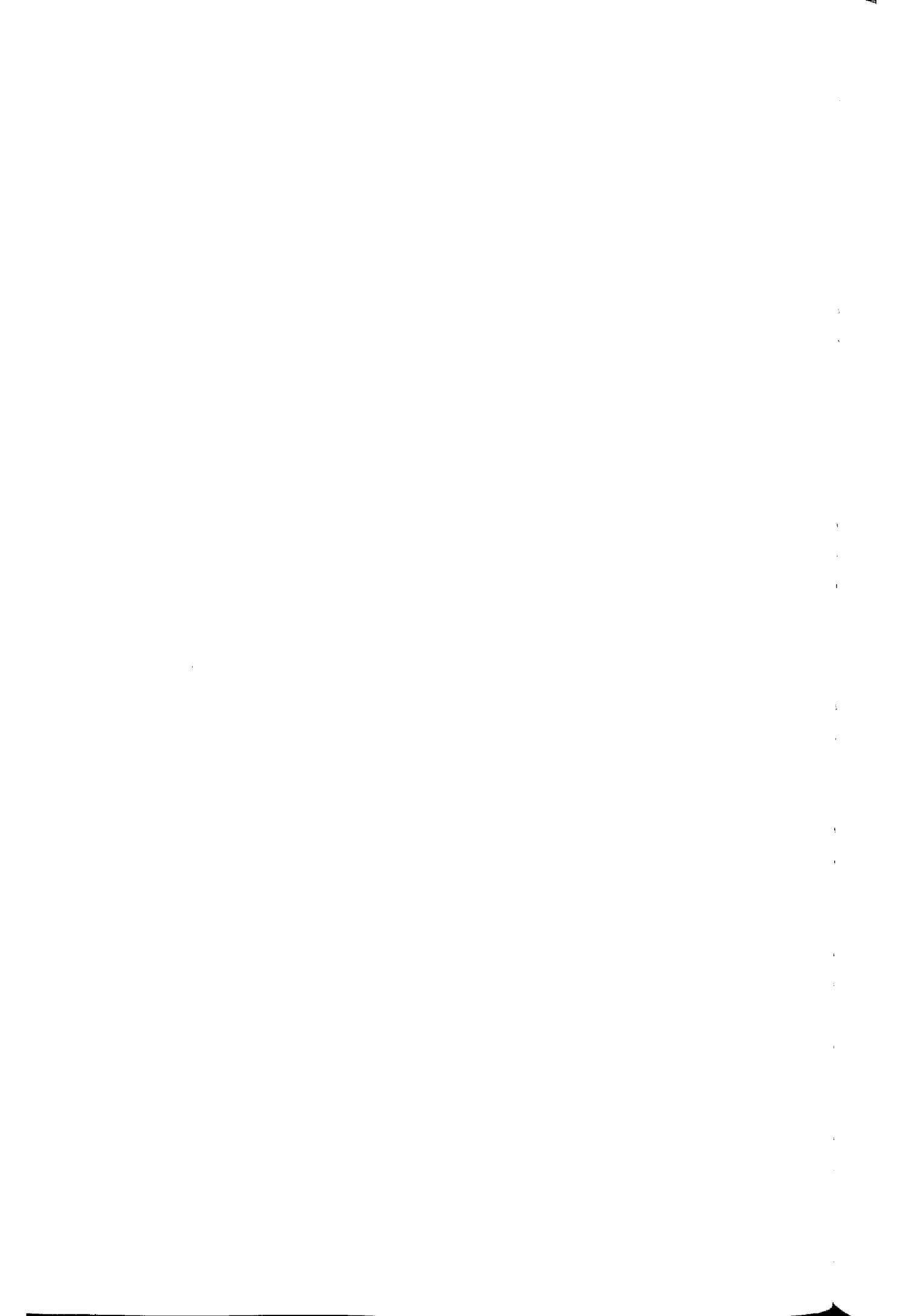
Also, we can further postulate that in a narrative flow a sentence *He lied* can easily be interpreted as a case of free indirect discourse, such that the thought about a lie is part of a stretch of narrative where a character's thoughts are being represented. In such a case, the stance remains that of the narrative “cognizer,” but it is difficult to distinguish whether the utterance should be treated as subjectified or representational, and what remains constant is the fact that the cognizer's epistemic stance is being represented somehow. One explanation for this ambiguity might be that in the absence of constructional signals, what remains salient is the viewpoint associated with the current cognizer, and the viewpoint assumes the availability of some epistemic stance, without necessarily distinguishing subtle levels of stance commitment. The conclusion might be that compositional constructions provide means for very precise allocation of stance, while grammaticalized forms package stance economically into expressions, but there is also a broad area of usage where interpreting a text can proceed smoothly with just a minimal assumption of epistemic viewpoint. This kind of usage may in fact be the default case, similar to Langacker's Ground, where the details of stance are not salient enough to warrant a complex construal. It might be argued, then, that the constructions that Verhagen refers to as intersubjective explicitly evoke multiple viewpoints in the Ground (which otherwise remain assumed as default), while the specific argumentative impact of any such construction further relies on one of the two options outlined above: either there are constructional means (verb forms, stacked clauses, refutation/repair sequences, etc.) that make stance allocation more specific, or stance and commitment are packaged into grammaticalized forms of subjectification.

References

- Biber, Douglas and Edward Finegan. 1988. Adverbial stance types in English. *Discourse Processes* 11:1, 1–34.
- Brinton, Laurel. 2008. *Comment Clauses in English*. Cambridge University Press.
- Bybee, Joan, with Joanne Scheibman. 2007. The effect of usage on degrees of constituency: the reduction of *don't* in English. In Joan Bybee. *Frequency of Use and the Organization of Language* 294–312.
- Coulson, Seana and Todd Oakley. 2005. Blending and coded meaning: literal and figurative meaning in cognitive semantics. *Journal of Pragmatics* 37, 1510–36.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford University Press.
- Dancygier, Barbara. 1992. Two metatextual operators: negation and conditionality in English and Polish. *Proceedings of the 18th Annual Meeting of the Berkeley Linguistics Society* 18, 61–75.
1998. *Conditionals and Prediction: Time, Knowledge and Causation in Conditional Constructions*. Cambridge University Press.
2010. Alternativity in poetry and drama: textual intersubjectivity and framing. *English Text Construction* 32, 165–84.
- Dancygier, Barbara and Eve Sweetser. 1997. Conditionals, distancing, and alternative spaces. In Adele E. Goldberg (ed.). *Conceptual Structure, Discourse, and Language*. Stanford, CA: Centre for the Study of Language and Information, 83–98.
2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Dancygier, Barbara and Lieven Vandelanotte. 2009. Judging distances: mental spaces, distance and viewpoint in literary discourse. In Geert Brône and Jeroen Vandaele (eds.). *Cognitive Poetics: Goals, Gains, and Gaps*. Berlin: Mouton de Gruyter, 319–70.
- Davies, Eirian C. 1979. *On the Semantics of Syntax: Mood and Condition in English*. London: Croom Helm.
- Fauconnier, Gilles. 1994 [1985]. *Mental Spaces*. 2nd edn. Cambridge University Press.
1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fillmore, Charles J. 1990a. Epistemic stance and grammatical form in English conditional sentences. *Proceedings of the 26th Annual Meeting of the Chicago Linguistic Society* 137–62.
- 1990b. The contribution of linguistics to language understanding. In Aura Bocaz (ed.). *Proceedings of the 1st Symposium on Cognition, Language, and Culture*. University of Chile, 109–28.
- Heller, Joseph. 1955. *Catch 22*. New York: Simon & Schuster, Inc.
- Horn, Laurence. 1978. Remarks on Neg-Raising. In Peter Cole (ed.). *Syntax and Semantics, Vol. 9: Pragmatics*. New York: Academic Press, 129–225.
1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61, 121–74.
1989. *A Natural History of Negation*. University of Chicago Press.
- Jeffries, Lesley. 2010. *Opposition in Discourse: The Construction of Oppositional Meaning*. London: Continuum.
- Kärkkäinen, Elise. 2003. *Epistemic Stance in English Conversation*. Amsterdam: John Benjamins.

- Langacker, Ronald W. 1985. Observations and speculations on subjectivity. In John Haiman (ed.). *Iconicity in Syntax*. Amsterdam: John Benjamins, 109–50.
1987. *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*. 2 vols. Stanford University Press.
1990. Subjectification. *Cognitive Linguistics* 1, 5–38.
2006. Subjectification, grammaticalization, and conceptual archetypes. In Angeliki Athanasiadou, Costas Canakis, and Bert Cornillie (eds.). *Subjectification: Various Paths to Subjectivity*. Berlin/New York: Mouton de Gruyter, 17–40.
- Oakley, Todd. 2005. Negation and blending: a cognitive rhetorical approach. In Seana Coulson and Barbara Lewandowska-Tomaszczyk (eds). *The Literal and Non-literal in Language and Thought*. Berlin: Peter Lang Publishing.
- Sperber, Dan and Deirdre Wilson. 1991 [1981]. Irony and the use-mention distinction. In Steven Davis (ed.). *Pragmatics: A Reader*. Oxford University Press, 550–63. [Originally in Peter Cole (ed.). *Radical Pragmatics*. New York: Academic Press, 295–318.]
- 1995 [1986]. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Sumnicht, Anne. 2004. A new look at negative raising. In Augusto Soares da Silva, Amadeu Torres, and Miguel Gonçalves (eds.). *Linguagem, Cultura e Cognição*. 2 vols. Coimbra, Portugal: Livraria Almedina, vol. II, 607–26.
- Sweetser, Eve. 1996. Mental spaces and the grammar of conditional constructions. In Gilles Fauconnier and Eve Sweetser (eds). *Spaces, Worlds, and Grammar*. University of Chicago Press, 318–33.
2006. Negative space: levels of negation and kinds of spaces. In Stéphanie Bonnefille and Sébastien Salbayre (eds.). *Proceedings of the conference “Negation: Form, figure of speech, conceptualization”*. Publication du groupe de recherches anglo-américaines de l’Université de Tours. Tours: Publications universitaires François Rabelais.
- Szymborska, Wisława. 1997. *Nothing Twice: Selected Poems/Nic Dwa Razy: Wybór wierszy*. Translated by Stanisław Barańczak and Clare Cavanagh. Krakow: Wydawnictwo Literackie.
- Thompson, Sandra and Anthony Mulac. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Elizabeth Closs Traugott and Bernd Heine (eds). *Approaches to Grammaticalization, Vol. II: Types of Grammatical Markers*. 2 vols. Amsterdam: Benjamins, 313–29.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65:1, 31–55.
1995. Subjectification and grammaticalization. In Dieter Stein and Susan Wright (eds.). *Subjectivity and Subjectivisation*. Cambridge University Press, 37–54.
2003. From subjectification to intersubjectification. In Raymond Hickey (ed). *Motives for Language Change*. Cambridge University Press, 124–42.
- Traugott, Elizabeth Closs and Richard Dasher. 2005. *Regularity in Semantic Change*. Cambridge University Press.
- Vandelanotte, Lieven. 2004a. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36:3, 489–520.
- 2004b. From representational to scopal “distancing indirect speech or thought”: a cline of subjectification. *Text* 244, 547–85.

2009. *Speech and Thought Representation in English*. Berlin and New York: Mouton de Gruyter.
- Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford University Press.
2006. On subjectivity and “long distance Wh-movement”. In Angeliki Athanasiadou, Costas Canakis, and Bert Cornillie (eds.). *Subjectification: Various Paths to Subjectivity*. Berlin/New York: Mouton de Gruyter, 323–46.
- Wierzbicka, Anna. 2006. *English: Meaning and Culture*. Oxford University Press.



Part II

Gesture and processing of visual information



4 Interactions between discourse status and viewpoint in co-speech gesture

Fey Parrill

Viewpoint is a delightfully messy area, as the introduction to this volume makes clear. While the majority of the chapters in this volume center on analyses of specific viewpoint phenomena, the central goal of this chapter is to provide some traction on the issue of viewpoint in multimodal language (speech and speech-accompanying gestures) – that is, I attempt to provide both a global picture of the world of viewpoint, as well as a more systematic description of how viewpoint in gesture can be understood.

In what follows, I begin by separating viewpoint into three phenomena: conceptual, linguistic, and gestural. I then describe a general notion of viewpoint (as discussed by researchers like Kuno, MacWhinney, and Chafe), and viewpoint in the *blending* or *conceptual integration* framework. I argue that viewpoint as seen in co-speech gestures can be helpful in unifying these approaches. Finally, I describe an experimental study demonstrating that both linguistic and gestural viewpoint can be affected by discourse status. This study highlights the ways in which the two modalities (speech and gesture) are linked.

4.1 Conceptual, linguistic, and gestural viewpoint

The term *viewpoint* is often defined as a conceptualizer's perspective on an event or scene (Chafe 1976; DeLancey 1981). This is a problematic definition, as it is unclear what perspective is. Many researchers define perspective as point of view, and the two terms are often used interchangeably, leaving us with a circular and unsatisfying definition. The construct of viewpoint is an important one, however ill-defined. It captures the fact that humans represent events from some experiencer's consciousness, and that they are able to project their own visual perspectives, thoughts, and feelings onto other entities. Most researchers seem to agree that viewpoint is, in its most basic form, spatial. That is, viewpoint is most essentially about a perceiving entity's physical perspective on an event or scene. The metaphors used in some discussions of viewpoint (e.g. *viewing arrangement*, *camera angle*, *vantage point*) reflect that fact. The perceiving entity's physical perspective is thought to result in the use of certain linguistic devices, thus making a spatial system linguistic. Viewpoint can also

be extended to include representations of epistemic states – that is, a particular physical location (say, close enough to another entity to be able to hear it speak) can make one privy to information that other entities do not have access to. Viewpoint can thus involve knowledge. This fact is often salient when considering written texts, which have narrators who have points of view, and representations of characters' points of view, and so on. In short, the construct of viewpoint is applied to a wide range of phenomena, some of which are about physical locations and some of which are about much more abstract conceptual structures.

The study described here uses cues from a part of the language system that is non-linguistic, namely co-speech gestures. These gestures are intricately linked to speech, but are not governed by a grammar and are not symbolic. Because co-speech gestures are spatial, however, they can provide a bridge between viewpoint in the sense of a conceptualizer's location, and viewpoint in a more abstract linguistic sense.

In considering how viewpoint might be reflected in communicative behavior, it is helpful to consider the phenomenon as having three distinct dimensions: the real or imagined physical location of a conceptualizer (referred to as *conceptual* viewpoint), the linguistic manifestations of that location (referred to as *linguistic* viewpoint), and spatial cues to that location as seen in gesture (referred to as *gestural* viewpoint).

4.2 Conceptual viewpoint

Conceptual viewpoint can refer to a mental representation based on visual perception of a current physical location. For example, from my office window I can see the Cleveland Museum of Art, and I would use certain language in describing its location to you as a result of the visual angle I have. More often, however, viewpoint is about an imagined physical location. That is, I might need to describe the Cleveland Museum of Art to you while in California. Cognitive science has long had the apparatus to explain how humans generate mental images, and we know that mental images retain many of the properties of real visual images (Finke 1989; Kosslyn 1994). A mental representation of a visual scene may thus have an implicit or explicit viewing angle, just as a real visual scene does (Kosslyn 1994). Recent research suggests that mental images and motor programs are integral to using language – that is, language use involves running a mental simulation (a partial reconstruction) of what is being talked about (Tanenhaus *et al.* 1995; Kaschak and Glenberg 2000; Glenberg and Kaschak 2002; Zwaan *et al.* 2002; Richardson *et al.* 2003; Borghi *et al.* 2004; Gibbs 2005; Bergen *et al.* 2007). For example, describing a person walking around the Cleveland Museum of Art will involve generating mental images of the building and activating motor programs for walking. Simulations can vary in

viewpoint as well. One can simulate an event from the point of view of an actor in the scene or from a more distanced point of view (that of an observer). This distinction – between the point of view of an entity within the scene and that of an observer outside the scene – has received a fair amount of attention, and is thought to have fairly direct links to certain linguistic phenomena. I will return to this point below.

As noted above, conceptual viewpoint can be extended to include access to knowledge, because of the link between being physically present in a certain location and being privy to information. Thus, conceptual viewpoint can be understood as the way a language user is mentally simulating an event, and simulations may include mental images, motor programs, and representations of mental states (both the conceptualizer's own and those of other entities). Conceptual viewpoint can be assessed using behavioral measures (e.g. eye-tracking, response time). For example, Bryant and Tversky (1999) have shown that when people are presented with two-dimensional diagrams, they tend to take a point of view outside the scene. When presented with three-dimensional models, participants tend to take a point of view within the scene. This difference in viewpoint results in reliable differences in response time to questions about items contained within the scene.

Linguistic devices that mark viewpoint can be considered instances of *linguistic* viewpoint. Theoretically, linguistic viewpoint reflects conceptual viewpoint, if only partially or schematically. A great deal has been written about viewpoint in language. Several studies have investigated how reading certain kinds of texts results in taking a particular point of view (e.g. Taylor and Tversky 1996), or how shifts in viewpoint affect text comprehension (e.g. Black *et al.* 1979). In general, this research suggests that specific linguistic devices serve as instructions for how an event should be simulated during language comprehension or how it is being simulated during production. The following is a brief sample of linguistic elements in English that have been claimed to reflect viewpoint.

- a. *Pronoun choice*: second person pronouns are more likely to induce readers to adopt a point of view inside the scene, while third person pronouns are more likely to induce readers to adopt an external point of view (Brunyé *et al.* 2009).
- b. *Frequency of mention of a particular referent in the discourse*: a frequently mentioned referent is likely to be the locus of viewpoint (Chafe 1976, 1994).
- c. *A referent serving as the origo for deictic elements*: *Cassandra came to Jane's house* as compared to *Cassandra went...* suggests that the locus of viewpoint is with Jane (Chafe 1994; MacWhinney 2005; Vandelanotte, this volume).
- d. *A referent appearing with certain classes of verbs*: verbs describing psychological states (*think, believe, feel*, etc.), verbs of perception (*see, hear*, etc.),

quotatives/markers of direct speech (*said, was like, went, was all*, etc.), and others (Chafe 1976, 1994; Kuno 1987; MacWhinney 2005; Dancygier, this volume).

- e. *Syntactic subject and word order*: a referent that appears frequently as the syntactic subject of utterances within the discourse is likely to be the locus of viewpoint (Chafe 1976, 1994; MacWhinney 1977, 2005). In languages like English, therefore, different syntactic structures (topicalization, clefts, passives) may function to permit flexibility in locus of viewpoint.
- f. *Modals, counterfactuals, definiteness, reciprocal verbs, reflexives . . .* (Kuno 1987; MacWhinney 2005; Ferrari and Sweetser, this volume).

Links between linguistic markers of viewpoint and conceptual viewpoint can be established using experimental paradigms as well. Black (1979) has shown that making an element the subject of a sentence establishes that entity as the locus of viewpoint, and that participants' memory for events within a text is better if point of view is consistent across sentences. Brunyé *et al.* (2009) have shown that if participants are presented with a sentence describing a task with the pronouns *I* or *you* (e.g. *I cut the apple*, as compared to *he cut the apple*) they are faster at a picture verification task when the picture shows the task from the point of view of an actor as opposed to an observer. Patterns such as these suggest that, during comprehension at least, there are links between the conceptual and the linguistic.

The picture becomes more complex when one considers data from gestures produced along with speech. The hands and body can be used in ways that encode point of view (McNeill 1992), or *gestural* viewpoint. Gestural viewpoint has received far less attention, despite the fact that it is equally prevalent, and equally valuable in attempting to understand conceptual viewpoint. It is also important to note that some uses of the hands and body that encode point of view are actually instances of *linguistic* and not gestural viewpoint, namely certain phenomena in signed languages. In signed languages, the hands, body, and face can encode point of view, sometimes according to grammatical rules (in which case, such behaviors should be considered linguistic), and sometimes in a way that is gradient and not rule-governed (gestural behaviors). Despite the fact that there are close parallels between some of these phenomena and phenomena in co-speech gesture (Quinto-Pozos 2007; Casey and Emmorey 2009), gestural examples in the sign stream will not be considered here (but see Janzen's and Shaffer's chapters in this volume).

As with linguistic viewpoint, it is common for gestural viewpoint to suggest that the conceptualizer is taking a perspective either outside the scene or inside it. Gestures in which the speaker uses her hands or body as though she is a character in the narrative (e.g. pumping the arms as though running while talking about someone running) have been called character viewpoint gestures (McNeill 1992). Gestures in which the speaker traces a character's path or indicates a location as though observing from a distance have been called

observer viewpoint gestures (McNeill 1992). This basic distinction between an internal and an external point of view can be found in conceptual accounts (e.g. MacWhinney's (2005) work on perspective taking; Tversky and colleagues' work on spatial descriptions: Taylor and Tversky 1996; Bryant and Tversky 1999; Emmorey *et al.* 2000), in linguistic accounts (e.g. Langacker 1991), and in research on gesture and signed languages (see Parrill 2009 for a comparison of different terms researchers have used in this arena). The internal/external distinction is thought to reflect a real difference in simulation – that is, simulation from the point of view of an actor or from the point of view of an observer. This difference in simulation would entail differences in mental imagery and motor programs (see Hostetter and Alibali 2008; Parrill 2010, for more detailed discussions of this issue) and might recruit different cortical circuits (Ruby and Decety 2001).

Viewpoint is clearly a complex construct, and most models of viewpoint do not even attempt to account for gestural data. In the next section, I describe a model of linguistic viewpoint that is somewhat distinct from that presented above. I show how this model might deal with gestural viewpoint, and suggest that the relationship between conceptual and linguistic viewpoint can be better understood by considering evidence from co-speech gestures.

4.3 Viewpoint in the blending framework

A notion of viewpoint closely related to that presented above can be found in the *blending* or *conceptual integration* framework (Fauconnier and Turner 2002). According to the blending framework, conceptualization recruits a series of partially structured mental models, or *mental spaces*. Viewpoint can be seen as residing in one of these mental spaces, referred to as the *Viewpoint Space*. According to Fauconnier (1997), the Viewpoint Space is “the space from which others are accessed and structured or set up” (p. 49). Certain linguistic structures are seen as *space builders*: they prompt language users to construct new mental spaces, and may also trigger a shift in which space is the Viewpoint Space. For example, in direct speech (e.g. *the man said, I ate a lobster*) the quotative *the man said* prompts the listener to shift viewpoint from the mental space of the speaker's reality (called the Base Space), to the mental space of the referent described by *the man*, which I will call the Story Space. Because information is accessed directly from the Story Space, *I* refers not to the speaker, but to the referent of *the man*. (See Sanders and Redeker 1996, for further discussion.)

Such explicit space builders draw attention to the fact that viewpoint can shift from one mental space to another across clauses. Many of the phenomena listed above can be captured with the mental spaces framework. However, viewpoint shifts can also be the result of more implicit linguistic prompts. Under the rubric of *implicit perspective*, Sanders and Redeker describe examples where the use of the indefinite article indicates that viewpoint has shifted across utterances.

To condense their example: in a sequence like *I lost my cat. A neighbor found a cat with a red collar*, the use of *a cat* requires a shift to the point of view of the neighbor, who does not know to whom the cat belongs.

These more implicit cases are of particular interest when attempting to unify the Viewpoint Space with the more general notion of viewpoint presented above. When one considers an isolated utterance, these two approaches to viewpoint are not equally clear in their claims. Within the mental spaces framework, the Viewpoint Space in an utterance like *the man ate a lobster* would be the same as the Base Space – that is, viewpoint would reside in the narrator's reality. According to the more general account, however, the syntactic subject indicates that viewpoint is with the man, which would entail that the Viewpoint Space is the Story Space. Cases of implicit perspective seem to indicate that viewpoint should not be assumed to be in the Base Space in the absence of more explicit space builders. But linguistic data alone leave this question open.

Gesture serves as an indicator of viewpoint as well, and gestural viewpoint can reveal differences in *conceptual* viewpoint that are not apparent from speech – that is, gesture can indicate that the Viewpoint Space is the Base Space or an embedded space (a Story Space), even when speech is identical.

4.4 Gestural viewpoint in the blending framework

The meaning of a co-speech gesture is largely dependent on context provided by speech (McNeill 1992, 2005). For this reason, understanding viewpoint in gesture requires two levels of analysis: what is encoded in speech, and how speech shapes the interpretation of what is encoded in gesture. In addition, when gesturing, the hands typically represent something else. As a number of researchers have pointed out (Liddell 1998, 2003; Dudis 2004; Parrill and Sweetser 2004), this additional level of complexity can be fruitfully understood in terms of *Real Space* blends. Real Space blends involve a particular mental space, the Real Space. The Real Space is a conceptual model of a language user's immediate physical environment, including a representation of her own body (Liddell 1998). Real Space, in conjunction with other mental spaces, serves as an input to a blend. Blending is a general cognitive process that allows humans to take partial structure from various mental spaces and create new mental structures. A blend allows the Real Space articulators (hands, body) to map onto different entities.

4.4.1 Narrator Viewpoint

When considering viewpoint in gesture, there are three basic possibilities for Real Space blends. First, the Real Space hand or body may map onto the narrator's hand or body. In Figure 4.1, the narrator produces a so-called *presenting*

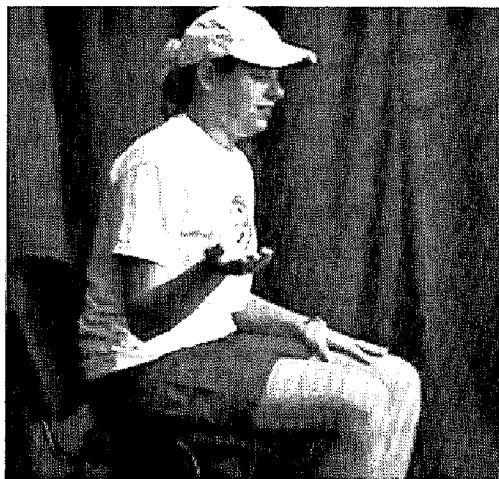


Figure 4.1 Narrator Viewpoint gesture (speech: *then he goes up a flagpole or something*)

gesture, while describing the actions of a character in a story. This is a common metanarrative gesture that has a metaphoric interpretation: the speaker is presenting an object to the listener for inspection. That object is interpreted as a segment of discourse. The network of mappings that underlies this example is described in Parrill and Sweetser (2004). For current purposes, however, the important feature is that the Real Space hand of the narrator maps onto the narrator's own hand in the blend. In this case, we would say that the gestural viewpoint is with the narrator. Such gestures will be referred to as Narrator Viewpoint gestures. According to the mental spaces framework, the Viewpoint Space is the Base Space, which models the narrator's reality. The more general account of viewpoint would claim that the entity described by *he* is the locus of viewpoint. Gesture can resolve this inconsistency. The gestural data suggest that conceptual viewpoint really is with the narrator: the narrator's actions are interpreted with respect to the narrator's Base Space, not the Story Space. The blending framework captures this difference by having the connector between an element in the Base Space and its counterpart in the Story Space be an identity relation – that is, the narrator's gesture maps onto the narrator's own action of offering a metaphorical object. Narrator Viewpoint gestures can thus be understood as gestures in which there is an identity relation between a Real Space action and the narrator's own action.

4.4.2 *Observer Viewpoint*

The second possibility is for the Real Space articulator to map onto an entity in the Story Space (the reality of the story world). McNeill (1992) has called such gestures Observer Viewpoint (O-VPT) gestures. In Figure 4.2, the narrator's hand no longer maps onto the narrator's hand, but onto a character as a whole.



Figure 4.2 Observer Viewpoint gesture (speech: *then he goes up the flagpole*)

The motion of the hand maps onto the character's trajectory. Gestural point of view is no longer in the Base Space because the hand no longer represents the narrator's hand. The interpretation of the action performed by the hand relies on the Story Space. However, the narrator's viewpoint is still implicitly present. This can be seen from the fact that she is taking a particular visual perspective on the scene. Such cases are not unlike linguistic examples of indirect speech (e.g. *he said that he ate a lobster*): the narrator is responsible for the phrasing of the utterance, but the character (*he*) is responsible for the content (see Sanders and Redeker 1996). The blending framework captures this difference by having the connector between the utterance in the Base Space and its counterpart in the Story Space be a consequence relation, rather than an identity relation. Unlike an identity relation, which involves equivalency between two things (e.g. a description of an utterance and an actual utterance, a gesture and an actual event), a consequence relation involves some kind of re-representation, but no direct equivalence. The same can be said of the Real Space action and its counterpart in the Story Space. The presentation of the trajectory is the responsibility of the narrator, but the content is the responsibility of the character. Thus, Observer Viewpoint gestures can be described as gestures in which there is a consequence relation between a Real Space action and an action in a Story Space.

4.4.3 Character Viewpoint

The third possibility is for a Real Space articulator to represent a character's own articulator. In Figure 4.3, the speaker uses her own body to depict actions

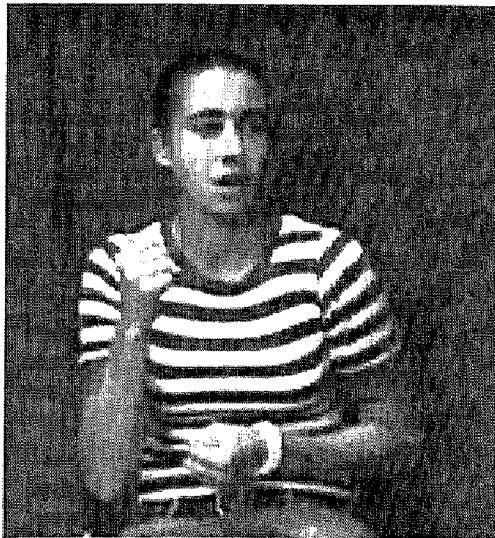


Figure 4.3 Character Viewpoint gesture (speech: *then he goes up the flagpole*)

of the character's body. Her hands map onto the character's hands and her torso onto the character's torso. McNeill (1992) has called such gestures Character Viewpoint gestures. Again, gestural point of view shifts from the narrator's space to a projected entity in the Story Space. In this case, the gestural point of view is the character's, rather than the narrator as an observer within the Story Space. Such cases are not unlike linguistic examples of direct speech (e.g. *he said, I ate a lobster*): the character rather than the narrator is responsible for both the phrasing and the content of the utterance (see Sanders and Redeker 1996). The blending framework captures this difference by having the connector between an element in the Base Space and its counterpart in the Story Space be an identity relation, rather than a consequence relation. Thus, Character Viewpoint gestures can be described as gestures in which there is an identity relation between Real Space action and an action in the Story Space.

4.4.4 *Interim summary*

This discussion has glossed over a great many details, including the fact that narrators occasionally combine Character and Observer Viewpoint gestures simultaneously, creating Dual Viewpoint gestures (McNeill 1992; Parrill 2009). In the previous sections, I have attempted to show that gesture can indicate the locus of conceptual viewpoint. In the next section, I move from a general discussion of the interactions between conceptual, linguistic, and gestural viewpoint to the presentation of an experimental study. This study suggests that a conceptual variable known to have effects on linguistic viewpoint – namely, whether

information is shared between speaker and addressee – can also have effects on gestural viewpoint. This is precisely what we might expect if gestural and linguistic viewpoint have a common conceptual source.

4.5 Gesture complexity and discourse status

Many of the linguistic devices traditionally analyzed as viewpoint phenomena are sensitive to discourse status. The label *discourse status* can refer to a number of complex phenomena: in this chapter, the term refers to assumptions speakers make about what information is available and currently accessible to an interlocutor. These assumptions determine how entities in a discourse are linguistically encoded and how referring expressions are interpreted by addressees (Clark and Haviland 1977; Clark and Marshall 1981; Prince 1981; Clark *et al.* 1983; Birner 1994; Birner and Ward 1998; Clark and Krych 2004). For the current study, one general pattern is of particular importance. When information is shared, less linguistic encoding is typically used in referring to an entity (Givón 1979). For example, the first mention of an entity might involve a complex noun phrase with the indefinite article (*a rabbit wearing a red and grey baseball uniform*). After several mentions, a simple pronoun (*it*) might be used to refer to that entity.

This pattern can be observed in speech-accompanying gestures as well (Levy and McNeill 1992). The fact that gesture is also sensitive to discourse status has been demonstrated experimentally in both narrative (Parrill 2011) and non-narrative language (Gerwing and Bavelas 2004; Holler and Stevens 2007). These experimental studies have analyzed features such as gesture complexity, precision, and informativeness, but have not considered viewpoint. David McNeill, however, has hypothesized that character viewpoint should be used when information is maximally salient or newsworthy (McNeill 1992) – that is, Character Viewpoint gestures should correlate with moments in the discourse where complex linguistic structures are necessary (e.g. the complex noun phrase above). In an analysis of one narration, McNeill shows that gesture complexity increases with linguistic complexity, with character viewpoint gestures occurring only with complex forms such as clauses and verb phrases (McNeill 1992). The current study extends this work in three ways. First, an experimental rather than an observational approach is taken. Second, this study focuses specifically on shared knowledge, rather than looking at newsworthiness or salience more generally. Third, the study focuses on a specific event within a narrative, for which both character and observer viewpoint can plausibly be produced. This control is essential, as there are some events for which viewpoint does not vary, regardless of context (Parrill 2010).

4.6 An empirical assessment of gestural viewpoint

4.6.1 Participants and method

Twenty-four University of Chicago students participated in the study for payment. All were native speakers of American English. Participants came to the experiment with a friend, who served as a listener. Each participant watched three cartoon stimuli and described them to his or her friend. Two were practice stimuli and will not be discussed further. The third (experimental) stimulus was a cartoon that a pilot study had shown would evoke both Character and Observer Viewpoint gestures. Participants watched the stimuli in one of two conditions: the *shared knowledge* condition and the *control* condition. In the shared knowledge condition, participants watched the stimuli with their friend. In the control condition, participants watched the stimuli alone. There were twelve participants in each condition.

If the addressee is present while the narrator watches a cartoon (as is the case in the shared knowledge condition), the narrator should be aware that the content of the cartoon is shared. If the production of Character Viewpoint gestures is impacted by the discourse status of the information a narrator is communicating, fewer Character Viewpoint gestures are expected in the shared knowledge condition.

4.6.2 Data and analysis

The speech each participant produced when describing the experimental stimulus was transcribed. The accompanying gestures were also transcribed and coded by two independent coders, who were blind to the experimental condition and hypotheses. Gestures that did not relate to the cartoon content (e.g. metaphoric gestures produced to regulate the discourse) were excluded from this analysis. Coders sorted cartoon-relevant gestures into those in Character Viewpoint and those in Observer Viewpoint. Coders agreed on the viewpoint of 89 per cent of the gestures. Disagreements were resolved through discussion.

4.6.3 Results

We first compared the mean number of gestures each participant produced. As shown in Figure 4.4, participants in the control condition produced more gestures than did participants in the shared knowledge condition (Mann-Whitney $U = 2, p < .0001$). This finding is in line with the general pattern discussed above: linguistic encoding tends to decrease when information is shared. As linguistic encoding decreases, so does gestural encoding. Because narrators in

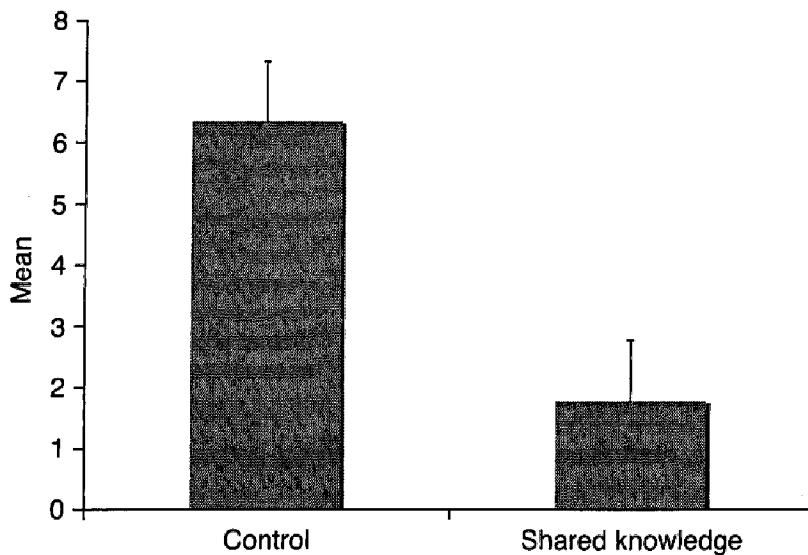


Figure 4.4 Mean number of gestures produced

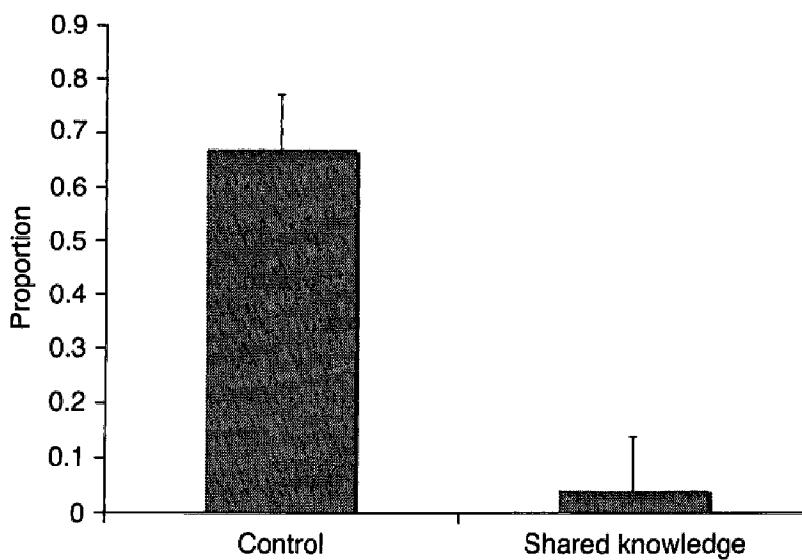


Figure 4.5 Mean proportion of C-VPT gestures produced

the control condition are gesturing more frequently overall, the *proportion* of each narrator's gestures that were in Character Viewpoint was used for a comparison across conditions. The mean proportion of character viewpoint gestures produced by participants in the control condition was also significantly higher ($Mann-Whitney U = 1, p < .0001$). This pattern is shown in Figure 4.5.

4.6.4 Discussion

As predicted, narrators who were aware that their listeners shared knowledge about the content of their narration produced fewer Character Viewpoint gestures. This finding supports the claim that gesture complexity in general, and Character Viewpoint as a special case of maximal gestural complexity, changes as a function of discourse status.

The experiment presented here is a rather unsubtle way of manipulating discourse status. As a result, there are other possible interpretations for the pattern observed. For example, the control condition is a relatively normal situation, whereas the shared knowledge condition is less naturalistic. It may be that motivation was reduced or inhibition was increased in the latter condition. However, the important issue is whether there is an alternative explanation that accounts not just for the lower gesture rate in the shared knowledge condition (which might well be explained by decreased motivation), but also for the difference in the type of gesture produced – that is, there is no explanation for why Character Viewpoint gestures would be affected by motivation while Observer Viewpoint gestures are not.

4.7 Conclusions

This chapter has endeavored to accomplish two things. First, I have tried to provide a general picture of the world of viewpoint that takes into account viewpoint in both speech and speech-accompanying gestures. I have suggested that viewpoint is properly considered as three phenomena: conceptual, linguistic, and gestural. Conceptual viewpoint (how a language user simulates an event or scene) is largely hidden, accessible only via linguistic and gestural viewpoint (or behavioral or brain-imaging measures). In considering how these two indices of conceptual viewpoint interact, I have shown that differences among gestural manifestations of viewpoint can be profitably understood using the mental spaces framework. I have also argued that gestural viewpoint can reveal shifts in the locus of viewpoint that are not apparent from speech alone. Another possibility, of course, is that the two systems are distinct, and that gestural viewpoint does not tell us anything about the system governing linguistic viewpoint. In one sense, the two systems obviously *are* distinct. As McNeill (1992, 2005) has pointed out, language is sequential, symbolic, and analytic, while gesture is holistic and imagistic. It is therefore not unthinkable that a conceptualizer might have one linguistic perspective and another gestural perspective. But to assume that gestural viewpoint tells us nothing about linguistic viewpoint seems unwise for two reasons. First, a growing body of research indicates that language relies heavily on imagery and motor systems, the same systems that give rise to co-speech gestures. Second, gestural and linguistic

viewpoint appear to be impacted by some of the same variables. That is, the second goal of this chapter (in addition to providing an overview of viewpoint) was to provide some empirical support for the claim that gestural viewpoint and linguistic viewpoint share conceptual sources. This study has suggested that discourse status, a variable known to be central for selection of linguistic viewpoint, also impacts gestural viewpoint. While this study was not designed to show that the two factors were simultaneously influenced, work by Gerwing and Bavelas (2004) does show such yoked effects. Thus, gestural viewpoint should be considered when making claims about linguistic viewpoint, as both ultimately reflect conceptual viewpoint.

References

- Bergen, Benjamin, Shane Lindsay, Teenie Matlock, and Srinivari Narayanan. 2007. Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science* 31:5, 733–64.
- Birner, Betty J. 1994. Information status and word order: An analysis of English inversion. *Language* 70:2, 233–59.
- Birner, Betty J. and Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins.
- Black, John B., Terrence J. Turner, and Gordon H. Bower. 1979. Point of view in narrative comprehension, memory, and production. *Journal of Verbal Learning and Verbal Behavior* 18:2, 187–98.
- Borghi, Anna M., Arthur M. Glenberg, and Michael P. Kaschak. 2004. Putting words in perspective. *Memory and Cognition* 32:6, 863–73.
- Brunyé, Tad T., Tali Ditman, Caroline R. Mahoney, Jason S. Augustyn, and Holly A. Taylor. 2009. When you and I share perspectives. *Psychological Science* 20:1, 27–32.
- Bryant, David J. and Barbara Tversky. 1999. Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25:1, 137–56.
- Casey, Shannon and Karen Emmorey. 2009. Co-speech gesture in bimodal bilinguals. *Language and Cognitive Processes* 24:2, 290–312.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li and Sandra A. Thompson (eds.). *Subject and Topic*. New York: Academic Press, 27–55.
1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.
- Clark, Herbert H. and Susan E. Haviland. 1977. Comprehension and the given-new contract. In Roy O Freedle (ed.). *Discourse Production and Comprehension*. Norwood, NJ: Ablex, 1–40.
- Clark, Herbert H. and Meredyth A. Krych. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50, 62–81.
- Clark, Herbert H. and Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber, and Ivan Sag. *Elements of Discourse Understanding*. Cambridge University Press, 10–63.

- Clark, Herbert H., Robert Schreuder, and Samuel Buttrick. 1983. Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 22, 245–58.
- DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57:3, 626–57.
- Dudis, Paul. 2004. Body-partitioning and real-space blends. *Cognitive Linguistics* 15:4, 223–38.
- Emmorey, Karen, Barbara Tversky, and Holly A. Taylor. 2000. Using space to describe space: perspective in speech, sign, and gesture. *Spatial Cognition and Computation* 26, 157–80.
- Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Finke, Ronald A. 1989. *Principles of Mental Imagery*, Cambridge, MA: MIT Press.
- Gerwing, Jennifer and Janet Bavelas. 2004. Linguistic influences on gesture's form. *Gesture* 4:2, 157–95.
- Gibbs, Raymond. 2005. *Embodiment and Cognitive Science*. Cambridge University Press.
- Givón, Talmy. 1979. From discourse to syntax: grammar as a processing strategy. In Talmy Givón (ed.). *Syntax and Semantics 12: Discourse and Semantics*. New York: Academic Press, 81–114.
- Glenberg, Arthur M. and Michael P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 96, 558–65.
- Holler, Judith and Rachel Stevens. 2007. The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology* 26:1, 4–27.
- Hostetter, Autumn B. and Martha W. Alibali. 2008. Visible embodiment: gesture as simulated action. *Psychonomic Bulletin & Review* 15:3, 495–514.
- Kaschak, Michael P. and Arthur M. Glenberg. 2000. Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language* 43:3, 508–29.
- Kosslyn, Stephen. 1994. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy*. University of Chicago Press.
- Langacker, Ronald W. 1991. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.
- Levy, Elena T. and David McNeill. 1992. Speech, gesture, and discourse. *Discourse Processes* 155, 277–301.
- Liddell, Scott. 1998. Grounded blends, gestures and conceptual shifts. *Cognitive Linguistics* 9:3, 283–314.
2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.
- MacWhinney, Brian. 1977. Starting points. *Language* 53, 152–87.
2005. The emergence of grammar from perspective taking. In Diane Pecher and Rolf Zwaan (eds.). *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge University Press, 198–223.

- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal About Thought*. University of Chicago Press.
2005. *Gesture and Thought*. University of Chicago Press.
- Parrill, Fey. 2009. Dual viewpoint gestures. *Gesture* 9:3, 271–89.
2010. Viewpoint in speech-gesture integration: linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25:5, 650–68.
2011. The hands are part of the package: gesture, common ground, and information packaging. In Sally Rice and John Newman (eds.), *Empirical and Experimental Methods in Cognitive/Functional Research*. Stanford: Center for the Study of Language and Information, 285–302.
- Parrill, Fey and Eve Sweetser. 2004. What we mean by meaning: conceptual integration in gesture analysis and transcription. *Gesture* 4:2, 197–219.
- Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical Pragmatics*. New York: Academic Press, 223–55.
- Quinto-Pozos, David. 2007. Can constructed action be considered obligatory? *Lingua* 117, 1285–314.
- Richardson, Daniel C., Michael J. Spivey, Lawrence W. Barsalou, and Ken McRae. 2003. Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27:6, 767–80.
- Ruby, Perrine and Jean Decety. 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neuroscience* 4:5, 546–50.
- Sanders, José and Gisela Redeker. 1996. Perspective choice and the representation of speech and thought in narrative discourse. In Gilles Fauconnier and Eve Sweetser (eds.), *Spaces, Worlds, and Grammar*. University of Chicago Press, 290–317.
- Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and Julie C. Sedivey. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–4.
- Taylor, Holly A. and Barbara Tversky. 1996. Perspective in spatial descriptions. *Journal of Memory and Language* 35:3, 371–91.
- Zwaan, Rolf A., Robert A. Stanfield, and Richard H. Yaxley. 2002. Language comprehenders mentally represent the shapes of objects. *Psychological Science* 136, 168–71.

5 *Maybe what it means is he actually got the spot*

Physical and cognitive viewpoint in a gesture study

Shweta Narayan

5.1 Introduction

Participants in a gesture study create meaning interactively. In the example analyzed in this chapter, one participant (Speaker), who can see an image, fails to understand it. She comes up with successively less wrong interpretations of the scene, but it is the other participant (Listener), who cannot see the image, who finally figures out the standard canonical interpretation of the image. In the course of the interaction, Listener's cognitive viewpoint shifts, as evidenced by his speech and physically viewpointed gestures. His viewpoint goes through four phases:

- 1) Taking Speaker's viewpoint: this is only clear in the *interaction* of speech and gesture.

LISTENER: ... to the left...? [*gestures to his right (Speaker's left)*]

- 2) Misunderstanding Viewpoint Space: Speaker's viewpoint is of someone looking at the picture. Listener thinks it is of characters within the depicted scene.

SPEAKER: On the right-hand side of the street, there are five cars...

LISTENER: Hold on, you can't say the right side of the street, because whatever side of the street you're on is gonna be the right side.

- 3) Switching visual viewpoint: Listener traces the street from Speaker's viewpoint; then pauses and repeats, from his own viewpoint.

- 4) Mismatch: his gesture contradicts his speech, and Speaker's earlier gesture.

SPEAKER: He's pulling out of a spot [*gestures car backing out*]

LISTENER: He's trying to pull out of the spot? [*gestures car moving forward*]

Listener, then, begins by aligning himself to Speaker's physical viewpoint, but through a series of viewpoint conflicts, he shifts away from this alignment. After the switch, he starts asking for elaboration instead of clarification, suggesting that he is actively building a space corresponding to the picture. Only after this does he disagree with Speaker, presenting the correct framing of the image:

“... maybe what it means is he actually got the spot.” His shifts in cognitive viewpoint are indexed by changing visual viewpoint.

Gesture is inherently viewpointed, since it is an embodied action that occurs in space, and any body in space has a richly structured frame of reference. It is not possible to gesture iconically without taking a viewpoint on the gestured scenario; often the viewpoint is that of the gesturer’s body in some blend. Because of this, gesture provides insight into interlocutors’ cognitive viewpoint that speech alone cannot. In the data examined here, the interaction of speech and gesture reveals the constructed, imagined relationship of participants’ bodies to a picture, and thereby also profiles the frames of reference that must have been preserved in the ongoing discourse blend for these relationships to exist.

Recent work on gesture and signed languages has shown that both can be profitably understood in terms of Mental Spaces Theory (Fauconnier and Sweetser 1996; Fauconnier 1997; Fauconnier and Turner 2002), and in particular in terms of Real Space blends (Liddell 2003; Dudis 2004; Parrill and Sweetser 2004). When a communicator makes a “grasping” hand motion in the air, for example, and is understood as depicting some character grasping an object, the construal of the gesture is a blend between the communicator’s motion in Real Space (her, or her interlocutor’s, understanding of the physical space around her) and the imagined character’s motion in the described Event Space. As we shall see, interpreting pictures is actually more complex than the gesture just described, but human ability to create and develop Real Space blends is basic to our ability to make and understand meaningful gestures.

5.2 The experiment

5.2.1 *The communicative situation*

In the study that this chapter’s gesture data are taken from,¹ two subjects sat facing each other in a room (see Figure 5.1). One (Speaker) could see a projector screen, and the other (Listener) could not. Nine different comics panels (including the one under discussion) were presented on the projector screen, one at a time. The Speaker’s task was to describe each panel so that the Listener could understand it. The Listener was encouraged to ask questions if anything was unclear. After each panel was described to both subjects’ satisfaction, the Speaker pressed a key² to move on to the next image. Images were presented in random order.

As seen in Figure 5.2, subjects were mostly facing each other, though Speaker could swivel to look at the projector screen (direction indicated by the arrow).

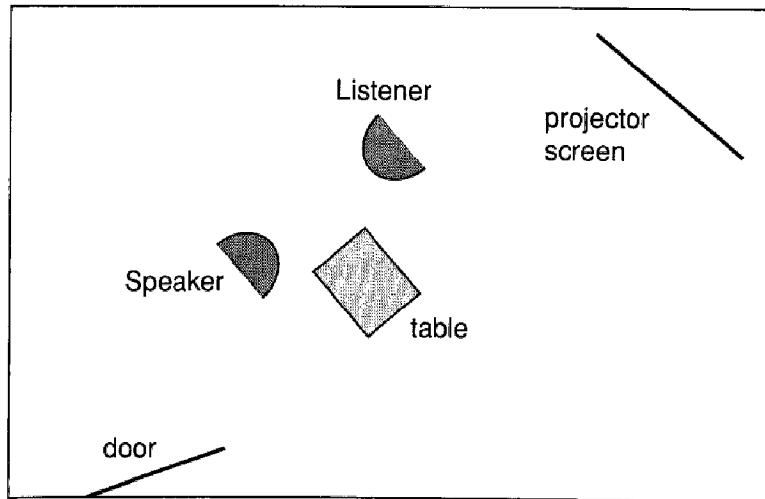


Figure 5.1 The room

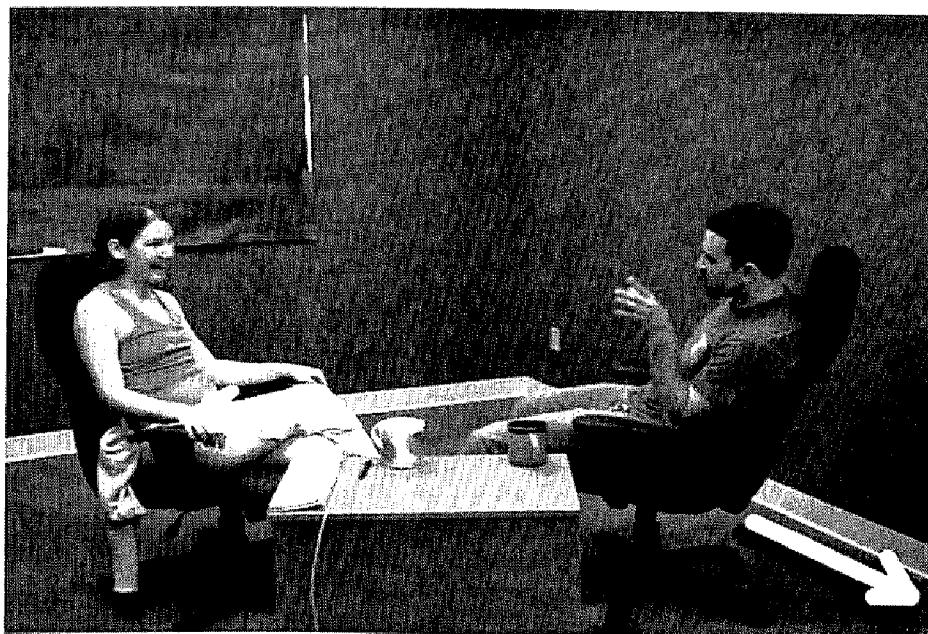


Figure 5.2 The communicative situation

5.2.2 *The stimulus*

Given the task, Speakers could potentially have done most of the talking, with Listeners giving a more limited contribution, perhaps only a few clarification questions. However, this was often not what happened; the process of meaning construction was more collaborative, despite the asymmetry in information availability. This is especially evident in the data from the session analyzed



Figure 5.3 Stimulus image (from Patrick Farley [2003], *Barracuda: The Scotty Zaccharine Story* [www.electric sheepcomix.com](http://www.electricsheepcomix.com))

here, where the Speaker became extremely confused about the image shown in Figure 5.3.

In this image, a yellow-green VW New Beetle has just made an illegal U-turn against oncoming traffic on a city street in order to pull into a parking space. To get from the static image to this mentally constructed dynamic scenario, readers need to do some fairly sophisticated cognitive processing. They need to *partition* (Dudis 2004) their understanding of the physical space of the comic into several mental spaces (Fauconnier and Sweetser 1996; Fauconnier 1997; Fauconnier and Turner 2002). Notice that the various words and image pieces are not depicting simultaneous events: the U-turn indicated by the screech and the vapor trail precedes (and partially overlaps with) the honk and hey events, and score presumably is the driver's final mental comment as the parking process is complete. So the image represents a succession of temporally and causally related spaces, which the viewer is invited to construct by relating the frames prompted in the scene. Making the "correct" mappings, to reach the canonical interpretation, requires frame-based knowledge (Fillmore 1985) of real-world situations, image-schematic (Hampe and Grady 2005) and force-dynamic (Talmy 2000) structure, and knowledge of comics conventions such as

vapor trails and thought bubbles. These provide the generic structure. Viewers thus at least need an image space, plus a space for each linguistic item (in this case, each of the single words “screech,” “honk,” “hey,” and “score,” which are not interpreted as visual parts of the constructed scene, but as separate auditory stimuli), and a “vapor trail” space indicating motion. Using their frame-based knowledge (annoyed drivers honk or shout when someone makes a U-turn in front of them; seekers of scarce parking spaces may exult when they finally find one), viewers can then integrate these mental spaces to create a dynamic, noisy scene from the static, silent stimulus.

Without this work, readers would draw some rather implausible inferences. For example, they might infer that there are bright yellow letters floating in the air, obscuring parts of the road and a car, or a curved area of tarmac that happens to be lighter than the rest of the street.

Speakers in this study had no trouble inferring that the letters spelling out *Honk* represent a sound, and that the sound was made not by the green car, but by the white car behind it. In addition, almost all Speakers understood from the yellow letters that the green car made a screeching sound, that the “vapor trail” indicated the path that the car followed, and that the co-location of the yellow letters and the path of the vapor trail indicated that the car screeched as it traversed this path. Speakers could generally infer from the combination of trail and letters that the motion was fast. Theoretically, they might also infer from the placement of letters *where* along the path the screeching sound occurred, and conclude that when the car gets to where it is drawn, it is no longer screeching; however, none of them mentioned that specifically.

5.2.3 *The Speaker's confusion*

In the clip analyzed here, the Speaker has a great deal more trouble than most subjects. In twenty-four subject pairs, she is the only one who had this much trouble; she seems to have missed the motion trail – she never mentions it. Possibly as a result, she never links the SCREECH with the green Beetle at all.

Her initial understanding of the picture is both very far from the interpretation I proposed, and extremely unsatisfactory to herself. She changes and modifies her interpretation several times during the course of the discourse. Initially, she tries to structure the image according to a video game frame.

(1) [0:01] Speaker

The biggest thing that stands out is this huge thing that says SCORE!

Oh, maybe it came from a video game.

Looks like it coulda come directly from a video game.

She never mentions a comics frame, but does implicitly evoke one and restructure the scene accordingly.

(2) [0:55] Speaker

There's honk, screech
and some guy saying
hey
or the, a bubble
that says hey coming out of a car

But as soon as she has done this, she mistakenly evokes an accident frame, and only a little later does she realize that the image is of a street scene but not of an accident.

(3) [1:06] Speaker

y- there's no visual . . . thing, of
honk . . . or of a, of a accident . . .
but like, the screech and the word honk make you think that there
may have been an accident.

(4) [1:16] Speaker

Oh, I g- okay, I'm understanding what's happening here.
Nervous laughter
So, it's a, it's a street scene . . .

She then gets confused about the car's direction of motion, concluding that it is pulling *out* of the parking space.

(5) [3:19] Speaker

RH hold diagonally in front of face
[One guy . . .][is actually
[*LH B, palm facing R, fingers touching R arm mid-forearm*]
[*LH moves down away from R arm*[
[Either[
[*LH moves down slightly more*

He must be pulling out of a spot.

LH returns to R forearm, then pulls downward away from RH

LH returns to R forearm

He's pulling out of a spot. And it is a . . .

LH pulls downward away from RH

Speaker's right forearm represents the upper edge of the road; it is held at an angle that mirrors the road's angle from a viewer's perspective, and its length is mapped to the visible segment of road. Her left hand represents the green

Beetle; it maps iconically to the car's position and orientation relative to the road, and then, when she moves it, to the car's inferred motion.

Because of Speaker's continued and changing confusion, this subject pair provides a rather spectacular example of collaborative construction of meaning. The comics artist provides cues to meaning construction that Speaker does not pick up on; but as we shall see, Listener, who cannot see those cues, nevertheless picks up on them from Speaker's description, and from the mental model he constructs. His comments feed back into Speaker's meaning construction, thereby allowing them both to reach an understanding of the image that is close to the canonical interpretation I have proposed.

5.3 Speaker and Listener viewpoints

So how does Listener break away from the interpretation that Speaker has constructed? How does he move from nodding and asking for clarification to proposing an alternative solution? This shift in interaction seems to be tied to his shifting cognitive viewpoint, which goes through four sequential phases, indexed by his viewpointed speech and gesture.

In the first phase, he takes Speaker's viewpoint, by moving his hand to his right (her left), while saying "left" (much as a dance teacher might when demonstrating for students). In the second, he misunderstands her Viewpoint Space – she is describing the picture as a visual object seen from her physical viewpoint, while he thinks she is describing the depicted scene from the imagined viewpoint of someone within that scene. In Mental Spaces terms, he assumes that her Viewpoint Space is the nested space of a driver-participant in the depicted traffic scene, when her Viewpoint Space is actually her own Base Space – that of a person looking at a picture. So when she says "the right hand side of the street," meaning the part of the street further to the right of the depicted image as she sees it, he objects that "whatever side of the street you're on, that's gonna be the right side."

In the third, he stops taking Speaker's viewpoint, and changes the direction of his gestures so that they would be coherent if *he* were looking at the picture (though the only picture he's "looking at" is imagined). In Mental Spaces terms, his speech and gesture indicate that his Viewpoint Space is that of his (reconstructed) understanding of the depicted scene, a blend in which some of the physical space to the front of his actual body is construed as being part of the imagined scene, such that the concrete gestures centered on his own body correspond to structure in his imagined space. Such construals of actual space are known as Real Space Blends (Real Space is a speaker's conceptual model of the physical space around her [Liddell 2003; Dudis 2004]; recall that in Mental Spaces Theory, speakers/thinkers never have objective knowledge, but only their own perceptual and cognitive models, at their disposal).

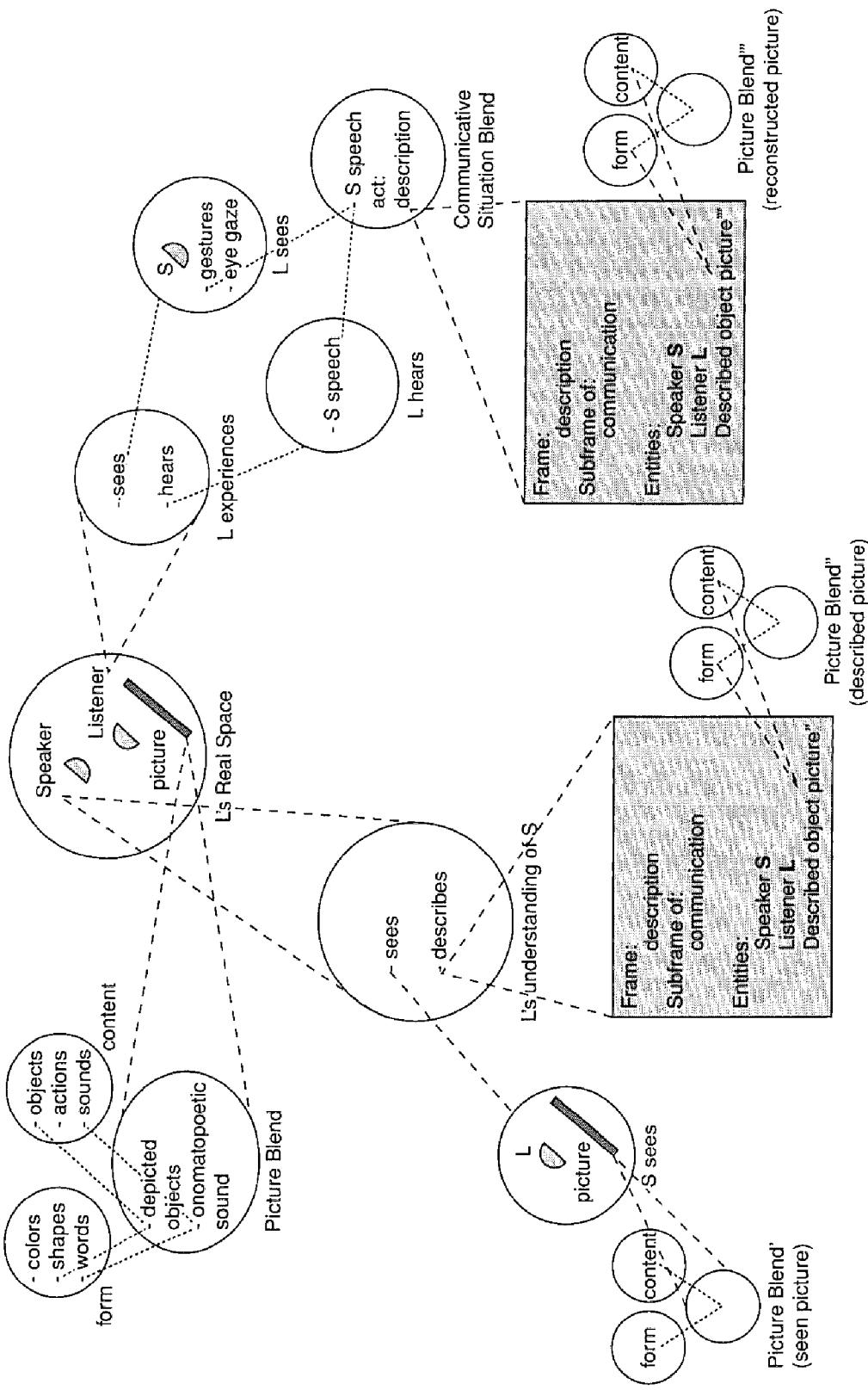


Figure 5.4 Listener's accessible space network

In the fourth phase, he shows a gesture mismatch with Speaker's gesture, repeating her words but with a gesture that moves in the opposite direction from hers. Only after this does he disagree explicitly with Speaker, presenting the correct framing of the image.

Listener's shifts in cognitive viewpoint are indexed by changing visual viewpoint, which are detailed in the sections below. To properly understand them, however, we must spend a little time on his constructed understanding of the communicative situation, before we move on to his understanding of the depicted scene.

This is more complex than it may seem at first, as can be seen by Figure 5.4, a minimal representation of Mental Spaces accessible to Listener and informing his understanding of the discourse. Listener needs to keep track of both his Real Space (*L's Real Space* in Figure 5.4, including Speaker and a screen) and his understanding of Speaker's experience and actions (*L's understanding of S* in Figure 5.4 – what is he interpreting her as describing, and reconstructing her as seeing?). He must also be tracking what he experiences (*L experiences*), which includes the nested spaces of what he sees and what he hears, some parts of which (saliently Speaker's speech and gestures) are blended into a *Communicative Situation Blend*.

In Listener's Real Space is also the visually inaccessible but saliently present projector screen behind him, which shows a comics image. He knows, then, that there is a picture on the screen with formal and semantic elements (expanded into his *Picture Blend* in Figure 5.4). The dotted lines between the spaces indicate that the *Picture Blend* Space is an elaboration of the Real Space, a partitioned part of it, rather than a nested space. The dots connect to the Blended Space rather than the Form and Content Spaces because the blend is primary. This blend is generic – Listener does not see the picture and cannot form a full understanding of the blend; he only knows it is there and possibly uses some of the generic blending conventions he has for comics images once he knows that the image is part of a comic.

In addition, he must be aware of what Speaker sees – especially that she has (presumably veridical) experience of that picture. This is not the same blend as the picture itself; it is a part of Speaker's experience as Listener constructs it, shown in Figure 5.4 as *Picture Blend'* (*seen picture*). Her task is to describe this blend, thus structuring her experience with a description frame. The act of describing creates another blend – *Picture Blend''* (*described picture*). Note that this is all Listener's understanding of Speaker's experience and action – her own mental space network would be entirely separate, in *her* head rather than his, and presumably would contain analogs of many of these spaces.

Because the task is to listen to her description, Listener's *Communicative Situation Blend* must be structured by a description frame. This description lets him create an elaborated understanding of the stimulus, shown

in Figure 5.4 as the nested *Picture Blend''' (reconstructed picture)*. This is a different instantiation of the description frame from the one which structures his understanding of *her* experience, because this one structures *his* Communicative Situation Blend. They do share identity mappings, however, since the situation is intersubjective, and the description she gives is the same one he hears.

In these terms, Listener's task is to bring *Picture Blend'''* into alignment with the actual picture, via his experience of *Picture Blend''*, and the assumption that image-schematic structure is preserved from the picture to Speaker's seen picture, to *Picture Blend'*. (In order to keep the diagram possibly comprehensible, I have not drawn in any of these identity mappings.) As we shall see, in order to do so, he starts off with speech and gesture that suggest his cognitive viewpoint is *S Sees*, and shifts his viewpoint to *Picture Blend''' (reconstructed picture)* over the course of the interaction. This shift, this pulling away from alignment with Speaker's viewpoint, is what lets him make the cognitive leap that she does not make.

5.3.1 Listener takes Speaker's viewpoint

Later in the exchange, Listener makes the following comments:

- (6) [1:50] **Listener**
so the larger
B hands, palms facing each other, go to top right corner
the larger angle of the road is
Both-hand beat
on – going off – to the . . .
B hands move further up and right
left-hand side?
B hands stay up and right
- (7) [1:56] **Listener**
And then, kinda shrinks down as it goes to the right-hand side?
B hands trace a line down and left, ending in bottom left corner

These examples show an apparent mismatch between Listener's speech and gesture. He says "right" while gesturing to his left, and vice versa. Listener is not, however, confused. He is facing Speaker, and when he gestures to his right, he is gesturing to her left; he does this, and then hesitates before saying "left," which indicates some difficulty rather than a mistake. It is worth noting that he is mirroring Speaker's earlier gestures at this point; the hand shape and movement of his gesture reflect her previous one, as shown in Figure 5.5.

I have called this "taking Speaker's viewpoint." Obviously, however, we do not have direct access to other people's mental spaces, and Listener's Viewpoint

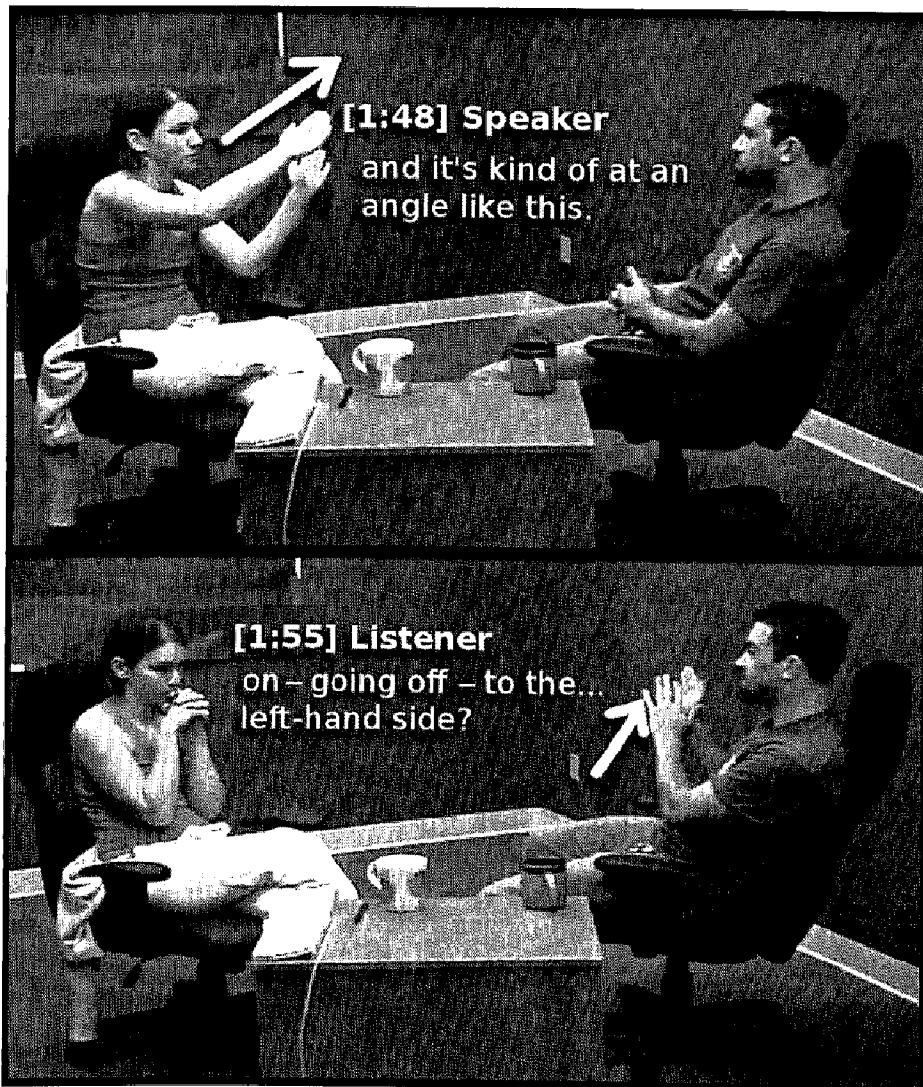


Figure 5.5 Mirroring gestures

Space is not actually that of Speaker's visual experience. It is Listener's partial, minimal reconstruction of that visual experience. This is the space labeled *S Sees* in Figure 5.4.

In terms of the discourse, what this means is that the spatial frame of reference of *S Sees*, and its image-schematic structure, are preserved and taken as canonical in the blend that Listener is currently running. This blend takes two reified, standard blends as input.

5.3.1.1 Reified blend 1: shifting to an interlocutor's frame of reference

We shift frames of reference all the time, by integrating our Base Space with a space created by some image-schema transformation (Lakoff 1987). One very simple, reified version of this happens in face-to-face conversation; we

Table 5.1 *Reified blend 1*

Input 1: A's Frame of Reference	Input 2: B's Frame of Reference	Generic image structure of: Bodies Frames of Reference	Blend: A shifts to B's Frame of Reference	
A's left side	B's left side	Right-left image schema	"Your left"	
A's right side	B's right side			"Your right"
A's head	B's head	Up-down image schema	Head	Up
A's feet	B's feet		Feet	Down

blend our Real Space with a 180-degree mental rotation (Kosslyn 1980) of Real Space, to create the blend that is our understanding of our interlocutor's Real Space. (*S Sees* in Figure 5.4 is just such a blend.) The process of rotation often takes some work, and may be the reason for Listener's disfluency in the segment at 1:50, before he says "left." But in the blend, his right-directed gesture is a representation of her left-directed gesture.

Reified blend 1, laid out in Table 5.1, is a simplified diagram of this blend, noting only four points of our richly structured frames of reference.³ It presupposes that A and B are upright, facing each other, as interlocutors prototypically are; however, the cognitive process of mental rotation is versatile and potentially preserves all relevant structure. Since that cannot be completely represented in text form, I have restricted the diagram to elements relevant to the current data.⁴ Note that a blend very similar to this is involved in the ASL viewpoint shift mechanism described by Janzen (this volume), where a mental rotation allows a signer to represent a mirror image of a scene without in fact moving his or her body at all.

5.3.1.2 *Reified blend 2: imposing a canonical frame of reference*

Another common process is of imposing a frame of reference on an object that does not inherently have one (Lakoff 1987), blending our frame of reference with its inherent image schema structure⁵: thus a picture has left-right structure parasitic on the viewer's body (the picture is either blended with the viewer's own left-right structure (as in this case) or with that of an imagined person facing the viewer. Reified blend 2 is laid out in Table 5.2. Like the previous Reified blend, it is a simplification of this process for current purposes.

Table 5.2 *Reified blend 2*

Input 1: Viewer's Frame of Reference	Input 2: Object's spatial orientation	Generic image-schematic structure	Blend: Imposed Frame of Reference
Viewer's left side	side	Right-left image schema	"The/its left"
Viewer's right side	side		"The/its right"
Viewer's head	top	Up-down image schema	"The/its top"
Viewer's feet	bottom		"The/its bottom"

Table 5.3 *Discourse Blend 1*

Input 1: Listener's Frame of Reference	Input 2: L shifts to S's Frame of Reference	Input 3: Imposed Frame of Reference	Shared image schema structure of: Bodies Frames of Reference	Blend: Listener shifts to Speaker's Frame of Reference as canonical
L's right side	"Your left"	"The left"	Right-left image schema	"The left"
L's left side	"Your right"	"The right"		"The right"

5.3.1.3 Discourse blend 1: "the...left-hand side"

Reified blend 1, then, lets Listener construe his right, where he is gesturing, as Speaker's left. By Reified blend 2, Speaker's frame of reference is the picture's canonical frame of reference. Therefore, integrating all these spaces into the discourse's current Real Space Blend, Listener's right is *the left*. This is shown in Table 5.3. Figure 5.6 highlights the image-schematic structure being preserved in this blend.

As has been noted elsewhere in this volume, physical viewpoint is a marker of cognitive viewpoint. At this point in the discourse under examination, by preserving Speaker's frame of reference in his discourse blend and taking it as canonical, Listener is subsuming his own physical viewpoint to his understanding of Speaker's physical viewpoint, indicating a close alignment of his cognitive viewpoint to hers. This is natural, given his dependence on her for information about the picture.

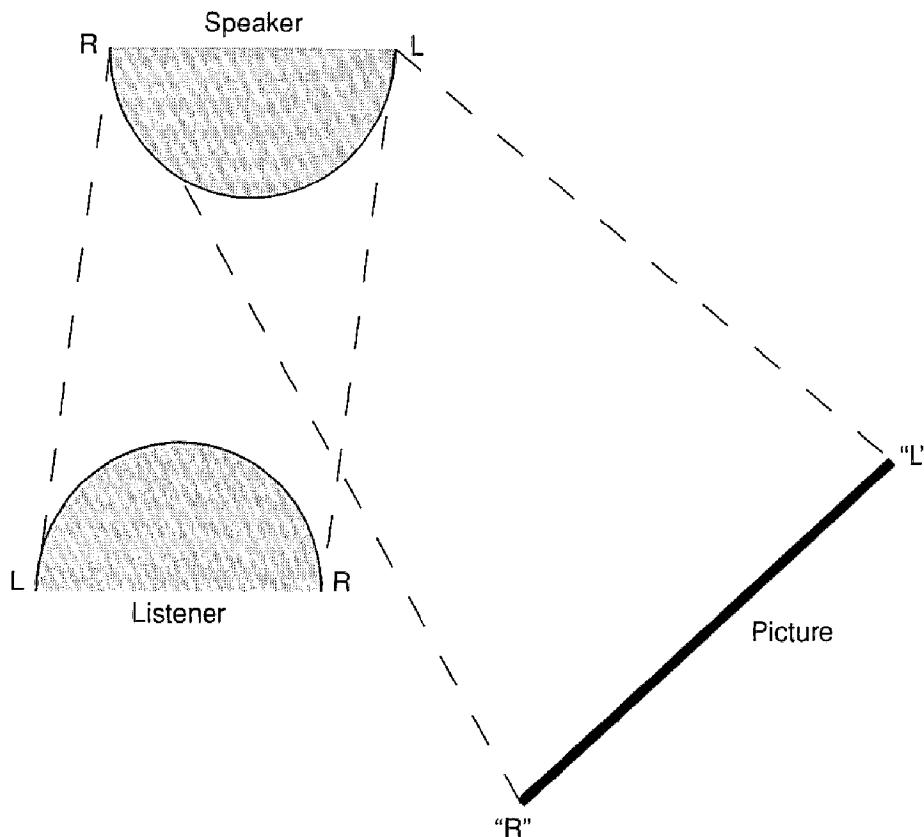


Figure 5.6 Image-schematic structure in Discourse blend 1

In the following sections, that alignment breaks down, until Listener gives up on it completely and begins to speak and gesture about the picture as though he were looking at it himself.

5.3.2 *Listener misunderstands Speaker's viewpoint*

The exchange continues as follows:

(8) *[2:40] Speaker*

Um, on the right-hand side of the street

There are...

Five cars. One of them is in the middle of the street,
driving.

[2:48] Listener

Hold on hold on h'lon

You can't say the right side of the street because whatever side of the
street you're on,

that's gonna be the right side, 'cause that's where you're driving.

Here we see two different viewpoints: Speaker's viewpoint is her Base Space – her Real Space, with herself looking at a two-dimensional image projected onto a screen. This allows her to talk about the panel boundary and the objects contained within; it is also a sign that she may not be comprehending the scene yet. The frame of reference that she preserves in her discourse blend is the one she has imposed on the picture.

Listener, however, is confused by her use of the word "right"; his confusion gives us a clue that he understands her as viewing the scene in the depicted image from inside – perhaps he thinks she is taking the viewpoint of someone driving a car along a street. It is only within this space, structured by the driving and street scene frames, that cars have to be on the right side of the street, "cause that's where you're driving." His constructed understanding of her blend preserves the frame of reference of the inferred driver in the depicted scene.

5.3.2.1 Discourse blend 2: Speaker's construction of the scene

In this analysis, my conception of Real Space differs from Liddell's (Liddell 2003), by which pictures are inherently part of Real Space. In Liddell's analysis, the mental spaces involved in understanding a viewed picture *are* part of Real Space, since visual processing makes depicted spaces accessible as part of our primary experience. Thus, for example, the image of a car in a picture is indeed an object in the physical Real Space of the viewer.

However, this analysis requires significant expansion to allow for the complex blending processes involved in understanding some pictures – no coherent understanding of this particular picture (or of many other comics images) could be reached by construing it just as a static object present in Real Space. Nor, therefore, does treating a picture simply as a Real Space object account for the misunderstanding discussed here – Speaker is construing the picture as a two-dimensional object viewed from outside, while Listener believes she is construing the picture as a three-dimensional space (and perhaps viewing it from inside the constructed scene). Since the misunderstanding arises from the subjects dealing with a picture in versatile ways, which suggest they have access to several spaces, more unpacking is necessary; pictures are not only objects in Real Space. The relationship between Real Space and picture spaces is perhaps better understood as a primary blend network, and this is how I analyze them.

Within this network, then, Speaker's blend at [2:40] preserves the frame of reference of her Real Space, as it has been imposed on the picture. This blend is diagrammed in Table 5.4.

5.3.2.2 Discourse blend 3: Listener's misunderstanding of Speaker's blend

If Listener had understood Speaker perfectly, then his Viewpoint Space, the one whose frame of reference was preserved in *his* blend, could simply be *L*'s

Table 5.4 *Discourse blend 2*

Input 1: Picture Space blend		Input 2: Speaker's Real Space	Input 3: Imposed Frame of Reference blend	Generic structure	Blend:
Form	Constructed meaning				
	Room Speaker				Room Speaker Speaker's frame of reference
Picture shape: square (has horizontal and vertical extent)	Listener Screen Speaker's left	"The left"	Right-left image schema	"The left"	Listener Screen "The left"
Grey wedge	Street	Speaker's right	"The right"		"The right side of the street"
		Depicted scene		Pictures (subtype: comics)	The street
Green ovoid Other shapes	Car (Beetle) Other cars, pedestrians, buildings, etc.				The Beetle Other cars, pedestrians, buildings, etc.
Words Bubbles containing words	Sound Speech				Sound Speech

understanding of S in Figure 5.4. However, *L's understanding of S* is not a salient part of the blend he is constructing at [2:48]; rather, he is elaborating his understanding of *Picture blend* "", already structured by the street scene frame, by blending in a space structured by the driving frame. As I said, this kind of blended picture interpretation, as a three-dimensional scene with participant viewpoints, has not so far been elaborated by Speaker.

The frame of reference that Listener's blend preserves is that of the Driver, an entity in the Driving Space. It is the Driver's viewpoint that Listener takes when he says that "whatever side of the street you're on, that's gonna be the right side, 'cause that's where you're driving."

Table 5.5 *Discourse blend 3*

Input 1: Picture Space blend		Input 2: Driving	Generic Structure	Blend:
Form	Constructed Meaning			
Picture shape:				
square				
Entities	Entities	Entities		Entities
Grey wedge	Street	Street		The street
	Driver	Driver	Driver:	Driver
	(implied)		instance of <i>human</i>	
				Driver's frame of reference
Green ovoid	Car (Beetle)	Driver's car	cars	Driver's Beetle
Other shapes	Cars	Cars		Cars
	Pedestrians	Pedestrians	Pedestrians	Pedestrians
	Buildings	(optional: buildings)	Buildings	Buildings
Words	Sound	Sound	Sound	Sound
Bubbles containing words	Speech	(optional: speech)	Speech	Speech
Relations	Relations	Relations	Relations	
		Driver sits in car to drive, facing forward		Driver sits in car to drive, facing forward
		Driver's Frame of Reference imposed on road		Driver's Frame of Reference imposed on road
Shapes along edge of grey wedge	Cars along side of street	Cars drive on “the right”	Cars drive on street	Cars driving on “ the right side of the street ”

This suggests that even at this early point in the discourse, Listener is creating a more coherent understanding of the image than Speaker is, as laid out in Table 5.5 as Discourse blend 3.

5.3.3 *Switching visual viewpoint*

The interchange featured in this section happens just after the one in section 5.3.2; the misunderstanding seen in section 5.3.2 may be cuing a more general realignment of Listener's viewpoint.

Table 5.6 *Listener's Real Space blend*

Input 1: Listener's Real Space	Input 2: Spatial orientation of Picture blend'''	Shared image schema structure of: Frames of Reference	Blend: Listener's Real Space blend of the reconstructed picture
Listener's right side	Vertical edges of picture	Right-left image schema	Vertical edge is on L's right side
Listener's left side			Vertical edge is on L's left side
Top of L's gesture space	Top of picture	Up-down image schema	Top of picture is at top of L's gesture space
Bottom of L's gesture space	Bottom of picture		Bottom of picture is at bottom of L's gesture space

(9) [3:02] **Listener**

So there's four cars,
RH 5, palm down, sweeps across gesture space, L bottom to R top
 there's four cars going up...
 Going up the, uhh...
RH 5, palm down, angles from R bottom to L top

At the beginning of this example, Listener is taking Speaker's viewpoint, as we saw in section 5.3.1. However, he then hesitates and changes his gesture so that it matches *his* left-right orientation, rather than Speaker's.

In terms of the ongoing blend that Listener is constructing, he is no longer making use of the *Shifting frame of reference* blend outlined in section 5.3.1.1, but rather creating a different Real Space Blend, diagrammed in Table 5.6, in which the *Picture Blend'''* Space is integrated with his Real Space in a location in front of him, and *his* frame of reference is imposed on the constructed image.

Creating a Real Space Blend gives a spatial orientation to *Picture Blend'''* – an imaginary, constructed “object.” By taking this Real Space Blend as an input to the Imposed Frame of Reference Blend (in a manner analogous to that in section 5.3.1.2), Listener imposes *his* frame of reference on his understanding of the picture. This is laid out in Discourse Blend 4, shown in Table 5.7.

5.3.3.1 Discourse Blend 4: Listener imposes his frame of reference on Real Space Blend of reconstructed picture

Listener's Viewpoint Space, then, has shifted entirely from *S Sees* to his own Real Space, and the reconstructed picture blended with it; even formal

Table 5.7 *Discourse blend 4*

Input 1: Listener's Frame of Reference	Input 2: Spatial orientation of blended "Object"	Generic image-schematic structure	Blend: L's Frame of Reference imposed on Real Space blend
Listener's left side	Vertical edge	Right-left image schema	"The/its left"
Listener's right side	Vertical edge		"The/its right"
Top of L's gesture space	Top	Up-down image schema	"The/its top"
Bottom of L's gesture space	Bottom		"The/its bottom"

elements now take the viewpoint of this blended space. His gesture and language no longer depict *L's understanding of S* (or spaces nested below that space). This suggests that Listener's cognitive viewpoint has become less closely aligned with Speaker's; he is no longer working to understand the image in terms of what Speaker sees, but to imagine his own view of the reconstructed picture.

Before Example 9, Listener has been asking for clarification and information. Once he switches viewpoint, the grammatical form of his questions changes: some are phrased as statements, which Speaker must accept or reject. The next two examples show Listener's last utterance before the [2:40] point, and his first one after the [3:02] point.

- (10) /2:39/ **Listener**
 how many cars are in the picture?
- (11) /3:09/ **Listener**
 Okay.
 They're all moving cars, no parked cars on the side of the street.

This suggests that, having shifted viewpoint to his own reconstructed picture blend, Listener is now talking about that constructed understanding of the image. Since his task is to align this space with that of Speaker's understanding of the picture (and, via Speaker, the picture itself), this makes sense.

5.3.4 *Mismatch*

At several points through the clip, Speaker has used a right arm hold, elbow down, diagonally across her face, to depict the top *edge* of the road, and has made iconic gestures representing objects *in* the road with her left hand, at points below her right arm. She does so again at [3:31], using her LH to

represent the green Beetle. Her fingers, which start in close proximity to her RH, map to the front of the car (which is close to the curb), and the heel of her hand maps to the back of the car (which is out in the street). Her gesture is unlikely to be ambiguous at this point, as she is elaborating on a pre-established iconic structuring of Real Space.

(12) [3:31] Speaker

And he's tryna [pull out] into the road.

R arm returns to diagonal hold

[LH B moves downward from R arm]

[3:32] Listener

he's tryna pull out into the road?

RH B, palm facing left, fingers forward, hand moves out from body

Speaker's left-hand gesture, co-timed with "pull out," is iconic for a car backing up, since the car could not be moving forwards legally in that orientation to the relevant side of the street, and since fingertips generally map to the "leading edge" of a moving object, most often the canonical front. The structuring of her Real Space blend, and thus the iconicity of the left-hand movement, should be unambiguous.

However, when Listener repeats her words, he makes a small gesture, mismatched to hers – and possibly mismatched to his speech. Assuming that both speakers are mapping the front of the car onto their fingers and the back onto their palms, Speaker's gesture iconically depicts the car backing out,⁶ but Listener's gesture depicts the car moving forward. In addition, if he has set up his Real Space Blend of the scene in a way consistent with Speaker's (as might be expected from his earlier, closely aligned viewpoint), then his gesture would map iconically to a car pulling *into* a space at the "top" of the road, not pulling out of a space.

This gesture suggests a viewpoint matching and elaborating that of Discourse Blend 3, preserving the frame of reference of the Beetle's implied Driver. While Speaker's gestures are on a two-dimensional space, suggesting that the picture in her Real Space Blend is still a flat object, Listener's gesture moves away from his body, suggesting a three-dimensional, egocentric representation, with the car at Ego location. His mismatch looks like a classic mismatch (Goldin-Meadow 2003) between gesture and speech, but it is hard to see how he could have made it if he had simply been mirroring Speaker. And indeed, like some of Goldin-Meadow's subjects' gesture–speech mismatches, it presages a new construal that has not yet emerged in speech.

Not long after this, Listener provides the interpretation that they both agree is correct.

(13) **[3:57] Speaker**

The car parked behind the bug is so hard to see, could even be a
taxicab, it's just a
yellow car, um . . .

It's hard to see because the word honk

Um

Is basically over the top of it.

[4:06] Listener

It's interesting that it says SCORE, maybe what it means is he actually
got the spot,

as opposed to . . .

looks at Speaker

[4:11] Speaker

Ohhh! Yeah!

[4:12] Listener

He scored the [spot]

[4:13] Speaker

[He's like [Score! I got a spot!]]

*[R arm held vertically; S hand, palm facing face,
moves diagonally across face]*

In the moments before he suggests this correct interpretation of the picture he cannot see, Listener stops responding to Speaker's speech and stares into space. His focus is clearly not on any space that pertains directly to Speaker; it seems to be on his reconstructed Picture Space Blend.

The breakdown we saw in section 5.3.2, where Speaker and Listener's iterative process of collaborative meaning construction failed due to a viewpoint misunderstanding, might actually be *helping* Listener here. His resulting disalignment of viewpoint from Speaker's, seen in sections 5.3.3 and 5.3.4, allows him to conceptualize the image in a way that she is not doing, and thus to create meaning that she is not constructing.

5.4 Discussion

In section 5.3, putting together speech and gesture data allowed us to trace the stages of Listener's conceptualization, which would only be very partially available from the speech track alone. We saw that Listener's visual viewpoint indexed a cognitive viewpoint that started closely aligned with Speaker's, then shifted; his attempts to build meaning were apparent in the interaction of his speech and gesture, as was a sense of the changing frame of reference he was working with as his cognitive viewpoint shifted.

An analysis of these data that focused merely on the speech would prove outright misleading; in only one of these crucial sections of the interaction could we have inferred much about viewpoint from speech alone (5.3.2: Listener misunderstands Speaker's viewpoint). Gesture data, therefore, are crucial to our understanding of Listener's viewpoint shifts, and therefore to our understanding of why his speech patterns shifted at the [3:02] mark, from confirmatory questions to declarative statements, leading to the construal that resolved Speaker's confusion.

Why might gesture prove crucial to the study of cognitive viewpoint? Certainly, one answer is that gesture, as a less conscious communicative track than speech, is crucial to the study of language in general; also, being visuospatial, it has much richer scope for iconicity than the speech stream (Taub 2001). But in addition to these general qualities – as mentioned at the beginning of this chapter, gesture is an embodied action, and thus inherently viewpointed. The iconic and deictic information coded in gesture can only *be* iconic or deictic with regard to some frame of reference (as well as some Focus Space, disambiguated by speech [Parrill and Sweetser 2004]). By imposing a frame of reference on the ongoing blend, then, the Viewpoint Space constrains possible gestures to those that can meaningfully occur within that frame of reference. Cognitive viewpoint imposes the frame of reference of a metaphorical “space” rather than a literal one; language is inherently viewpointed too, but rarely as unambiguously as gesture.

Gesture takes place in a physical space, with observable physical viewpoint – we can see what the subject is looking at, or in what direction he is placing his arm to iconically represent an object. Language is often hugely ambiguous between reference in different spaces; it is actually useful for interlocutors to consider that they are talking about the “same” thing in referring to S’s imagined picture and L’s imagined picture, or to the physical picture on the screen and the imagined contents of the picture. But for the analyst, trying to unpack these spaces that language conflates, gestural viewpoint provides invaluable cues to distinguish parts of the network from each other.

And by imposing a frame of reference on the blend, the Viewpoint Space constrains cognition; it is the space *in whose terms* cognizers conceptualize the rest. Viewpoint, then, is indexed physically by gesture and provides a constraint on the meaning that can be constructed; in turn, gesture provides a channel through which viewpoint can be analyzed. This tight link between speech, gesture, and cognition is unsurprising; as Sweetser (1998) notes, language is a manifestation of a cognitive system “concerned with interaction and the situation of the person.” With regard to aspects of cognition such as viewpoint, which depend crucially on our embodied understanding of space, the surprise would be if gesture analyses were *not* crucial to our understanding of communication.

Yet another general lesson reinforced by this example is that cases where communication and reasoning processes break down are often more revealing to the analyst than cases where all goes without a hitch. Smooth interaction is generally transparent to participants – and often transparently understandable to the analyst as well. On the other hand, breakdown and renegotiation force presupposed and transparent mechanisms to the fore, and profile the cognitive underpinnings of framing and viewpoint structure.

References

- Coulson, Seana. 2001. *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge University Press.
- Dudis, Paul G. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15:2, 223–38.
- Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fauconnier, Gilles and Eve Sweetser. 1996. *Spaces, Worlds, and Grammar*. University of Chicago Press.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di semantica* 6:2, 222–54.
- Goldin-Meadow, Susan. 2003. *Hearing Gesture: How Our Hands Help Us Think*. Cambridge MA: Harvard University Press.
- Hampe, Beate and Joseph E. Grady (eds.). 2005. *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlin/New York: Mouton de Gruyter.
- Kosslyn, Stephen M. 1980. *Image and Mind*. Cambridge MA: Harvard University Press.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.
- Liddell, Scott K. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press.
- Parrill, Fey and Eve Sweetser. 2004. What we mean by meaning: conceptual integration in gesture analysis and transcription. *Gesture* 4:2, 197–219.
- Sweetser, Eve. 1998. Regular metaphoricity in gesture: bodily-based models of speech interaction. In *Actes du 16e Congrès International des Linguistes* (CD-ROM). Elsevier.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Cambridge MA: MIT Press.
- Taub, Sarah F. 2001. *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge University Press.



Part III

Multiple viewpoints in American Sign Language



6 Reported speech as an evidentiality strategy in American Sign Language

Barbara Shaffer

6.1 Introduction

This study looks at one of the many ways that speakers of American Sign Language (ASL) infuse their discourse with their viewpoint. As people speak, they convey information and simultaneously impart their viewpoint, opinions, and beliefs about what they say. Some utterances convey more speaker perspective than others, and thus, warrant special consideration. One such subjective linguistic device involves the use of what are termed by Aikhenvald (2006) and others “evidentiality strategies.” The research presented here contributes to the growing body of literature providing an analysis of the function and structure of signed languages from the perspective of Cognitive Grammar (e.g. Janzen and Shaffer 2002; Wilcox 2002; Liddell 2003; Dudis 2004). In particular, this study extends the very promising path of ASL research using Conceptual Integration Theory (Fauconnier 1985; Fauconnier and Turner 1996; Dudis 2004). In his 2004 study of body partitioning and Real Space Blends, Dudis showed that a Mental Spaces approach was particularly useful to illustrate how signers make use of various body parts to convey complex concepts spanning both Space and Time.

In this study, I show how such an approach is also useful in describing how deaf native signers indicate the source of the evidence they have for claims they make in their discourse. An evidential is any of the set of devices in a given language used to indicate the nature of evidence for a statement, and, more broadly, how a speaker chooses to mark the veracity of that statement (Chafe and Nichols 1986).

In some languages, speakers indicate the source of their information with a grammatical marker called an evidential, sometimes referred to as a grammaticalized evidential (Aikhenvald 2006). In English, and many other languages, this information is often encoded through the use of words or phrases such as *I've heard* and *they say/said*, as in Examples (1) and (2).

- (1) I've heard he takes steroids.
- (2) They say she's really nice.

In (1) and (2) the speaker is simultaneously making a claim (“he uses steroids,” “she’s really nice”) and indicating that there is some (evidential) basis for this claim. The specific source of the evidence is not named, nor does the speaker explicitly state whether or not she believes the statement, though the choice of a particular evidential can, through invited inference, suggest much about the speaker’s own beliefs. It is common for such “hearsay” evidentials to cast some doubt on the truth of the proposition (see Travis 2006 for a description of one such evidential, *dizque*, in Spanish).

Aikhenvald (2006) prefers the term “reported evidential” for an evidential that is a “verbal” report, such as the English periphrastic *I’ve heard*. She suggests that “quotative evidential” should be reserved for reported evidentials where the identity of the “reporter” is specified. For the purposes of this chapter, “reported speech” and “reported evidential” will be used to describe those forms that function to index information to a spoken/signed/written source and to an interpreter of that source (typically the speaker) (Mushin 2001), whether the source is explicitly stated or left unspecified (as in “someone said”).

Another commonly studied type of evidential is the so-called inferential evidential. Its meaning and use differ in significant ways from reported evidentials. In (3), the periphrastic *I’ve heard* is used as a reported speech evidential to inform the listener that the speaker has heard, or read, or otherwise learned that rain is in the forecast. In (4), on the other hand, the speaker uses evidence available (clouds, wind, people walking by with umbrellas, etc.), on the basis of which an inferential conclusion is drawn.

- (3) I heard it’s going to rain today.
- (4) It looks like it’s going to rain today.

Inferential evidentials are frequent in ASL discourse. In (5), a bystander has previously seen a married couple arguing, has heard that the husband has moved to a hotel, and now notices that the husband is no longer wearing his wedding band. He makes the following statement of deduction using SEEM:

- (5) [TIM, JENNIFER DIVORCE]-TOPIC
[SEEM]-HEAD NOD/BROW FURROW
It appears that Tim and Jennifer are going to get a divorce.

As in (5), the evidential marker is often seen in the comment of a topic-marked construction. Shaffer (2004) noted that the speaker’s subjective stance is often found in the comment of such constructions.

Evidentials that report the statements and stances of others – what Aikhenvald would call reported evidentials and the related “quotatives,” which explicitly name the source – are also a frequent and interesting phenomenon in ASL discourse. They are also quite complex, and little, if anything, has been written

about them in the literature. As such, they will be the primary focus of this chapter. In the following sections I will describe this interesting use of reported speech in ASL and introduce some elements of the conceptual integration framework to illustrate how signers create blends where constructed discourse is used to manifest reported speech evidentials.

6.2 Conceptual Integration Theory

Drawing on Conceptual Integration theory, as outlined in Fauconnier (1985) and Fauconnier and Turner (1996), and applied to signed languages by Dudis (2004, 2007), I will describe how signers make use of the space around them in the expression of reported evidentials in ASL. In these constructions, the signer creates a blend where he uses constructed, yet compressed, prior discourse (see Dudis 2007) to recreate a segment of a conversation he shared with another speaker. It is in that constructed discourse that he shares his evidential stance.

Liddell (2003), Dudis (2004), Wulf and Dudis (2005), and others have described how ASL signers manifest conceptual blends. They use the conceptual integration framework described by Fauconnier and Turner (1996, 1998). Dudis (2004: 226) introduces essential elements of the approach below.

[F]our interconnected mental spaces are involved in the blending process. Two of these mental spaces serve as inputs to another mental space, the blend. . . . [O]ne of the inputs is real space which, as Liddell (1995: 21–23) describes, is a mental space built via perceptual processes. This is distinct from the other input in that elements in the former are understood by the conceptualizer as being part of his or her immediate environment; Liddell (1995) uses the term “grounded” to label this particular property of real space. In signed discourse, the real space of the addressee would consist of her conceptualization of the signer via visual input. That is, the signer in the addressee’s real space is a mental space element. Other objects that are visually accessible are also potentially real-space elements. The empty physical space is also a real-space element.

The other input is referred to as a non-grounded mental space. Dudis calls this a narrative space, containing elements introduced during a given narrative. Of significance to the discussion here are the unique ways that signers can partition their bodies (such as their hands and eyes) to show the actions, and reactions, of someone else. When the signer is part of the blend a visible |actor| results (Dudis 2004). To illustrate this, Dudis describes how a signer might make use of body partitioning in a narrative about riding a motorcycle up a hill. The signer may create a Real Space Blend in which parts of his own body are understood to be those of the |motorcyclist| in the narrative. He notes that “the signer’s body contributes a great deal of information about the |motorcyclist|—the way he moves, his facial expression, how he responds to events, and so forth” (Dudis 2004: 224).

In this chapter, I will describe how signers use conceptual blending as an evidentiality strategy. In such blends (for simplicity, called “what I heard” blends here), a past discourse context is activated with two or more discourse participants, referred to as the |informer| and the |past self|. Many blends of this type make use of body partitioning, where one of the signer’s hands is understood to be the signing hand of the |informer|, a discourse partner from some past conversation. Other parts of the signer’s body are understood to be the signer’s (|past self|) in that same past discourse event. When such a blend is activated, the |informer| relates some piece of important (and possibly controversial) information, and the |past self| reacts to it. This is typically done via constructed dialogue. It is the conveyance of the constructed dialogue that is evidential in nature, as shown in (6). All the examples below come from a larger set of data consisting of dyads of native ASL users engaging in natural conversations. Filming took place in both the United States and Canada.

- (6) YES YES AND PRO.1 HEARD (gaze right) \exists SIGN.TO₁ (leans forward) #WHAT**

(gaze center) V-R-S (nod) START CUT (left hand)

CUT (right hand) CUT (left hand)

[HAPPEN SEE NAME]-TOP

[KNOW WHO]-TOP

CUT REFUSE ANSWER

[WHY]-TOP SOMETIMES SIT NAKED BODY SECOND BAD
SWEARING

[TRUE]-Y/N

Yes, and I heard, well, I was told, and was shocked to hear, that the VRS (companies) are starting to disconnect calls. If they see a name that they know they don’t answer, because sometimes people have answered naked, or used profanity. Is that true?

In (6), the signer and her addressee are discussing the use of video relay services (VRS). She says “I heard,” then pauses, shifts her eye gaze from the addressee and gazes to the right, activating the blend. Next she signs, “I was told” (literally: “it is signed to me”), and then responds (to the past discourse partner) to what she was told with a look of surprise and the reply “what!?” She then returns her eye gaze to her addressee, thus deactivating the blend, and only then relates what she was told in the past discourse event and invites her addressee to comment on it.

6.3 The complex problem of evidentiality

Evidentiality and related discourse phenomena have been the subject of numerous studies. Unfortunately, little consensus exists as to the nature of the category or its natural, logical limits. It seems to be a case of “we know it when we see it,”

but beyond that, much is left to the linguist attempting to describe it. This is due, in large part, to the fact that evidentiality is a discourse strategy that is expressed in different ways depending on language type and use. In the introduction to their well-known book, *Evidentiality: The Linguistic Coding of Epistemology*, Chafe and Nichols (1986: vii) somewhat whimsically describe evidentiality as “the ways in which ordinary people unhampered by philosophical traditions, naturally regard the source and reliability of their knowledge.”

A second complicating factor has to do with the variability in functioning of evidentials, perhaps confounded by the first issue of delimiting the category. While most researchers agree that evidentiality is best considered distinct from modality (but cf Palmer 1986), little else is clearly agreed on. For example, Mithun (1986: 89) states that evidential markers “specify the source of evidence on which statements are based as well as their degree of precision, their probability, and expectations concerning their probability.” She goes on to say that the specification of the source allows the speaker to “abdicate” some of the responsibility for the reliability or truth of the proposition, and allows the listener to come to his or her own conclusions regarding the truth of the statement. As Mithun also notes, the reliability of a statement can be qualified by specifying the probability of its truth. Evidentials are often used as a hedge whereby the speaker can again abdicate responsibility for truth. If the speaker turns out to be wrong, the hedging has guarded against a violation of the Gricean maxim of truthfulness.

De Haan (1999b: 7–8), on the other hand, states that evidentials are “a priori unmarked with respect to a commitment to the truth of the speech utterance on the part of the speaker. Evidentials merely assert that there is evidence to back up the speaker’s utterance.” While de Haan acknowledges that an evidential may invite a particular inference by the listener, he believes that such connections are secondary.

With respect to discourse function, de Haan (2005) states that the role of an evidential is to denote the relative distance between the speaker and the action. An indirect evidential (which Lazard [2001] and others aptly refer to as a “mediative”) will be chosen when the speaker wishes to state that the action took place outside his or her deictic sphere, whereas the choice of a direct evidential indicates that the action took place within that deictic sphere. I will return to the deictic notion of evidentiality in a moment. Here, however, I will provide a quote from Lazard to contrast with de Haan’s description of evidentials in order to illustrate the variability among researchers’ discussions of the function of an evidential.

The opposition is not direct vs. indirect knowledge, old vs. new knowledge, or assimilated vs. unassimilated knowledge. Rather, it is an opposition at the morphosyntactic level between forms indicating nothing about the source of the information and forms referring to the source of the information without specifying it. “Ordinary”, non-evidential forms state the facts purely and simply. Evidential forms, on the other hand,

point to the speaker's *becoming aware* of the facts. In the case of inference it implies "as I infer"; in the case of unexpected perception it implies "as I see". The speaker is somehow split into two persons, the one speaking and the one who has heard, inferred, or perceived. (Lazard 2001: 362)

From these descriptions of evidential use a common theme emerges. An evidential indexes a certain distance and relationship between the speaker and the proposition, as de Haan suggests, or perhaps more succinctly: "between the speaker and his or her own discourse, or between the speaker as the person acquiring evidence and the person expressing it" (Lazard 2001: 362). In essence, evidentials function to reduce the speaker's responsibility for the information they convey.

Because evidentials point to the time a speaker "becomes aware" of facts, a conceptual blending analysis of them is extremely illustrative. And, because evidentials are deictic in nature, pointing both to a propositional content and to a source of evidence, it seems logical that in a signed language, that deictic relationship would be spatially marked. I return to the deictic nature of evidentials in the next section, but first will provide several more examples of reported speech evidential blends.

6.4 The construction of reported speech evidential blends

Before going further, let me describe in more detail how signers use their bodies and the space around them to express reported speech evidentials. I will use Example (6) from above. Figure 6.1 shows the positioning of the interlocutors and the discourse space between them that they use to construct their utterances.

The signer on the right is relating something she previously heard about a video relay service provider.¹ She has heard that relay providers keep a log of users' names and numbers and will not provide service to individuals who misuse the service. To share (and frame) this information, the signer activates a "what I heard" blend. Here she shifts her eye gaze to a locus to the right of the central signing space she and her interlocutor share. Figure 6.2a shows the signer's eye gaze towards her interlocutor and the discourse space between them. Figure 6.2b shows her eye gaze shift to the right (activating the blend), as she begins the report of what has been said previously. She begins by saying "I heard" (while looking at her addressee), shifts her gaze to the right, then states "it is told to me" (literally, "it is signed to me"), and then relates her reaction to what she has been told. Her reaction is translated here as "what?!" She then returns her eye gaze to the addressee and relates what she has heard.

In such a construction, the signer creates a blend where the original discourse, now being reported, occurred. She reconstructs the act of being told, and her

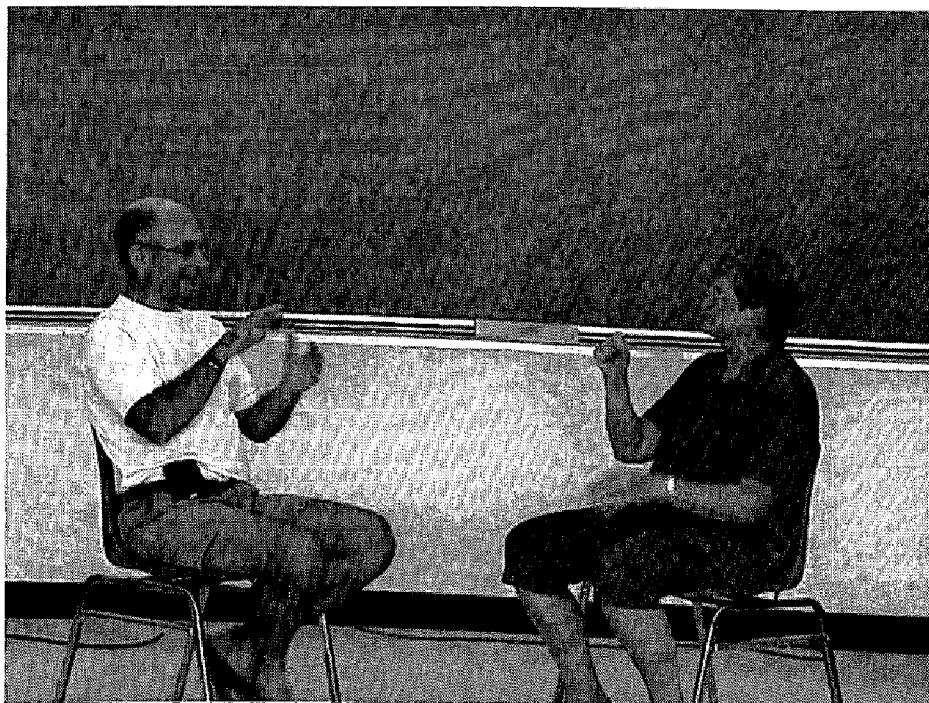


Figure 6.1 Discourse setting for Example (6)

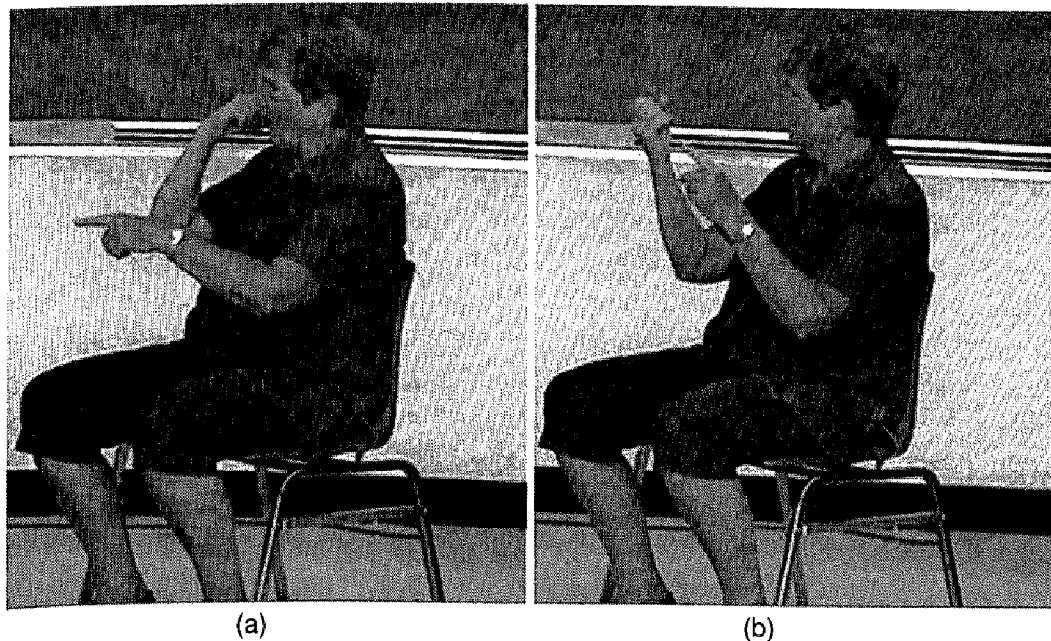


Figure 6.2 "I hear." The blended space where "it is signed to me" is produced

reaction to the information. She visually creates, for the interlocutor, the scene where the prior conversation occurred. She shifts her gaze to that scene (and to a distinct physical space), thus temporarily disengaging from her interlocutor and cueing him to follow her into the Event Space (where the conversation she is relating took place) that she has created. She literally *points* to the evidential source (with her eye gaze) in the blend, as she relates the conversation and her reactions to it.

The signer takes an entire discourse segment, presumably several sentences in length, and shortens it, what Dudis would call a temporal compression (Dudis 2007). What the unidentified person (the |informer|) said to her is now compressed into a single ASL sign, translatable into English as “signed to me.” Her response (as |past self| experiencing it) to it may also be compressed, and is now simply “what?!” She then deactivates the blend, returns her gaze to her interlocutor, and goes on to report, in the present discourse event, the details of what she has heard. In (6), for example, she reports that VRS companies are now denying certain customers access based on prior misuse of the service. Finally, she invites her discourse partner to comment on what she has relayed.

Of particular interest in this blend is the body partitioning feature previously described in detail by Dudis (2004: 223): again, these are “the partitionable zones of the body available to ASL signers. These zones are the body subparts that can participate in mappings that create blends (Fauconnier and Turner 1996) that have as one input Real Space, a mental space built by perceptual processes (Liddell 1995).” In the “what I heard” blend, the signer produces “it is signed to me” with her right hand. We easily understand her right hand at this point to represent her past interlocutor’s hand. We also know that the past interlocutor actually told her something, and did not simply say, “you are signed to.” Of note, the phrase “it is signed to me” is in a passive construction (see Janzen *et al.* 2000, 2001; Janzen and Shaffer 2007), where the right hand represents the hand of the demoted, yet still present, agent of the passive. The signer (the |past self|) is the addressee of the utterance. She is in view and profiled in the utterance. Her facial expressions depict her past discourse reaction to what she was told. Next, her right hand (which no longer represents the hand of her interlocutor, but rather her own hand) signs her own past reaction to the information she was (just) given. The blend is then deactivated, and she returns her eye gaze to her interlocutor and continues her discourse. The entire sequence is seen in Figure 6.3, including Figure 6.3e, which shows her eye gaze return to her interlocutor, signaling that the blended space is no longer active.

On returning her eye gaze to her interlocutor, the signer reports what she has heard. She does not quote her source, but rather paraphrases, and elaborates on it as needed, in order to provide enough information and thus accomplish her communicative goals. Her reference to the unidentified interlocutor (the

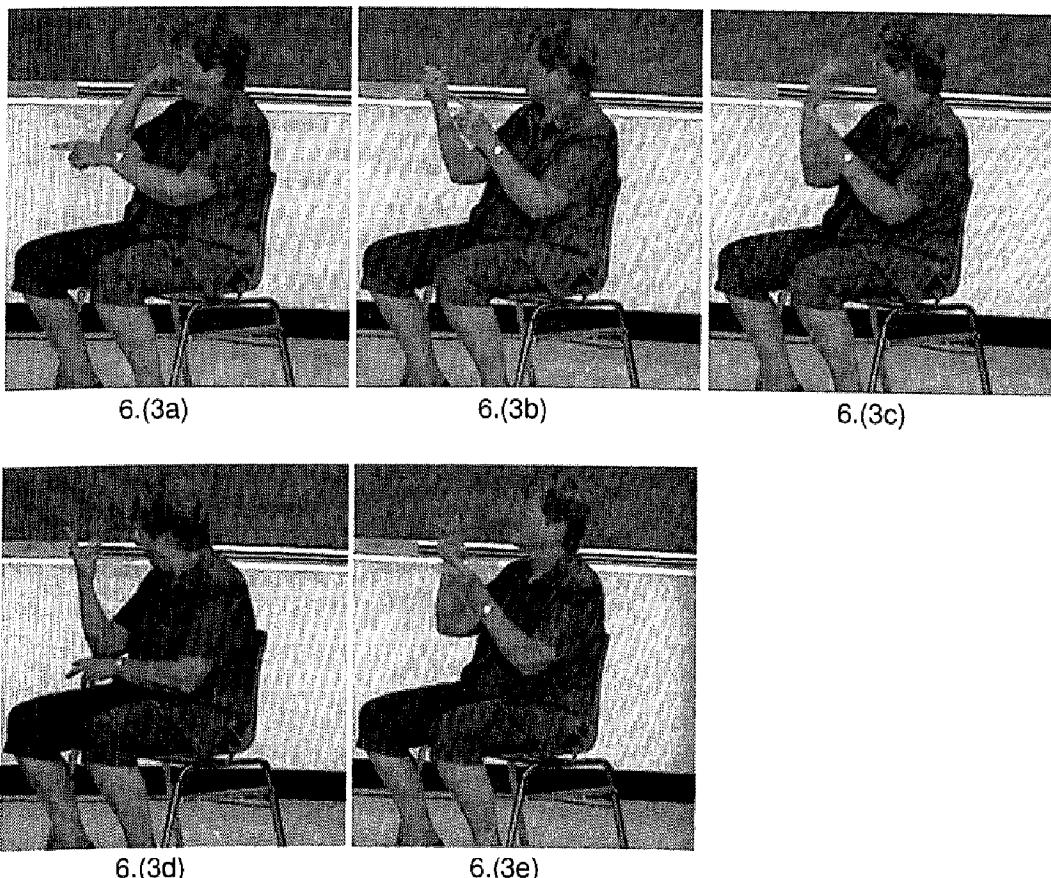


Figure 6.3 (6.3a): “I heard”; (6.3b) and (6.3c): “it is signed to me”; (6.3d): “I said ‘what?!’” and (6.3e): VRS

|informer|) in the blended past space does not introduce a new discourse participant, but rather provides information about *the signer’s* judgment regarding the veracity of the truth of her own proposition. In essence, the signer can prime her interlocutor with her reaction and thoughts before relaying what she has heard. The blended space could be interpreted as a hedge, as Mithun (1986) describes it, serving to distance the signer from what she is saying, and abdicating her responsibility for the truth of the proposition.

In contrast to Example (6), where the signer partitions her body to indicate that someone is signing to her and to depict her responses to what is said, in Example (7) (and others like it), the signer constructs a blend where he only shows his reaction to the information he has learned and relates that information. We do not see any compressed segment of the “evidence.” In some examples the source of evidence is named, while in others it is just referred to generally (such as “someone said”). The signer’s eye-gaze towards the conceptualized location of that past discourse event conversant and his reaction towards the

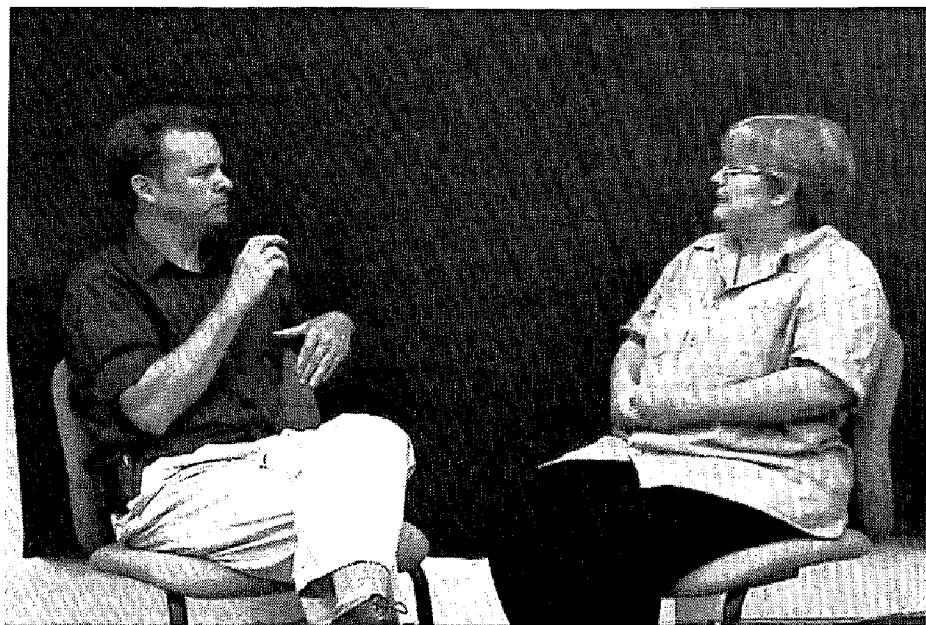


Figure 6.4 Discourse setting for Example (7)

blended space are the only traces of the other interlocutor. The discourse space for Example (7) is seen in Figure 6.4.

Prior to the discourse segment, transcribed in Example (7), the signer is looking at the addressee, as shown in Figure 6.4. He begins this segment with “someone said,” and then shifts his gaze to the right.

- (7) **(eye gaze right) SOMEONE SAY**
PRO.1 OH.I.SEE GOOD IDEA
(eye gaze to addressee) SUPPOSE PRO.1 N.C. CALL VRS index
left CHARLOTTE
GO.TO OTHER STATE
LESS CONFLICT INTEREST
KNOW EACH.OTHER
LESS CONFLICT INTEREST
(eye gaze right) PRO.1 OH.I.SEE
 Someone told me something interesting. I thought it was a good idea. If I’m in North Carolina and call the video relay service I won’t get the one in Charlotte, I’ll get one in another state. There’s less conflict of interest that way, because we don’t know each other. So there’s less conflict of interest. Interesting.

Here, the signer says “someone told me,” and then activates the blend. He shifts his eye gaze right, and signs his reaction (OH.I.SEE, translated as “interesting”) to what he was told in the previous discourse setting. However, there is no

constructed discourse from his conversant, only a constructed discourse space where the signer looks, and where he reacts to what he was previously told. He then returns his eye gaze to his interlocutor, deactivating the reported speech blend, while simultaneously signing “good idea,” and then begins to convey the reported information.

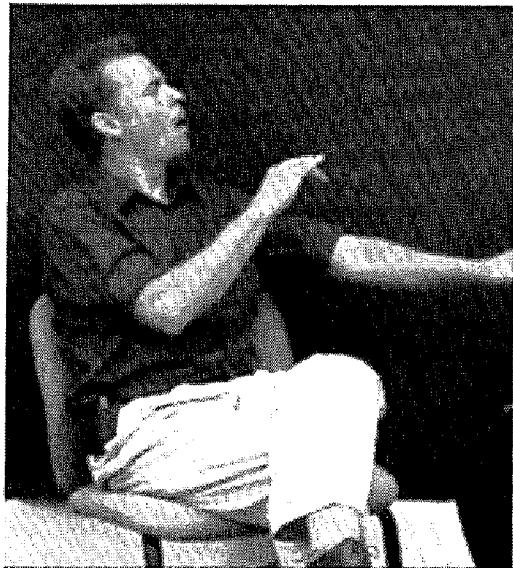
Interestingly, at the end of the reported segment, he returns his eye gaze to the previously active blend space, thus reactivating it, and again signs OH.I.SEE “interesting.” Figure 6.5 illustrates the shift in his eye gaze and deictic indexing of the previous Event Space.

Similarly, in Example (8) (from the conversation shown in Figure 6.1), a “what I heard” blend is activated, as the (Canadian) signer introduces his past speech event interlocutor (the [informer]) as “an American, I can’t remember who.” The signer shifts his eye gaze to the Past Event Space and we understand his eyes to be his [past self] eyes as he is viewing and reacting to the [informer]. The [past self] addressee shows his surprised response to what he learned. The blend is deactivated and only then does the signer convey what he learned from the American about deafness, and the grieving process parents go through on learning that their infant is deaf. Again, the signer primes his addressee to consider the information in the same manner with respect to its truthfulness as he has.

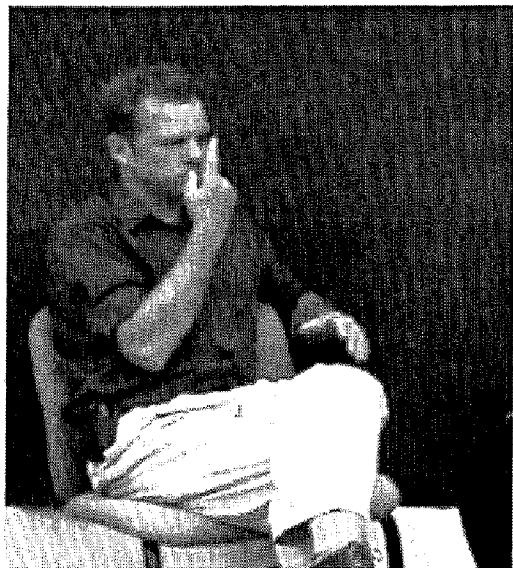
- (8) [AMERICAN THINK WHO PRO.3 COMMENT]-TOPIC [I
LOOK.AT]-BACKWARDS HEAD LEAN
RESEARCH FIND MOTHER.FATHER FIND DEAF 48H-R-S LESS
GRIEF

I saw someone, an American I think, say, and thought “huh, interesting,” that if parents find out their child is deaf in the first 48 hours they experience less grieving.

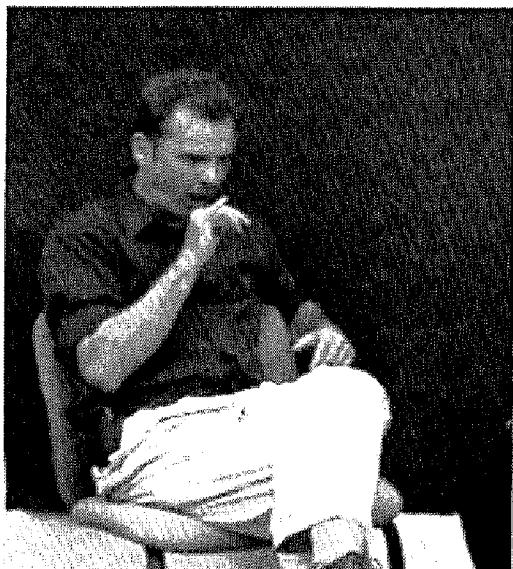
In Example (9), also from the discourse setting seen in Figure 6.1 above, the male signer is again relating the “evidence” presented to him by another American scholar (also about deafness, and the grieving process), along with his reaction to it. Here, the individual is specifically named, and he states that he learned the information via a lecture. Other than the naming of the source, Example (9) is essentially identical to (8). There is no physical trace of the evidential source. We are not intended to believe that the verb “LECTURE” is a depicting verb (Dudis 2007) showing how someone made a presentation, or that it is part of the blended space. The only blended element seen is the signer (the [past self]) and his reaction. And, of course, we understand that Arlene Kelly and all elements of her presentation are part of the blend as well.



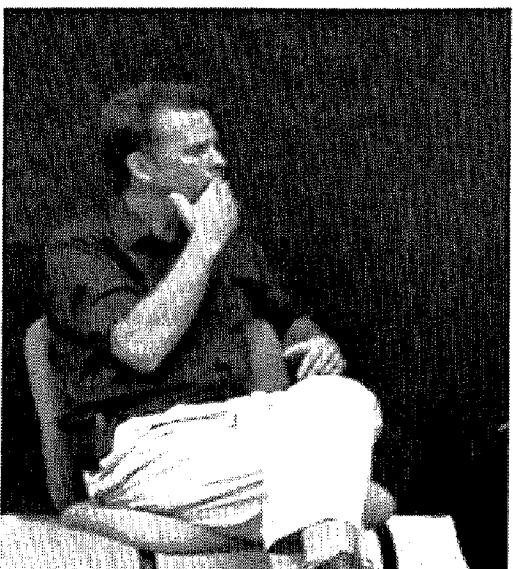
6.(5a)



6.(5b)



6.(5c)



6.(5d)

Figure 6.5 (6.5a): discourse segment begins; (6.5b): “someone said” (eye gaze toward constructed discourse space); (6.5c): “interesting” (with eye gaze to constructed space); (6.5d): eye gaze returns to interlocutor and signer resumes discourse with “good idea”

- (9) [YOU KNOW ARLENE KELLY DR. COME LECTURE]-TOP
PRO.3 COMMENT ME LOOK (head bob)
IF GO INTERNET CLICK LIST
OH.I.SEE NAME,
ME FORGET QUOTE NAME PUBLICATION VIDEOTAPE
NO NOT VIDEOTAPE CD ROM DVD WHATEVER DISC ADVER-
TISE
PARENTS ORDER GET
[PARENTS BIRTH DEAF]-TOP
[FIND FREE]-YES/NO QUESTION
[PUBLICATION FREE]-YES/NO QUESTION
[INFORMATION]-YES/NO QUESTION
GET FREE

You know Dr. Arlene Kelly. The one who came and presented. She said, and I thought "huh, interesting," if you look on the internet, there's a site. I forgot the name of the publication and videotape. Well, no, it's not a videotape it's a CD-ROM, or a DVD. A disc anyway. It's an advertisement for a free disc. That way, if parents have a deaf child they can order this free publication. They can get this information (about deafness) free.

Example (10) is also from the ongoing conversation about deaf education. It has just been stated that people believe that if a deaf child is allowed to sign, he or she will not learn to speak. The signer believes this is not the case. He reports the "evidence" while stating that he cannot remember who said it. The signer still breaks his eye gaze with the interlocutor, and looks to a past discourse space. However, in Example (10), he quickly shifts his gaze back to his interlocutor. No reaction to the information from the previous discourse setting is provided. Further investigation is warranted here, to determine if utterances with unnamed sources of evidence are consistently structured like (10).

- (10) NOT TRUE
(eye gaze right) [WHO COMMENT]-WH QUESTION
(eye gaze center) SAY MANY PARENTS PRO.3 DEAF (child)
ALLOW SIGN
SPEECH DEVELOP FINE
That's not true. I don't remember who said it, but they said that many parents of deaf children allow their kids to sign and their speech develops well.

6.5 Evidentiality and deixis

As we have seen, reported speech as an evidentiality strategy in ASL has several related characteristics. First, it functions to report information. Second, because the speaker shows his or her reaction to the reported information before (and sometimes after) making the statement, the listener is primed to react to the information in the manner the speaker wishes – that is, if the speaker wants his listener to believe the report, he indicates his affirmation of what he is about to say. If the signer is unsure of the veracity of the statement, he indicates that accordingly. And, finally, depending on the speaker's goal, he can choose to make the [informer] more or less prominent in his discourse, giving the source increased or decreased status, thereby strengthening (or not) the status of the utterance.

Also note that reported speech evidentials in ASL deictically index the source of the information. Regardless of the language under investigation, a reported speech evidential could easily be argued to index a prior speech act. In fact, Schlichter (1986: 58) argues exactly this: "Wintu employs evidential deixis to place an event in the context of the other events" that are entailed by it. Just as deictic terms for place, time, and person are shorter terms standing for longer descriptions, so too are evidentials. However, in ASL, the "pointing" happens quite literally and iconically. The speaker suspends his discourse and indexes, with his eyes, the physical and mental space in which he reconstructs a previous discourse segment. He points to the prior speech act, and to the source of the evidence he is relying on. As Schlichter (1986: 57) aptly notes (for Wintu):

Deixis . . . has two components: a locating component in which the speaker places the event under discussion in time and space relative to the speech act; and an interpreting component in which it is up to the addressee to correctly interpret what he hears by reversing the first and second person pronominal reference and making all the adjustments necessitated by his own identity, individuality, and position in time and space as different from that of the speaker. When the speaker uses deictic expressions he organizes the information he gives with respect to himself as the referent, although he may occasionally, for purposes of politeness, take the point of view of the addressee.

In fact, de Haan (1999a) notes that direct evidentials typically arise from one of two primary sources: (1) deictic or demonstrative morphemes; or (2) tense or aspect markers (and, of course, tense itself is a deictic category). De Haan (1999a) suggests that when a speaker uses a visual evidential (developed from a demonstrative marker), he or she is saying that the action was witnessed personally because it occurred in the same deictic sphere as the location of the speaker.

ASL discourse participants utilize the spatial cues available to them to signal to their interlocutors that a previous discourse setting is being constructed, and

that the utterances and emotional reactions to what is said belong to that previous discourse. Their interlocutors correctly interpret the dialogue and emotional responses entailed as displaced from the moment of speech. Fleischman (1989) and others have argued for an egocentric view of languages, where the speaker is understood to be the center of all language use (de Haan 1999a). It seems clear that the displaced nature of ASL reported speech evidentials and deictic reference to that dialogue fit with such a conceptualization.

6.6 Summary and conclusions

Clearly, this chapter represents the mere tip of the iceberg for what I believe will be an interesting and fruitful area of investigation. As researchers such as Dudis (2004) and Janzen (this volume) have shown, signers make constant use of their bodies and the space around them as they construct their discourse. They are able to partition their own bodies to represent the bodies and actions of themselves and others as they construe them. This line of research is still in its infancy, but we do know that body partitioning and simultaneous constructions can be quite complex. In fact, Janzen (this volume) shows us that in some blends, signers are able to use body partitioning in extremely complex ways, representing at least two third person referents. In other words, the signer may not be represented in the body partitioning in certain blends.

This study describes what this looks like when signers use reported speech as an evidentiality strategy. Typically, here, the signer creates a blend where his |past self| becomes aware of new information and reacts to the new knowledge. As Schlichter (1986) noted for Wintu, ASL discourse partners are able to make all the necessary adjustments to correctly interpret pronouns, directional verbs, passives, and other deictic referents in such constructions, and have no difficulty understanding what is being said or the viewpoint being conveyed. In addition, interlocutors easily know that when someone signs “it is signed to me” in a past discourse space, that much more was said, and that, logically, the words “it is signed to me” were never said at all. In other words, they understand that “it is signed to me” is a compression (Dudis 2007) of the relevant parts of conversation that will soon be conveyed.

It is of note that in all examples in this data set, the |informer| is always at roughly the same height as the signer. That is, while “the American scholar,” while in the same room as the signer, was likely quite far away, and likely elevated relative to the signer during the described event in the blended space, the |informer| in the blend is positioned at the same height in the signing space as the |past self|. This could be evidence of some early grammaticalization of these constructions. Of course, more research is clearly warranted to investigate the many possible uses of reported speech evidential blends by ASL signers, both

those that utilize body partitioning to represent other signers in past discourse settings, and those that only depict the signer's "past self" reactions.

The grammaticalization processes noted here involve the expected physical reduction of form; as just mentioned, the imagined past communicative scene is not only abstracted away from the actual words used in it, but is "compressed" to the extent that it does not show height or distance of communicator. They also involve classic semantics of subjectification, in that the ground of the communicative act – here, a past act of communication – becomes less objectively expressed and profiled, and more implicit (Traugott 1982, 1989, 1995; Langacker 1990; Traugott and Dasher 2002; Ferrari and Sweetser [this volume]). The visual nature of ASL allows mental space structure, and degrees of elaboration of mental space structure, to be accessible to the analyst in ways that they might not be in a spoken or written language text. As Narayan (this volume) points out, spoken language is often very ambiguous about location in a space network – a good reason for spoken language analysts to attend to gesture. In ASL, the signer's gaze is present for us to see, with all that it tells us about space structure.

References

- Aikhenvald, Alexandra Y. 2006. *Evidentiality*. Oxford University Press.
- Chafe, Wallace and Johanna Nichols (eds.). 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
- de Haan, Ferdinand. 1999a. Evidentiality and epistemic modality: setting boundaries. *Southwest Journal of Linguistics* 18:1, 83–101.
- 1999b. Visual evidentiality and its origins. Unpublished manuscript, University of New Mexico.
2005. Encoding speaker perspective: evidentials. In Zygmunt Frajzyngier, Adam Hodges, and David S. Rood (eds.). *Linguistic Diversity and Language Theories*. Amsterdam: John Benjamins, 399–422.
- Dudis, Paul G. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15:2, 223–8.
2007. Types of depiction in ASL. Unpublished manuscript, Gallaudet University.
- Fauconnier, Gilles. 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge MA: MIT Press.
- Fauconnier, Gilles and Mark Turner. 1996. Blending as a central process of grammar. In Adele Goldberg (ed.). *Conceptual Structure, Discourse and Language*. Stanford, CA: Center for the Study of Language and Information, 113–30.
1998. Conceptual integration networks. *Cognitive Science* 22:2, 133–87.
- Fleischman, Suzanne. 1989. Temporal distance: a basic linguistic metaphor. *Studies in Language* 13, 1–50.
- Janzen, Terry, Barbara O'Dea, and Barbara Shaffer. 2000. Passive constructions in ASL. *Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society Proceedings* 434–45.

2001. The construal of events: passives in American Sign Language. *Sign Language Studies* 1:3, 281–310.
- Janzen, Terry and Barbara Shaffer. 2002. Gesture as the substrate in the process of ASL grammaticalization. In Richard Meier, Kearsy Cormier and David Quinto-Pozos (eds.). *Modality and Structure in Signed and Spoken Languages*. Cambridge University Press, 199–223.
2007. The agentive presence in ASL passive constructions. Paper presented at the Second international AFLiCo conference: *Typology, Gesture, and Sign*. Université Lille 3, France, May 10–12.
- Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1, 5–38.
- Lazard, Gilbert. 2001. On the grammaticalization of evidentiality. *Journal of Pragmatics* 33, 358–68.
- Liddell, Scott K. 1995. Real, surrogate, and token space: grammatical consequences in ASL. In Karen Emmorey and Judy Reilly (eds.). *Language, Gesture, and Space*. Hillsdale, NJ: Lawrence Erlbaum, 19–41.
2003. *Grammar, Gesture and Meaning in American Sign Language*. Cambridge University Press.
- Mithun, Marianne. 1986. Evidential diachrony in Northern Iroquoian languages. In Wallace Chafe and Johanna Nichols (eds.). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, 89–112.
- Mushin, Ilana. 2001. Japanese reportive evidentiality and the pragmatics of retelling. *Journal of Pragmatics* 33, 1361–90.
- Palmer, Frank. 1986. *Mood and Modality*. Cambridge University Press.
- Schlichter, Alice. 1986. The origin and deictic nature of Wintu evidentials. In Wallace Chafe and Johanna Nichols (eds.). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, 46–59.
- Shaffer, Barbara. 2004. Information ordering and speaker subjectivity: modality in ASL. *Cognitive Linguistics* 15:2, 175–95.
- Travis, Catherine E. 2006. Dizque: a Colombian evidentiality strategy. *Linguistics* 44:6, 1269–90.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann and Yakov Malkiel (eds.). *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, 245–71.
1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 57, 33–65.
1995. Subjectification in grammaticalization. In Dieter Stein and Susan Wright (eds.). *Subjectivity and subjectivisation*. Cambridge University Press, 31–55.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge University Press.
- Wilcox, Sherman E. 2002. William C. Stokoe and the gestural theory of language origins. In David F. Armstrong, Michael A. Karchmer, and John Vickery Van Cleve (eds.). *The Study of Signed Languages: Essays in Honor of William C. Stokoe*. Washington, DC: Gallaudet University Press, 118–30.
- Wulf, Alyssa and Paul Dudis. 2005. Body partitioning in ASL metaphorical blends. *Sign Language Studies* 5:3, 317–33.

7 Two ways of conceptualizing space

Motivating the use of static and rotated vantage point space in ASL discourse

Terry Janzen

7.1 Introduction

Descriptions of how American Sign Language (ASL) signers use space in their narrative discourse have revealed that there are two distinct ways that spatial scenes are oriented. In more traditional descriptions (e.g. Friedman 1975; Lillo-Martin 1995), signers conceptualize a static scene space when engaged in narrative discourse, such that the interactants and objects in the conceptualized scene are positioned around the signer's articulation space and remain static. The signer can move from vantage point to vantage point around the spaces, which prototypically correspond to human interactants' perspectives on the scene, by "body shifting" towards each interactant's position in the scene space. More recently, however, research has revealed a second type of scene space conceptualized by the narrator in ASL, called "mentally rotated space" in Janzen (2004, 2005, 2008). In this case, the conceptualized space is mentally rotated so that interactants' vantage points in a narrative passage are aligned with that of the signer. The signer does not use body shifts in order to view the conceptualized space from different angles, but instead shifts the space mentally as if it were on a turntable. In the case of mentally rotated space, both signer and addressee must be cognizant of the spatial organization of the conceptualized space from each interactant's vantage point as the signer brings each of these views into alignment with her own view of the imagined space (see Figure 7.1).

Each of these two complex uses of space has consequences for grammatical organization in ASL. For example, when using static space, non-present referents occupy distinct and differential points in space. Reference to any one of these spaces unambiguously evokes that referent. When a scene space is rotated, on the other hand, a referent occupies different locus points in the signer's articulation space, depending on which vantage point is active at a given moment. The identity of the referent intended by a pointing pronominal or eye gaze direction may only be accessible once the vantage point being enacted has been identified. An extreme grammatical effect in this regard is

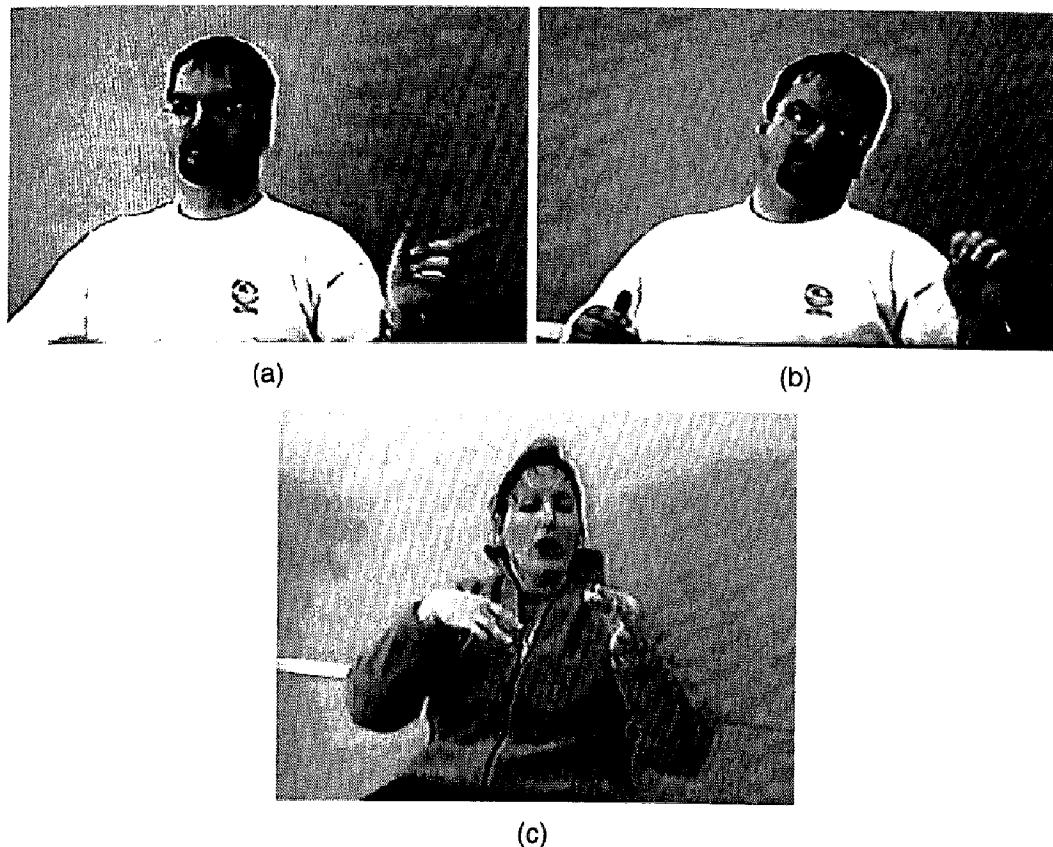


Figure 7.1 In (a) and (b) the signer indicates a "static" space orientation, where he points to a space (a) and uses a body shift (b) towards that space to represent the viewpoint of the referent occupying the space. To represent a different viewpoint, the signer would shift towards a different space. In (c) the signer does not shift physically to another space, but aligns the viewpoint of a referent other than herself (here a police officer) with her own Real Space view on the spatial scene. To represent a different viewpoint, the signer would next realign that viewpoint with her own, again with no physical shift to a different spatial position.

that two references to the same point in the signer's articulation space can variously refer to two distinct referents, depending on which vantage point is active.

The present study seeks to identify when each of the two conceptualized spatial orientations is realized in ASL discourse, and if possible, why. The literature on perspective marking through body shifts in ASL, here referred to as the use of static space, has assumed that this has been a mechanism in narrative discourse, with the exception of Winston (1995), who is not so much concerned with the discourse genre, but who describes body shifts in space when the signer is engaged in a comparative discourse frame. What

emerges in the present work is that body shifts in a static space were not used in narrative discourse. Instead, signers invoked mentally rotated spaces in narratives, and reserved static space use for non-narrative comparative frame discourse.

In what follows I discuss aspects of how space is used in signed language discourse when the signer portrays either her own viewpoint on a scene or the viewpoints of others, and in particular how shifts from one to another take place. We look in detail at how signers mentally rotate a conceptualized space in narratives and how a static space is employed in comparative frame sequences, along with some effects each has on grammatical structures. We then examine an example in which these two spatial orientations interact – that is, when a signer builds a comparative frame space, then inserts a short narrative for which he shifts to a mentally rotated space, and then returns to the static space of the comparative frame once again. This example is significant because it demonstrates the extent to which the two orientations towards space are distinct, both in their occurrence in different discourse genres and in their corresponding differential grammatical features.

7.2 The corpus of data: narratives embedded in conversational ASL

In this study we look at interactive ASL discourse found in a ten-hour corpus of videotaped ASL conversational data, to compare the use of static versus mentally rotated space and the discourse circumstances within which each type is found. The conversations in this corpus contain numerous narrative passages of varying lengths, as is typical of discourse among friends. The ASL signers are all members of the Deaf community in Winnipeg, Canada, but despite some potential for these data to include dialect features of the region, these signers' use of space is thought to be characteristic of ASL generally. The examples of both static and mentally rotated space described and discussed below are taken from several of the texts in the corpus, chosen because they are particularly illustrative, but it should be noted that across the corpus – that is, for all ASL users recorded – the use of space as outlined below was remarkably consistent.

As mentioned, if two such distinct functional uses of space exist, we might wonder what motivates the signer to use one over the other. Some additional questions might be asked as well: Do signers "prefer" one mechanism of using space over the other? And perhaps most importantly, do the two types have differing discourse effects? The present study examines these questions by analyzing spatial features in the ASL corpus mentioned above. Ultimately, this examination illustrates how gesture, conceptualized vantage point, and grammar are interlaced in conversational discourse in a signed language.

7.3 Using space to portray shifting vantage points

In building a static space the signer designates an entity x at some locus a and then shifts physically towards a in her signing, or articulatory, space to portray the view of the scene space that x takes. When more than one entity is designated as such, the signer may shift towards each locus in turn, while typically looking at the other locus (if the two entities are portrayed as interacting). This description of the use of space for perspective shifts in narrative discourse is exemplified by Lillo-Martin (1995: 162) as follows:

- (1) $\langle \text{aMOM } \text{aPOV } \text{1PRONOUN BUSY} \rangle^{\text{ashift}}$
 Mom (from Mom's point of view), I'm busy.
 Mom's like, I'm busy!

In this example the signer shifts to a locus associated with MOM, designated as subscript “a”, and then presents a comment as reported speech. The body shift towards the mom locus in the signer's real articulation space is an overt marker, both for identifying the referent whose perspective is being taken and that the perspective on the scene space is shifting away from that of the signer herself. The potential for a signer to shift perspectives in this manner has been demonstrated not only in ASL, but in numerous signed languages – for example, British Sign Language (Morgan 1999), Danish Sign Language (Engberg-Pedersen 1993, 1995), French Sign Language (Sallandre and Cuxac 2002), and Irish Sign Language (Ó Baoill and Matthews 2000) – which suggests at the very least that in many signed languages, real space blends are reflected in signers' bodily interaction with the actual space that surrounds them.

Even though this use of space – with designating loci around the signer's space as spatial placeholders for referents in the signer's narrative discourse and body shifts towards these loci to align with and portray actions (including linguistic action) from their various vantage points – appears to be pervasive, we cannot conclude that its use, with this function in particular, is universal among signed languages for three reasons. First, perspective marking and perspective shift mechanisms have not been described for the majority of signed languages in use worldwide, so there is at present no way of knowing whether all signed languages use this mechanism to at least some degree. Second, some signed languages appear to use body shifts to located referents to a much lesser extent than has been reported for ASL – for example, in Swedish Sign Language (Ahlgren and Bergman 1994; Nilssen 2008) – and Engberg-Pedersen's (1993) account of perspective marking in Danish Sign Language includes a body shift as only one possibility. Even some descriptions of the use of space in ASL have alluded to perspective shifting without body shifts (Padden 1986; Meier 1990), but such descriptions have not fully explored the details of these observations

nor the conceptual motivations behind them. Third, the findings of Janzen (2004, 2005, 2008) show that in narrative discourse signers consistently use mentally rotated space conceptualizations, and the present study demonstrates that, in contrast, static space – locating referents and body shifting towards them – is reserved for comparative discourse frames.

7.3.1 Real Space and articulation space

Real Space is defined by Liddell (2003) as someone's conceptualization of an actual space in an immediate environment obtained by their perceptions of the space and the items within it. Blends are produced when non-present entities, either imagined or recalled from memory, are mapped onto real space, which then may be interacted with bodily either through gestures as Narayan (this volume) demonstrates, or through linguistic means in signed languages, where articulators – hands and body, for the most part – reach out into actual space. Dudis (2004) illustrates this in his descriptions of complex blends, such as with the motorcyclist in a signed construction. Here the signer conceptualizes a motorcyclist traveling up a hill as mapped onto a Real Space in front of him, with his hands depicted as grasping the handlebars as part of his articulated linguistic structure. This understanding of blended spaces and articulation fits well with the claim in Janzen (2007) that real-world events and objects are never directly mapped onto articulation space, but are always mediated by conceptual space, and thus are readily affected by construal – that is, subjective point of view. This fact is captured in Shaffer's (this volume) study of evidentiality in that the interlocutors in her study of ASL turn from the present discourse space to a past discourse space from which some evidence for a claim is drawn. This conceptualized space is positioned in Real Space to the side away from the addressee, but very little of this past interaction with an interlocutor is overtly specified; thus the signer construes a less-distinguishable “other,” distant in both (past) time and space, that is the source of evidence for the present claim.

7.3.2 Conceptualizing space as mentally rotated space

In a mentally rotated space, the signer portrays others' vantage points on a scene by aligning a conceptualized scene space as someone else might view it with her own view of the space. Instead of moving physically towards a locus where referent x is designated as with the use of static space, the signer realigns the entire scene so that referent x 's view comes into alignment with her own view. Typically, this entails rotating the space 180 degrees, because interactions between people, especially those involving discourse, are more often than not face to face. In a signed language, this is perhaps even more the case because of the visual language medium: signers must look at each other in order to see each other's contributions to the discourse.¹

In one narrative the signer tells of an encounter with police officers and an incident on the highway.² In the story, a police officer motions for everyone to pull off the road onto the shoulder because of an incident taking place further up the road. The driver of the vehicle that the narrator was riding in, her mother, complies, as do other vehicles. The people in the signer's car and the police officer standing on the road ahead face one another, and their exchange – actions and reactions – is recounted in the narrative. Here, as in the narrative passages throughout the corpus, the signer mentally rotates the scene space from the vantage point of someone in the car – usually the mother, but occasionally the signer herself³ – to the vantage point of the police officer. Choosing a mentally rotated space orientation over static space has numerous consequences in terms of the grammatical features that appear in signers' discourse. If the signer were to choose to use a static space, then reference tracking, which has been extensively reported on (e.g. Friedman 1975; Padden 1986; Aarons *et al.* 1994; Lillo-Martin 1995), would involve the positioning of referents at various locations in the signer's articulation space and then using discrete pointing, body shifts, eye gaze, and so on (see Winston 1995), towards these designated spaces, with the idea that only one entity can occupy a given space at any time. For example, an entity positioned in a space to the signer's left remains cognitively salient at that location; the signer can further refer to it by pointing, eye gaze, and so on, from her own point of view, or, supposing that she also positions a referent in a rightward space, she may body shift towards that rightward space and refer to the same exact leftward space to enact the rightward-placed referent's viewpoint. And as already described, a leftward body shift evokes that leftward-placed referent; once evoked, if the signer turns to view the scene space, we understand that the leftward-placed referent's vantage point is also being enacted. But for each of these options, the conceptualized leftward-designated entity remains in that space.

Contrary to this, when the signer chooses mentally rotated space, each vantage point comes into line with the signer's stance, so that portraying a leftward-designated entity's perspective on the scene (from the vantage point of the signer) brings that entity's view to the signer's central view – that is, the signer does not body shift leftward to physically reposition herself at that entity's vantage point, but takes on that entity's vantage point as if it were her own. The effect of this is that all other entities associated with the scene also shift in the conceptualized space relative to whose view is being portrayed. To illustrate from the narrative described above, when the mother, depicted as holding onto the steering wheel, looks down the road ahead to see the police officer standing there, the signer fixes her gaze at a point in the space in front of her that corresponds to the relative location of the police officer in the conceptual scene. When she enacts the officer's gesture that tells them to pull off the road, she looks at *the exact same point in space*, which now corresponds

to the vehicle she is riding in, from the vantage point of the officer who is looking down the road at them.

7.3.2.1 Multiple articulatory spaces associated with a single referent

So far, we have described two quite different ways that an ASL signer might portray shifts in perspective-taking. Most described in the literature is what we refer to as the use of static space – that is, where a scene space is set and remains in place while the signer aligns with the vantage points of other referents in the static space by shifting her body towards the loci that these referents occupy. Analysis of the narrative passages in the present study, however, reveals an alternate approach to shifting perspective, in which the conceptualized scene space does not remain static, but is rotatable, such that rather than the signer moving in articulation space to align with the vantage points of referents positioned around a static space, the conceptualized scene space is rotated to bring third person referents' viewpoints in line with the signer's view of the scene. This has several important consequences for referent tracking, the grammatical elements associated with each spatial orientation and usage, and even, it turns out, for the phonological structure of complex predicate forms. Another significant finding in this study is that contrary to what has been reported previously for ASL, static space orientation, along with representing referents around articulation space and pointing to them and body shifting towards them to take their vantage point, does not occur in narrative passages. Instead, mentally rotated spaces are employed, and the use of static space and its associated mechanisms for third person referencing are reserved for the domain of comparative frames, which are not found within narratives.

Because the two spatial orientations described above are mental constructs and thus invisible to the observer, a question that might be asked is how we know when one or the other – static space or mentally rotated space – is being employed by the signer. The answer is twofold: we can look at the phonological and grammatical features of the discourse and, ultimately, as this study shows, we can look to the discourse genre of the signer. In this section we examine some of the phonological and grammatical elements of mentally rotated space usage.

In the passage described above, the driver of the vehicle and her passengers (including the signer) are face to face with a police officer standing on the road at some distance ahead of them. A striking sequence is signed in two contiguous constructions. In the first, the signer says that the police officer motions for them to pull onto the shoulder of the highway, and in the second, she says that their, and every other, vehicle complies and pulls over. This is shown in Figure 7.2.

The action shown in Figure 7.2(a) is of the police officer motioning with a sweeping leftward movement, which is clearly articulated from his vantage

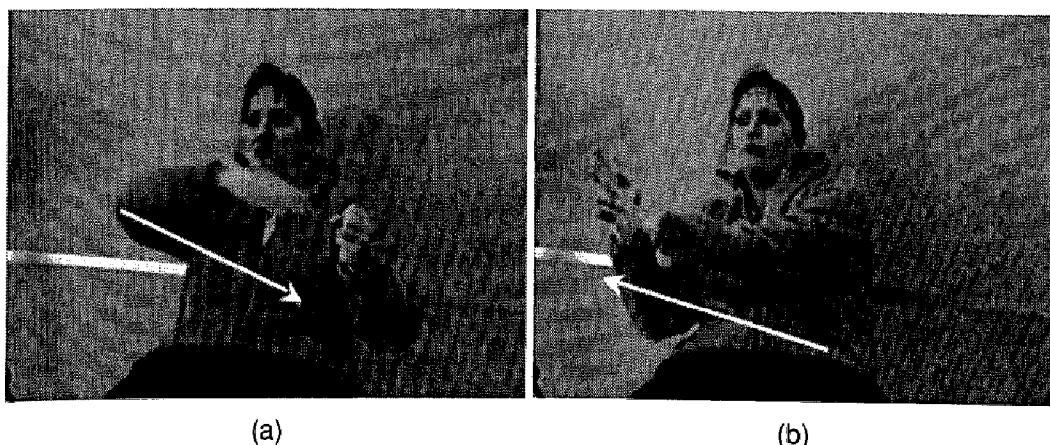


Figure 7.2 Depicting the police officer's action in (a): a leftward-moving gesture; and the signer's vehicle location in (b): a rightward-moving hand shape.

point, as he would be facing the signer's vehicle, indicating for the car to pull over to *his* left. The construction is actually made more complex, however, by the inclusion of body partitioning (Dudis 2004), in this case taking the form of the signer's face and right hand representing the police officer, and the left hand partitioned off to represent the signer's mother holding onto the steering wheel. Complex partitioned constructions such as this have not been adequately described or explicated for ASL as yet, and it is beyond the scope of this chapter to do so here, but for our purposes I might briefly suggest a solution. In the utterance just prior, the signer has said that her mother was driving along, but slowed down as they approached the officer on the road. Body partitioning in signed language articulation allows for what are termed "simultaneous constructions" (see the works in Vermeerbergen *et al.* 2007), where two distinct elements are articulated simultaneously, in part because the articulation system (two hands, face, body) can visually represent more than one element at the same time, within certain cognitive limits. It is unusual for the signer to be able to articulate simultaneously two entirely different actions undertaken by two distinct actors, although, again, within certain limitations it might be possible, but more often when such actions are conceptualized as occurring simultaneously, articulation in a signed language affords the potential, at least, for iconically representing the two in some simultaneous way. However, it appears to be the case that one or the other action is profiled in a given clause at the expense of the other, which is backgrounded, even though some morphemic representation of the action can still be given overtly in the simultaneous construction clause. In this case, the focus of the previous clause had been the driver, but in the next clause the focus shifts to the officer's action

and the driving is backgrounded. A remnant of this is retained by the left (non-dominant) hand, sometimes called a “buoy” (Liddell *et al.* 2007), as a kind of placeholder for the previous action, while the profiled action is articulated with the signer’s right (dominant) hand. Here we are not concerned with the buoy, but focus our attention on the phonological features of the motion articulated by the signer’s dominant hand.

As noted, the motion moved in a leftward direction, as shown in Figure 7.2(a). This can only be interpreted as portrayed from the vantage point of the police officer, since in facing the oncoming traffic he would want them to move onto the shoulder off their lane of traffic. They are on the left side of the highway coming towards him, so this would be off to his extreme left.

They comply, pulling off towards the right (Fig. 7.2b). In this clause, the signer is evidently viewing the same conceptualized space now from the mother’s vantage point, so that what was a leftward-indicating movement from the officer’s viewpoint is now a rightward action on her part: she pulls over correctly onto the right shoulder of the highway. Significant is that one entity, the vehicle the signer is riding in, occupies two spatial locations in the articulation space of the narrative, but both of these locations are phonological manifestations of a single conceptual Real Space location. The only explanation possible is that she has rotated her conceptualized space mentally from that of the officer to that of the driver. This utterance, composed entirely of a verb complex that indicates both direction of movement and a nominal feature that represents a string of entities (in this case interpretable as vehicles), has no lexical or pronominal words to identify the referent whose viewpoint is being taken. Instead, the signer relies on spatial means to identify the shifted view on the scene.

Thus we see a significant spatial effect of using either static space or mentally rotated space. In the use of static space, entities positioned at some locus in the articulation space remain in that space, unless of course they are depicted as traveling to a different location that is then represented by a different point in space, or the discourse segment regarding that entity has ended. But if, in a static space orientation, a signer wishes to refer to that located entity from any available vantage point around the space – that is, from her own vantage point or that of another third person referent located at some other spatial position – that entity occupies just one *locus* in the entire space relative to all others.

However, in the use of mentally rotated space, the conceptualized positions of entities are dependent on the overall conceptualized view of the scene. Thus if from one vantage point an entity is positioned in a leftward space, then from the vantage point of a referent just to the right of the signer, which is brought in line with the signer’s actual view on the surrounding space, that entity location is now further leftward because there is a greater relative distance between

this new viewer and the entity being viewed. The extreme case is a 180-degree rotation, where originally an entity occupies a leftward space, but once the scene is rotated to the vantage point of someone directly across from the signer, the entire scene shifts relative to the new vantage point, and the entity being viewed is now in a rightward space. Thus the grammar of ASL, which has previously not allowed a referent to occupy two spatial locations simultaneously, must be expanded to accommodate this possibility when vantage point shifts are evoked through mentally rotated space.

The idea of simultaneity here is a complex one, which warrants some comment. It has long been understood that signers can position a non-present or abstract entity at some locus in their articulatory space such that the space (a Real Space) anchors subsequent reference to that entity. That referent is evoked through numerous means: pointing to the space, gazing towards the space, and so on; even positioning the hands in the space can imply a connection to that referent. To succeed, these referencing mechanisms must be in keeping with the referent's accessibility (see Vandelanotte, this volume) – in other words, it must be distinguishable from other entities that will also be positioned around the signer's articulation space. Based on this, it has been understood that the signer cannot arbitrarily reposition the referent while continuing to treat it as a highly accessible entity through pronominal points or eye gaze, although note van Hoek (1992), who suggests that if a referent who had been in one location in the narrative, say Los Angeles, relocates to New York, the signer may refer to this single referent subsequently in either of the two spaces, depending on whether she is talking about her when she was in each of the two cities. But in the present context, the referents are (more or less) stationary within a single event. Thus when the signer takes one viewpoint, a referent may be positioned to her right, but when she takes another viewpoint by rotating the scene mentally, that (rotated) referent is now in her leftward field. She can alternate between the two viewpoints in her discourse, meaning that she will alternate referring to the two Real Space locations depending on which viewpoint she is taking, but importantly, throughout the discourse stretch both the signer and the addressee must maintain a mental representation of the referent + Real Space in two spaces simultaneously.

In terms of overt markers of viewpoint shift, an overt shift marker in a static space orientation appears in the construction – for example, a body shift towards the referent location.⁴ But when mentally rotated space is employed there is no corresponding overt shift marker because the rotation is mental; the signer must use other means to enable the addressee to track referents around the scene space.

Briefly, in the corpus analyzed here, there were three mechanisms used by signers to do this. One mechanism was apparently to rely on interlocutors' knowledge of ASL structure. That is, native or native-like ASL users would

likely be quite familiar with linguistic characteristics of the narrative genre in the language, so would interpret seeming discrepancies, such as two distinct spatial positions for a single entity, as corresponding to the same entity but from different vantage points. In other words, the spatial consequences of mentally rotated shifts would add up to correct interpretations about which vantage point was at play at any given moment in the story. Second, at times in the narrative passages, it seemed as if the signer was unduly naming referents with full Noun Phrases (NPs) when the referents were already clearly topical. Note that above it was suggested, following Slobin (2006), that topical nominals are not typically overt in ASL clauses, but in at least some of these narrative passages such NPs occurred frequently – for example, in a sequence such as *POLICE, action of police* (which has a kind of “topic, comment” feel). I would suggest that this is a pragmatic choice on the part of the signer, who may feel that information must be explicit at that point in order for the viewpoint to be understood. Interestingly, NPs were used and not pronouns, likely for reasons akin to the use of two spaces for a single referent, which may parallel Vandelanotte’s (this volume) “distancing indirect speech/thought” (DIST) type of speech and thought representation, where the speaker is more apt to use full NPs to refer to referents when she aligns other discourse viewpoints with her own. If, from one vantage point, the signer pointed (as a pronominal) at the spatial location associated with the referent she was face to face with in the scene space, then rotated the space, and then, from the second referent’s vantage point, used a pronominal point towards the space associated with the other referent, these two indices would point towards the same spatial location in articulation space, but would be intended to refer to two different referents. This may work once the vantage point is clear, but not to indicate whose vantage point is going to be taken next.

Third, occasionally, some stance feature of a referent may be invoked. For example, at one point after the police officer had pulled the oncoming vehicles off the road, the signer’s mother, at the wheel, was depicted as slouched in her seat and leaning on her elbow, propped up against the driver’s door, with the window rolled down. Several times the signer re-enacted this stance before telling something the mother either said or did, and each time the return to the mother’s vantage point was abundantly clear.

7.4 The role of static space use in discourse

As mentioned, mentally rotated space, and not static space, is used to identify perspective shifting in the narratives embedded in the conversations in the corpus. While it may be premature to conclude that static space is not used in ASL for this function, the signers in these data consistently use static space and mentally rotated space for two different purposes. Static space is used when a

signer compares the attributes or actions of two referents, but only when these referents are not interacting with one another as story characters in a narrative passage. In other words, when the signer compares entity *x* with entity *y*, an arbitrary space for *x* could be chosen to the signer's right, and a second arbitrary space for *y* could be chosen to the left. This is significant in that when the signer does this, he or she is not then situating the referents in any conceptualization of an existing space, but abstractly, because a list of the referents' comparative attributes does not occur in any existing space. In this case, when the signer uses a body shift towards referent *x* in rightward space *a*, the perspective remains the signer's own, and *x*'s perspective on a scene space is not taken up. Therefore, it is clear that static space and body shifts are reserved for comparative purposes in the signer's present, whereas mentally rotated space is used for vantage point shifting, prototypically in the narrative past.

7.5 The interaction of static and mentally rotated spaces

The use of space in signers' discourse is decidedly complex. This complexity has been illustrated by the example above of the police officer and the signer's vehicle, in that two different loci in articulation space are used to designate a single referent within the same discourse frame. The signer's vehicle is represented by a nominal hand shape moved to a rightward position when the vantage point is that of the signer herself along with others as characters in the story being told; but the same signer's vehicle in the very same scene is the object of the police officer's leftward gesturing. Thus, he motions for them to move over to *his left* and they comply, moving over to *their right*. The only way that this can be made coherent is to consider that the portrayal of vantage point determines how space is used, and when this is understood the seemingly conflicting use of that space is not in fact incoherent after all.

But in another signer's discourse from the corpus, a further level of spatial complexity is apparent. During one conversation, the signer compares characteristics of his boss's approach to his work with his own brother's approach – both are in the same line of work but have different businesses. To do this, he positions his boss in a space to his right and his brother in a space to his left. The signer then refers to these two spaces by pointing and body shifts toward the spaces, consistent with Winston's (1995) description of referencing in comparative spaces. The clarity of these two positioned referents is shown in Figure 7.3.

In this instance, the signer does not use body shifts towards each space for the purpose of portraying the vantage point of either referent on a scene, but rather to list and describe attributes of each while maintaining the role of "informer" in the discourse. The signer's eye gaze during these body shifts is primarily directed towards the addressee.⁵ As the signer compares attributes, he shifts

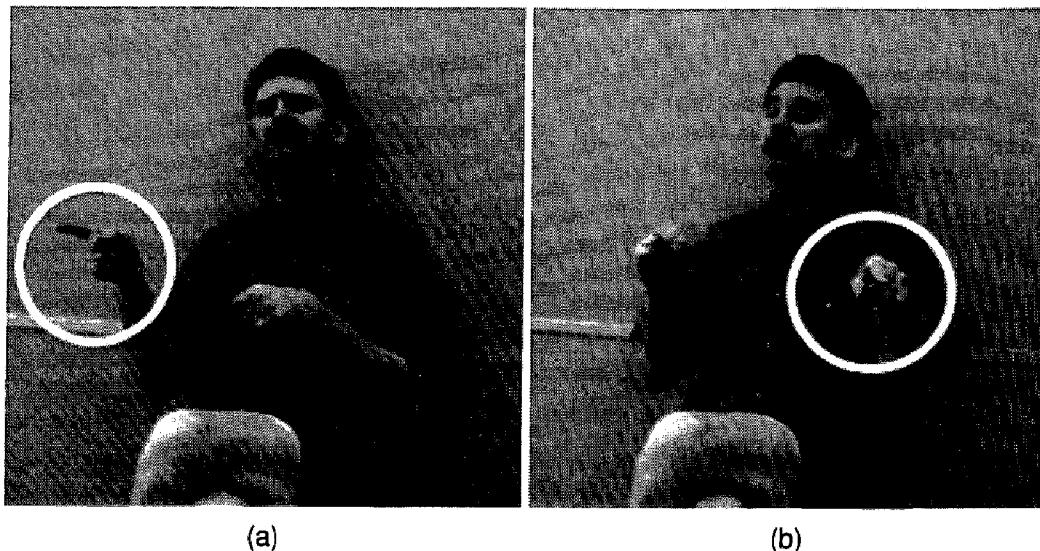


Figure 7.3 Positioning the boss in a rightward space in (a); positioning the brother in a leftward space in (b).

rightward and leftward towards each space alternately. Once again, the two referents being compared are not conceptualized as located in two real-world spaces relative to one another; rather they occupy two abstract spaces in the signer's discourse world.

However, at one point in the conversation the signer shifts from the comparative frame to a narrative description of a particular interaction he had with his boss, to illustrate an attribute he has said that his boss possesses. At this moment, in enacting his boss's actions and speech during that event, the signer moves from a rightward body shift (when describing the attribute) back to a central position, portraying the boss's and his own interaction with mentally rotated space. In this narrative sequence, he and his boss are understood to be facing one another, with the effect, similar to the interchange described earlier between the driver and the officer, of directing the discourse of each interactant in the scene's discourse outwardly towards the other person. This interchange in the narrative passage takes place entirely in the central area of the signer's articulation space. At the end of this short narrative, the signer moves once again to shifting rightward and leftward to continue his attribute description.

This sequence significantly illustrates that real-world space and non-real-world space are conceptualized differently for the signer. When linguistic expression is itself visual and spatial, as is the case for signed languages, there is remarkable potential for features of conceptualized spaces to map iconically onto articulation space.⁶ As this signer shows, the conceptual distinction between real-world and non-real-world spaces is expressed overtly and, I argue, this expression has become entrenched as grammar in ASL.

Further, this example illustrates the complex nature of referential spatial marking and, subsequently, reference tracking in ASL. When the signer begins the narrative he abandons the rightward and leftward spaces he has designated for the two referents in the comparative frame space. When he speaks to his boss, he directs this discourse towards a spatial area in a more central space, and when he enacts his boss's discourse he rotates the conceptual space so that this discourse is equally central. The rightward and leftward comparative space loci remain active, however, because the signer reverts back to them once the narrative sequence is completed, without needing to re-establish the referent identities.

In fact, at one point in the narrative the signer demonstrates the availability of the comparative frame space. As discussed above, referent identification during the use of mentally rotated space must still occur, whether implicitly or explicitly, as necessitated by discourse pragmatics. How does this signer identify the boss's action (including speech actions) in the narrative sequence? In one instance, at the beginning of an utterance, he points to the rightward location from the comparative frame space and then indicates what his boss said. This is significant because the boss already has a distinct spatial representation in the overall complex articulation space, such that the signer is able to reach into that space and into that frame with an identifying pronoun. But note that within the narrative frame, the boss does not occupy that rightward space, so the articulated pronoun is quite abstract.

7.6 Discussion

Overwhelmingly, the conversational discourse in the present study shows that while signers use both static space and mentally rotated space, the two uses of space serve quite different functions. What has been widely reported in the literature on perspective and perspective shifting in signed languages is that other referents' perspectives are expressed through body shifting towards the loci in the signer's articulation space where these referents have been positioned – in other words, through the use of static space. The data in this study do not support this belief. Here, shifts to other referents' vantage points are expressed through mentally rotated space consistently, while the use of static space is reserved for comparing the characteristics of referents who occupy spaces in an abstract sense. Winston's (1995) description of comparative discourse frames has already shown that the characteristics of static space use occur in that context, but the present study clarifies the extent to which these characteristics distinguish comparative discourse from narratives, which do not use space in the same way, even though previous accounts of perspective shifts in ASL narrative passages have assumed that they are present in that discourse genre as well.

One interesting effect that this differential use of space has is that in mentally rotated space referents can occupy more than one region of articulation space in a single discourse event. It may not be the case that two spaces can be occupied by a single referent simultaneously in the sense of “simultaneous constructions” (Vermeerbergen *et al.* 2007), where two items are articulated by each of the signer’s two hands – for example, if the signer pointed to two spaces simultaneously to represent the two locations that a single entity occupied – but it must be true that the signer’s mental representation of that referent allows for at least two spatial loci representations concurrently, even though it is likely that only one would be in focus at any given point in time. This means that we are either viewing the referent from one vantage point or the other, but not from both vantage points simultaneously. This is evident in both examples discussed above. In the narrative involving the police officer, the signer’s vehicle was indicated both to her left and to her right, depending on whether the perspective was that of the officer or the passengers in the vehicle. She alternates between these two vantage points, with the result that both the signer and addressee must maintain a mental representation of a single referent associated with two different regions of space concurrently, as both appear and reappear during the discourse segment. The example of the signer’s boss requires similar reference tracking, except that in this case one spatial region is more abstract and does not assume that referent’s vantage point on any scene.

The fact that two spaces may be used to locate a single referent has been discussed in van Hoek (1992). Van Hoek suggests that this may occur if the referent is being referred to at two different periods of time, or has occupied two different real-world locations, or if there are two “versions” of the same referent – for example, the person herself and a picture of the person – but the present study demonstrates that in addition to van Hoek’s categories, shifting vantage points on the same referent in a single setting and time also produce this effect in the grammar.

7.6.1 *Conceptualized space and abstract space mapping*

A final observation concerns the designation of spaces that coincide with two levels of abstraction in terms of referent location and articulation space. On the one hand, the positioning of referents in the comparative frame appears relatively concrete. These spaces are fixed, stable, and relational. They are fixed in the sense that the space is reserved for that referent and nothing else; stable because numerous spatial indicating mechanisms may be employed to evoke the referent over a long stretch of discourse; and relational because each spatial region is maximally distinct from the other, being on opposite sides of the signer’s articulation space, and this spatial distinction is itself meaningful. It is

interesting then that these spaces in articulation space are employed to represent abstract referents. The referents in comparative frames are not conceptualized as occupying or operating in some actual past or present space, but rather their attributes or characteristics are being discussed, and this “view” of these referents is not a view of physical space.

On the other hand, referents within a mentally rotated space description are not quite so spatially bound. They may be described as oriented in space, but they may not. In the highway narrative, the passengers in the vehicle see a police officer up ahead on the road, but as a referent, he is never designated as occupying a particular locus in the signer’s articulation space. The mother is depicted as looking ahead down a conceptualized highway (that is never articulated overtly, but is interpretable from the scene), and the police officer – the object of their gaze – is identified with the noun POLICE followed by his action, but the addressee must construct the relational space; this is not done explicitly by the signer. Thus in these two discourse contexts, the more abstract elements are given clear spatial designations in articulation space, whereas physical referents from an actual (although filtered through conceptualization) past event are not overtly localized in space.

7.7 Conclusion

While the discussion above concerns two ways that the ASL signer orients her discourse in space, whether as a static or mentally rotated spatial orientation, it is evident that these two means are illustrative of a more general principle that language users are guided by an embodied cognitive viewpoint. Narayan (this volume) agrees with this, but suggests that for spoken language users, this viewpoint surfaces in their gestures because their spoken discourse does not provide a medium for such overt viewpoint expression. She makes the point, then, that gesturers cannot gesture iconically without portraying an inherent viewpoint on the scene. My position, based on the type of data described above, is that when language expression is articulated by the hands, face and upper body as a whole, as is the case for signed languages, the differences between gesture and linguistic material fall away, resulting in an embodied viewpoint surfacing in language structure. In other words, when the medium of language is conducive to the expression of viewpoint, that viewpoint will surface. And further, because the signer’s language expression is via a body in space and the signer *is* a body in space, expressed viewpoint is pervasive, entrenched, linguistic, and grammaticalized.

A remaining question is: why have past analyses of perspective in ASL consistently concluded that body shifts in static space function to mark perspective shifts? The answer to this may necessarily be left to further analysis,

but it is worthy of mention that ASL signers themselves often describe perspective shifts as requiring physical body shifts in static space. It is possible, however, that perspective shifts utilizing body shifts in space occur in contexts where a signer may not have been able to make herself clear via mentally rotated space to an addressee, and thus she employs this strategy to clarify. It is the case that the identity of the referent in a mentally rotated space is not made overt by simply referring to a designated locus in the signer's articulation space, so perhaps reference tracking is compromised by a miscalculation of what is identifiable or assumed to be shared information. The suggestion might also be made (Rick Zimmer, personal communication) that body shifts in a static space may be used in this regard when the addressee is a second language user of ASL; thus the signer is looking for a means that is more overt to identify the referent in a scene and subsequently to show their perspective.

Further research may explore some of these elements of interaction and language use, and a larger corpus of more varied text types may also elucidate the extent to which the conclusions found in the present study hold. Nonetheless, the ASL signers in this corpus give us some clear insights into the complexity of conceptualized scenes and reference tracking, and how a signed language like ASL encodes these complexities in observable constructions.

References

- Aarons, Debra, Benjamin Bahan, Judy Kegl, and Carol Neidle. 1994. Subjects and agreement in American Sign Language. In Inger Ahlgren, Brita Bergman, and Mary Brennan (eds.). *Perspectives on Sign Language Structure. Papers from the Fifth International Symposium on Sign Language Research. Vol. I.* 2 vols. Durham: The International Sign Language Association, 13–27.
- Ahlgren, Inger and Brita Bergman. 1994. Reference in narratives. In Inger Ahlgren, Brita Bergman, and Mary Brennan (eds.). *Perspectives on Sign Language Structure. Papers from the Fifth International Symposium on Sign Language Research. Vol. I.* 2 vols. Durham: The International Sign Language Association, 29–36.
- Dudis, Paul G. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15:2, 223–38.
- Engberg-Pedersen, Elisabeth. 1993. *Space in Danish Sign Language: The Semantics and Morphosyntax of the Use of Space in a Visual Language*. Hamburg: Signum Verlag.
1995. Point of view expressed through shifters. In Karen Emmorey and Judy S. Reilly (eds.). *Language, Gesture, and Space*. Hillsdale, NJ: Lawrence Erlbaum, 133–54.
- Friedman, Lynn. 1975. Space, time, and person reference in American Sign Language. *Language* 51, 940–61.
- Janzen, Terry. 1998. Topicality in ASL: information ordering, constituent structure, and the function of topic marking. Ph.D. dissertation, University of New Mexico, Albuquerque.

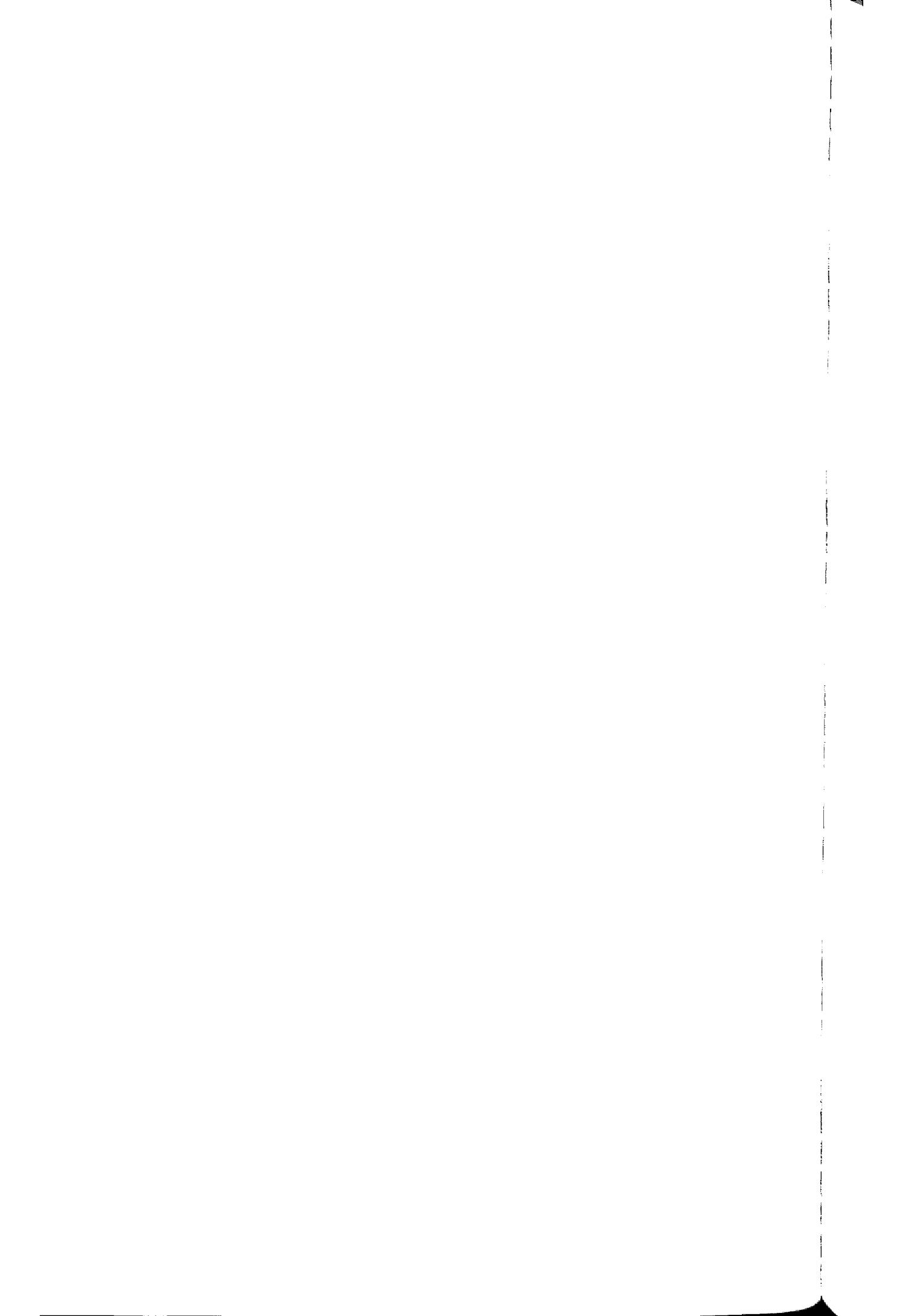
2004. Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics* 15:2, 149–74.
2005. *Perspective Shift Reflected in the Signer's Use of Space*. CDS/CLCS Monograph No. 1. Dublin: Centre for Deaf Studies, University of Dublin, Trinity College.
2007. The expression of grammatical categories in signed languages. In Elena Pizzuto, Paola Pietrandrea, and Raffaele Simone (eds.). *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*. Berlin/New York: Mouton de Gruyter, 171–97.
2008. Perspective shifts in ASL narratives: the problem of clause structure. In Andrea Tyler, Yiyong Kim, and Mari Takada (eds.). *Language in the Context of Use: Usage-Based Approaches to Language and Language Learning*. Berlin/New York: Mouton de Gruyter, 129–54.
- Liddell, Scott K. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press.
- Liddell, Scott K., Marit Vogt-Svendsen, and Brita Bergman. 2007. A crosslinguistic study of buoys: evidence from American, Norwegian, and Swedish Sign Language. In Vermeerbergen *et al.*, 187–215.
- Lillo-Martin, Diane. 1995. The point of view predicate in American Sign Language. In Karen Emmorey and Judy S. Reilly (eds.). *Language, Gesture, and Space*. Hillsdale, NJ: Lawrence Erlbaum, 155–70.
- Meier, Richard P. 1990. Person deixis in American Sign Language. In Susan D. Fischer and Patricia Siple (eds.). *Theoretical Issues in Sign Language Research. Vol. I: Linguistics*. 2 vols. University of Chicago Press, 175–89.
- Morgan, Gary. 1999. Event packaging in British Sign Language discourse. In Elizabeth Winston (ed.). *Storytelling and Conversation: Discourse in Deaf Communities*. Washington, DC: Gallaudet University Press, 27–58.
- Nilsson, Anna-Lena. 2008. *Spatial Strategies in Descriptive Discourse: Use of Signing Space in Swedish Sign Language*. CDS/SLSCS Monograph No. 2. Dublin: Centre for Deaf Studies, University of Dublin, Trinity College.
- Ó Baoill, D. P. and Patrick A. Matthews. 2000. *The Irish Deaf Community. Vol. II: The Structure of Irish Sign Language*. 2 vols. Dublin: The Linguistics Institute of Ireland.
- Padden, Carol. 1986. Verbs and role-shifting in American Sign Language. In Carol Padden (ed.). *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching*. Sliver Spring, MD: National Association of the Deaf, 44–57.
- Sallandre, Marie-Anne and Christian Cuxac. 2002. Iconicity in sign language: a theoretical and methodological point of view. In I. Wachsmuth and T. Sowa (eds.). *Gesture and Sign Language in Human–Computer Interaction. Proceedings of the International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001*. Berlin: Springer, 173–80.
- Slobin, Dan I. 2006. Issues of linguistic typology in the study of sign language development of deaf children. In Brenda Schick, Marc Marschark, and Patricia Elizabeth Spencer (eds.). *Advances in the Sign Language Development of Deaf Children*. Oxford University Press, 20–45.
- van Hoek, Karen. 1992. Conceptual spaces and pronominal reference in American Sign Language. *Nordic Journal of Linguistics* 15, 183–99.

Vermeerbergen, Myriam, Lorraine Leeson, and Onno Crasborn (eds.). 2007. *Simultaneity in Signed Languages: Form and Function*. Amsterdam/Philadelphia: John Benjamins.

Winston, Elizabeth A. 1995. Spatial mapping in comparative discourse frames. In Karen Emmorey and Judy S. Reilly (eds.). *Language, Gesture, and Space*. Hillsdale NJ: Lawrence Erlbaum, 87–114.

Part IV

Constructions and discourse



8 The constructional underpinnings of viewpoint blends

The *Past + now* in language and literature

Kiki Nikiforidou

8.1 Introduction

The co-occurrence of a past tense verb with a proximal temporal deictic has been identified as one of the formal features of free indirect style (FIS)¹ – that is, the type of discourse where there is a “transference of subjectivity from the discourse agent to the discourse referent” (Adamson 1995: 197). Within FIS, the *Past + now* serves to present events from the point of view of a “self” or a consciousness (other than the narrator) that is contextually available and prominent, as that character’s thoughts, speech, or perceptions. As illustrated by the corpus examples, in non-literary contexts as well, *Past + now* marks a shift in perspective from the narrator to a vantage point close to or inside the narrated event(s).

In this chapter, I adopt a constructional approach (Fillmore *et al.* 1988; Kay and Fillmore 1999), arguing that it can accurately pinpoint the syntactic and semantic source of the viewpoint effect associated with the pattern in all its manifestations. From a conceptual integration perspective, the constructional analysis proposed here serves as a detailed blueprint of the mappings that give rise to a particular kind of blended space; more precisely, I shall argue that constructionally triggered coercion (resolving the interpretational conflict in the *Past + now* pattern) can be rendered, in blending terms, as a specific type of compression relation in the blended space that represents the shifted interpretation.

Both literary and non-literary uses occur in a narrative, monologic (i.e. non-conversational) text type (Fillmore 1981; Banfield 1982). In a constructional framework, such discoursal properties (including textual and register characteristics) conventionally associated with a form are incorporated into the meaning pole of the corresponding construction, alongside purely semantic information (Goldberg 1995: 7; Fried and Östman 2004). “Meaning” in Construction Grammar thus stands for “all the *conventionalized* aspects of a construction’s

function, which may include not only properties of the situation described by the utterance, but also properties of the discourse in which the utterance is found . . . and of the pragmatic situation of the interlocutors" (Croft and Cruse 2004: 258).

In all versions of Construction Grammar,² constructions, as learned pairings of meaning and form of a non-derivational nature, are taken to pertain to all levels of grammatical analysis, from morphemes to phrasal patterns; the network of constructions "captures our grammatical knowledge of language *in toto*, i.e. it's constructions all the way down" (Goldberg 2006: 18). Although most construction-based analyses have focused on sentence-level phenomena (at best encompassing bi-clausal constructions such as conditionals – for example, Fillmore 1990; Dancygier and Sweetser 2005), the need to extend construction grammar to larger pieces of discourse ("all the way up," we might say) has been noted in the literature. Östman (2005: 125), for instance, suggests that certain discourse patterns represent conventionalizations of specific linguistic properties, which place them on an equal footing with the conventionalized patterns known as "grammar." In Östman's terms, a discourse construction specifically represents an association of a particular text type (such as argumentative, descriptive, narrative) with a particular genre (for example, recipes, obituaries, fairy tales).

Returning to FIS, in addition to the *Past + now* pattern (the "was – now paradox" in Adamson's [1995] terms), other formal features associated with this kind of narrative include the following:

- a) Constructions such as inversion, exclamative sentences, topicalization, adverb-preposing and right dislocation – in general, phenomena normally associated with main/non-embedded clauses (e.g. *The way to the Regent's Park Tube station – could they tell her the way to Regent's Park Tube station – Maisie Johnson wanted to know* [V. Woolf, *Mrs Dalloway*]).
- b) Parentheticals with verbs of communication or mental state (e.g. *Only wool gathering, she protested, flushing a little.* [V. Woolf, *To the Lighthouse*]).
- c) Definite articles and demonstratives with no referent in the preceding discourse, and also pronouns, even reflexive ones, without an antecedent (cf Brinton 1995) (e.g. *Within himself his will was coiled like a beast, hidden under the darkness* [D. H. Lawrence, *The Rainbow*]).

In the present framework, all these formal realizations of FIS³ would be treated as different constructions (or families of constructions), which in a narrative may share a semantic-pragmatic value – that is, the shifted viewpoint. The overarching requirement for a narrative context (or "text type") and the actual coexistence of all of these features in some genres (e.g. the nineteenth-century novel) might even uphold the suggestion that FIS has grammaticalized into a large-size construction, consisting of a text type that comprises all the configurations above and is associated with specific literary genres. Here,

however, arguments are restricted to the constructional properties of the *Past + now* pattern.

The interpretation of *Past + now*, as suggested above, may be seen as a result of coercion, technically defined as the resolution of conflict between constructional and lexical meaning (Michaelis 2005a, 2005b). In blending terms now, resolution of (apparent) conflict is often achieved through the mechanism of compression, whereby elements that are conceptually separate in the input spaces are construed as one in the blended space. The construction at hand cues a particular kind of compression, namely compression of a time relation. The dynamic, continuously updated character of such blending networks renders them particularly suitable for representing meaning in a narrative, where formal clues may often give conflicting instructions even within the same sentence (as is the case with FIS). The fact that, in *Past + now* and in FIS, as a whole, these instructions relate precisely to viewpoint, which has been defined as a feature of mental spaces,⁴ further justifies the semantic anatomy of this construction in terms of blending. A blending analysis serves to embed FIS in a paradigm of speech and thought representation modes, highlighting differences and similarities in a principled way. As shown, for instance, in Vandelanotte (this volume), FIS shares with the direct and indirect modes the existence of two deictic centers, but only in FIS and in the direct mode is viewpoint located in the space of the reported clause, thus distanced from the current (narrator) space.

The perhaps obvious point is that a theory like Construction Grammar, which focuses on the subtleties of the relationship between meaning and form, can and should feed directly into a theory such as blending, which produces fine-grained representations of semantic content. The constructional ramifications of mental space networks involving viewpoint and stance should be addressed explicitly (see also Dancygier, this volume), since particular space configurations would receive empirical support from the (possibly) systematic reflexes of space structure into linguistic form. The present chapter aims to contribute in this direction.

8.2 The *Past + now* in non-literary use⁵

All varieties of FIS (i.e. represented speech, thought, or perception, e.g. Examples (1), (2) and (3) respectively) are uniformly marked by the coexistence of past tense verbs with present time deictics. Despite the differences among individual analyses of FIS, researchers agree that this pattern is one of the formal features of the style, correlating with an interpretation in which events (whether speech, or thought, or perception events) are presented as experienced rather than reported (Banfield 1982; Adamson 1995; Brinton 1995; Wright 1995; Bosseaux 2004).

- (1) Mr Woodhouse . . . commended her very much for thinking of sending for Perry, and only regretted that she had not done it . . . It was a pity, **perhaps**, that he had not come last night; for, though the child **seemed** well **now** – very well considering – it would probably have been better if Perry had seen it.

(Jane Austen, *Emma*, cited in Adamson 1995)

- (2) **Today** she **did not want** him.

(D. H. Lawrence, *The First Lady Chatterley*, cited in Banfield 1982)

- (3) They **now saw**, tied to the fence, Ratliff's buckboard and team.

(William Faulkner, *The Hamlet*, cited in Banfield 1982)

A corpus search (British National Corpus) reveals, however, that far from being restricted to FIS, the pattern occurs in a wide variety of non-literary texts as well. In summary (see note 5), for combinations of selected verbs with *now*, the search showed that although their number is higher in literary texts, they occur in different types of non-literary contexts with the same semantic-pragmatic effect. The analysis of the corpus data further showed that the preferred order of verb and deictic varies among verbs: The *was-now* order, for instance, is the preferred one, regardless of whether *be* functions as the copula, the *be* of the progressive, or of the passive (all of these were counted as instances of the construction at hand), while the preferred order for the other verbs is with the deictic preceding the verb (e.g. *now felt*, *now saw*, *now thought*). While this parameter is significant in a usage-based study and important for the formal description of the construction, it will be ignored in the rest of the analysis, since it does not seem to entail a significant semantic difference.⁶ The frequency and distribution of the pattern as a function of the particular verb also show significant variation; compared to *was-now*, whose hits exceed 3,000, *now felt*, *now thought*, and *now saw* all have fewer than 100 occurrences. These results are certainly consistent with certain combinations being more conventionalized instances of the general pattern, which may be stored independently and in addition to it (as suggested by Goldberg 2006: 55).

In the non-literary instances, the *Past + now* pattern once again serves to shift the perspective to a vantage point close to or inside the narrated events, with an effect of zooming in on the events. The other consciousness, which anchors the viewpoint alignment, need not be so readily available, as in (4), or can be collective, as in (5); the perspective, however, is no less shifted, and the perception of the English team in (5) no less presented as experienced (or “represented”) than in (3) above. Similarly, interpreting (4) conjures up a participant at the scene, for whom the fear is experienced (or at least closely observed), rather than merely reported.

- (4) For over an hour the meeting wrestled with an undefined problem in the order. None had thought to specify how many shirt buttons could be undone and fear **was now** patently manifest that some uncontrollable disorder would be let loose should more than one button be opened to reveal ‘chest hair, or even a medallion’.

(M. Young, *An Inside Job: Policing and Police Culture in Britain*)

- (5) The batsman failed to survive the over though, Gooch managing to hold onto a slip catch. Waqar Younis thus joined Wasim Akram, the victorious Lord’s pair together again. England, defeated by these two then, **now saw** Akram go for a duck, a nifty piece of stumping by Russell from a curving, cutting ball from his captain.

(Wisden Cricket Monthly, periodical)

In FIS, therefore, the interpretational shift results unavoidably in viewpoint alignment with the consciousness or Self prominently available in the novel (Fludernik 1993: 204), and is on occasion sustained throughout the narration, with the aid of the other formal means characteristic of the style. In non-literary examples, on the other hand, in the absence of a prominent consciousness to whom the shifted content can be attributed as speech, or thought, or perception, the interpretation simply amounts to narrating events from a close-by, insider’s perspective (presupposing a participant or close observer, whether explicit or implicit).

Represented speech and thought are easier to identify, and may be considered more prototypical instances of shifted narration. Represented perception, on the other hand, presents more of a challenge in being unambiguously interpreted as the physical or mental content (conception) of a consciousness located at a different vantage point from the narrator’s. Evidence to the effect that all instances of the *Past + now* pattern mark a shift in perspective, even in the absence of other formal clues – as has been suggested here for Examples (3), (4), and (5) – comes from corpus examples such as (6) and (7); these accommodate the shift in the following discourse, formally marking it with expressions appropriate only to the shifted interpretation.

- (6) when most Jewish thinkers’ minds were somewhere between the atrocities of the Holocaust and the fearfully questionable use of the Bomb. Further, Israel **was now** a reality – “next year in Jerusalem” for millions had been fulfilled; but the reaction of the Arab world was very hostile – Suez was **months away** . . .

(D. Loranne and C. Rawlins, *Leonard Cohen: Prophet of the Heart*)

- (7) A week later Maxine arrived for her next consultation. So eager was she to experience another session of past life regression that she was fifteen minutes early for her appointment! I explained to her that,

although I could make no promises, since she **now felt** more confident about the technique, it was quite possible that **this** session would reveal the cause of her phobia. As she was now used to the technique, Maxine slipped quickly and easily into the regression. **This** time she told me that she was a lad called . . .

(U. Markham, *Hypnosis Regression Therapy*)

The proximal deictics in (7), unambiguously anchored to the present of the characters, and the temporal modifier in (6), whose landmark is clearly the time of the narrated events, support the shifted interpretation originating with the *Past + now* and are certainly consistent with an analysis that treats it as a viewpoint construction.

When the verb does not express physical perception, but rather denotes a cognitive or mental process, the interpretation favors an end-of-the-process reading, which profiles the final moment of the process or the initialization of the resulting state – for example, (8) and (9); Example (7) could also have such a reading.

- (8) with himself in the first and third person. “They suspect me of something – it is in the air – I am keeping something back. Vincent is hiding something that cannot stand the light.” This narrative flexibility indicates the extent to which he **now saw** himself as a character in his own drama. To put it another way, he was aware of the idea of divided consciousness, much discussed in his day, and here and elsewhere he can be seen applying it to his own actions.

(P. Callow, *Van Gogh: A Life*)

- (9) “You British!” said Penny Black, shaking her head in despair. She herself was British, in fact, but having spent several years as a graduate student in California, where she had been converted to radical feminism, she **now thought** of herself as spiritually an American, and tried as far as possible to speak like one.

(D. Lodge, *Nice Work*)

In Example (5), the event is presented as the “current,” co-temporal with the act of perception, experience of the third person character. In (8), in contrast, the “current” mental attitude of van Gogh is construed as the end phase of a cognitive process that has been going on for some time. This interpretation appears, in fact, to be one way of resolving the semantic conflict between punctual *now* and an inherently non-punctual, stative predicate (see also the discussion on the progressive in section 8.3).⁷

Literary FIS as a genre has been associated with third person narration, the choice of pronouns resolved differently from spatiotemporal deictics. However, some of the formal features of FIS appear in first person narratives as well, with

the same semantic-pragmatic effects – for example, non-anaphoric reflexives in novels narrated in the first person (Brinton 1995: 179). The *Past + now* pattern also appears in first person narration, earlier in fact than in third (Adamson 1995), and in non-fictional autobiography, as in (10):

- (10) The same things which I knew before **came now** in another manner,
with Light and Sense, and Seriousness to my Heart . . .
 (Baxter, *Reliquiae Baxterianae*, cited in Adamson 1995)

The zooming in on the past events, the shift in perspective regularly effected by the *Past + now* pattern in all the preceding examples is equally effected in first person examples. As observed by Adamson (1995: 203–4), the most common form of narrative in everyday language is in fact first person, and in such narrations, the *I* is both deictic and pronominal, standing simultaneously for narrator and character. In this context, subjectivity is transferred to our past self, especially if the narrated events are acts of consciousness or states of feelings. Although FIS is by definition restricted to third person (Bally 1912), this being a component of the grammaticalized profile of the genre, its historical prototype may well lie in first person narratives, fictional or not, featuring viewpoint shifting constructions. The first vs third person distinction, therefore, might not be so absolute or even critical for shifted narration in general.

Before turning to the constructional analysis, it should be noted that the *Past + now* pattern may also appear in the progressive form – for example, (11). The small number of hits for the *Past progressive + now*, equally low in literary and non-literary texts, both for the same verbs that were checked in the simple past and also for other, otherwise very frequent, verbs, such as *come*, *try*, suggests in fact that the (simple) *Past + now* is the more conventional pattern of the two. The shift in perspective and the alignment with the consciousness of the other is present in examples with the progressive as well, and, if anything, is even more pronounced in a way consistent with the function of the progressive I outline in the following section.

- (11) She must have done something for which she was now being made to pay. Spending most of each day in the house, she had no one except Maria to talk to and Maria was not worth the effort. She would rather talk to Pilade though she knew **this** was only a sophisticated way of talking to herself.

(A biography of Kylie Minogue)

In the next section, I examine separately the contribution of each component – that is, the past, the progressive, and the proximal deictic, showing that, although motivated, the overall interpretation should be attributed to the pattern as a whole.

8.3 *Past + now as a construction*

A constructional analysis, which assigns the interpretation to the pattern as such, competes with an account in which either or both of the parts in this expression are taken to be polysemous – that is, having extended meanings that allow them to co-occur. Starting with *now*, *Cobuild*, for instance, recognizes a function in which “now is used with the past tense, especially in novels and stories, to refer to the particular time in the past that you are speaking or writing about, as opposed to any later or earlier time” – the emphasis on “particular time” and “as opposed to a later or earlier time” presumably motivating the extension from the basic sense.⁸ However, even if a polysemy treatment for *now* may motivate its presence in the previous examples, it fails to predict the essential condition associated with the pattern, namely its occurrence in a narrative context only, so that the dictionary definition above finds it necessary to specify this explicitly. Indeed, it is only in a monologic narrative context that the *Past + now* combination is possible, as shown by the unacceptability of (12) and (13).

- (12) (?? I'm telling you,) George *now* appeared in the doorway.
- (13) (?? I do not agree;) the error was *now* patently manifest.

In (12) and (13) we have a conversational context – that is, a context with an explicit speaker and addressee, which, according to Benveniste (1966) and Banfield (1982: 171), is precisely one feature that negatively characterizes the sentence of narration – the other being the absence of present (in the absence of the first part of the sentence, *now* can be taken to refer to the present of the narration – rather than the present of speaking – and hence the sentences become instances of shifted narration). In fact, even if *now* can function as a synonym of *at that time*, as suggested by its lexicographic treatment, it only (as opposed to its phrasal synonym) correlates with the shift in perspective that is characteristic of the expression. In addition, a polysemy, case-by-case analysis obscures the systematicity of this interpretation that is available to other proximal deictics as well (e.g. *today*, *this morning*, *this minute*, *this year*); although the research reported here has focused on the *Past + now* combination, the effect is the same with the other temporal deictics (e.g. [14]); specifying that the particular interpretation is available to all proximal deictics in this pattern allows us to avoid the proliferation of the same sense for a class of lexical items. Indeed, this has been one of the central arguments in favor of a constructional analysis in general (cf Goldberg 1995, 2006).

- (14) Where **was** he **this morning**, for instance? Some committee, she never asked what.

(Virginia Woolf, *Mrs Dalloway*)

In a constructional approach, therefore, *now* can be argued to retain to a large extent its basic semantics, while any differences are attributed to the constructional context. More precisely, I suggest that *now* still means “anchored to a present,” except that its co-occurrence with the past tense verb rules out the possibility that this present is the present of the speaker; instead, it is interpreted as referring to the “present” of the narration, whether this is the present of a character in it or a depersonalized time frame to which the narration refers (as in [4]). In the first case, the event is interpreted as the *current* perception or experience of the third person character; in the second, as a close-by reconstruction.

A constructional analysis may therefore be justified on the basis of the conventional association of this interpretation with a specific discourse type, and of the interaction of specific lexical items with a grammatical category (i.e. the past), which results in a fairly productive semi-schematic construction.⁹ This said, the current trend in lexical semantics is to shift the focus “from words as building blocks to usage events, in all their contextual detail” (cf Cuyckens *et al.* 2003: 21). In this sense, a polysemy account of *now* may well be compatible (and mutually informative) with the constructional approach advocated here. Recognizing a constructional pattern, however, allows us to address (and resolve) the apparent conflict in the interpretation by recourse to the mechanism of coercion, which I discuss shortly in more detail.

The other component in the construction, namely the past, resists a polysemy analysis on intuitive grounds. The past we find in shifted narrative, whether (auto)biographical or fictional, is clearly the same past we find in the non-shifted one. On the other hand, a polysemy link may be naturally assumed between the past of autobiographical or biographical narration – for example, (15), (16) – and the past of fictional narration – for example, (17).

- (15) I was at work all day yesterday.
- (16) He was the first of the Beatles to leave the band.
- (17) The other three all lay flat in the cockpit now. Harry sat on the steering seat. He was looking ahead, steering out the channel . . .

(Ernest Hemingway, *To Have and Have Not*)

In the one case, the event is temporally anchored to the present of the narrator, or truly past with respect to the time of narration, while in the other, it is not. However, the unmarkedness of the past in fictional narratives has prompted many researchers (Benveniste 1966; Banfield 1982; Adamson 1995; Wright 1995) to adopt a semantic account that is based on distance rather than time: an event can be distant by being past or by being fictional, motivating the similarity in form. This literary-based analysis tallies with the cognitive treatment of the tense system in English (e.g. Langacker 1991: chapter 6). Langacker suggests

that the two formal oppositions – that is, past morpheme vs absence of the past morpheme, and presence vs absence of a modal – are conceptually characterized with respect to an epistemic model structured in terms of immediate reality, known (and unknown) reality, and non-reality. More precisely, the tense opposition is based on the concept of proximity so that “instead of ‘present’ and ‘past’ we can speak more generally of a proximal/distal contrast in the epistemic sphere. The import of the unmarked member is that the designated process is immediate to the speaker. Its overtly-marked counterpart – what we can now call the distal morpheme – conveys some sort of non-immediacy,” temporal in the prototypical case, but possibly of other types as well (Langacker 1991: 245). In this view, the past morpheme indicates simply that the profiled process is not immediate to the speaker (an essentially epistemic characterization), and although its prototypical value invokes the time model, this is only *one* manifestation of its basic epistemic meaning. The occurrence of the past in fictional narratives, therefore, falls well within this basic meaning, and whether an extension from the temporal prototype or yet another contextually triggered elaboration, constitutes a well-motivated use.¹⁰

What matters for present purposes is that the narrative past represents a well-entrenched, conventional, and certainly motivated construction. It is obviously the past of narration that appears in the relevant slot of the *Past + now* pattern. And it is the construction at hand, rather than the polysemy of *now*, which accounts for the presentation of a non-immediate, not directly accessible event as if it were directly accessible; put differently, the cancelled part of the semantics of the proximal deictic (which no longer points to the speaker/narrator) is attributed to the construction as a whole, to its conventional co-occurrence with a past tense verb in a narrative context, resulting in a particular (shifted) interpretation. In construction grammar terms, this is therefore a case of coercion resolved on the basis of the “override principle” (Michaelis 2005b: 51; also Goldberg 1995: chapter 2). Coercion refers to the clash between the morphosyntactic and/or semantic properties of a word and those of the construction in which the word is embedded; the override principle stipulates that if a lexical item is semantically incompatible with its syntactic context, the meaning of the lexical item conforms to the meaning of the structure in which it is embedded.¹¹

The constructional status of the *Past + now* pattern, so far argued exclusively on linguistic grounds, predicts that the association of the particular form (in a particular context) with the shifted interpretation is fully conventional; its discoursal properties should therefore be readily recognizable and the semantic/pragmatic effect directly accessible, independently of and in addition to its component parts. The diachronic variation in the use of the pattern may provide further support to its status as a constructional unit; a search through the TIME magazine corpus, for example, reveals that *was–now* is used with significant variation in frequency, with the peak of the curve centering in the decade of the 1940s (these results are summarized in Table 8.1 below). This variability in

journalistic discourse points to at least an idiom-like status for the construction at hand, consistent with its conventional association with particular semantics-pragmatics.¹²

Table 8.1 was–now in *TIME magazine*

<i>Decade</i>	1920	1930	1940	1950	1960	1970	1980	1990	2000	Total
<i>No of hits</i>	65	147	440	180	167	71	64	38	27	1199

The final component in this construction is the progressive. In line with earlier work (e.g. Lyons 1977, 1982), cognitive treatments (e.g. Langacker (1991: 207–11) analyze the progressive as coding a particular kind of construal, namely an internal perspective on the event, “as if one is watching it unfold rather than viewing it holistically as a unitary entity.” The English progressive is thus described as an imperfectivizing construction, restricting the profile of a process to a series of component states that does not include the initial and final ones. The profiled part of the meaning is an internal state, or series of states, of the ongoing event, presupposing therefore an insider’s vantage point. Note that this approach, whereby the progressive marks a particular construal, overrides the need to attribute special experiential pragmatics to this form distinct from its “truth-conditional” aspectual meaning, a distinction that has influenced earlier analyses (e.g. Ehrlich 1990).

It is presumably this semantics that has motivated the identification of the progressive as “one resource in the systematic combination of grammatical elements characterizing the linguistic structure of this style (i.e. FIS)” (Wright 1995: 153). We may therefore consider that it is a totally expected component in FIS (and shifted narration in general), making a predictable, compositional contribution to the overall interpretation. Still, the absence in such contexts – for example, (18) – of the normal reference time for the use of the progressive (*he was having dinner, when the phone rang*), licensed again by the narrative context, argues for this being a distinct use of the progressive.

- (18) He slowly nodded, as if saying, Yes, but that isn’t the point. He **was looking** hard at her. She **was being given** a warning, and from someone who **was taking** the responsibility for it. He might be a rather pitiable young man, and certainly an overtired and inadequately fed one, doing this job because he could not get another, but the weight of his position – the unhappy weight of it – **was speaking** through him . . .

(Doris Lessing, *The Fifth Child*, cited in Wright 1995)

Given the durative interpretation of the progressive (cf “profiles an internal state or series of component states”), punctual/momentary *now* appears to clash with it more prominently than with the simple past. Once again this semantic conflict

is resolved constructionally in favor of the embedding context – that is, *now* appears to neutralize the punctual aspect of its meaning and simply become a marker of the shifted viewpoint, in addition to the progressive. The double coercion thus renders the construction a highly marked and, therefore, powerful index of viewpoint. This is evidenced by the fact that in the instances found in the corpus the shift in perspective originating in the *Past progressive + now* is sustained in the following discourse more often than with the simple *Past + now*; while the shift may also persist with the simple past (Examples [6], [7]), this persistence is less frequent than with the progressive, where the shift is maintained much more systematically. The sustained shift may be marked by grammatical or lexical means, but it may also occur in their absence – that is, it may be only the *Past progressive + now* that triggers the shift and sustains it in the following discourse. Consider Examples (19), (20), and (21).

- (19) veiled threat to Richard? A reminder that he should not take it for granted that he would in time succeed to England, Normandy and Anjou? If Richard insisted on keeping Aquitaine, would he have to give up his claim to inherit the rest? That Henry **was now thinking** of Richard's keeping his duchy is suggested by the plans he was making to install John as King of Ireland, but **what price would the Old King demand in return for this concession?** Richard was alarmed. Immediately after Christmas he obtained permission to return to Aquitaine and we . . .

(J. Gillingham, *Richard the Lionheart*)

- (20) A few fragments of carrot were still lying about near the spring, but he had left these untouched and was eating the grass not far from the gnarled crabapple tree. They approached and he looked up. Hazel said nothing and began to feed beside him. He **was now regretting** that he had brought Bigwig. In the darkness before morning and the first shock of discovering that Fiver was gone, Bigwig had been a comfort and a stand-by. But **now**, as he saw Fiver, small and familiar, incapable of hurting anyone or of . . .

(Richard Adams, *Watership Down*)

- (21) or in a crowd? Alas, solitude is not very likely, there is so little of it in life, so what can we expect after death! After all, the dead far outnumber the living! At best, existence after death would resemble the interlude she **was now experiencing** while reclining in a deckchair: from all sides, she would hear the continuous babble of female voices. Eternity as the sound of endless babble: one could of course imagine worse things, but the idea of hearing women's voices forever, continuously, without end . . .

(Popular lore, *Esquire*)

In (19), the shift is sustained by the inversion construction appearing in a non-direct speech context, whose content is attributed to Henry, while in (20), it is supported lexically by another *now*. But even in (21), without any additional marking, the following text is interpreted as the thoughts of the third person character. The possible readings of the text preceding or following a passage containing explicit markers of shifted narration is an issue addressed in the literature, especially from a literary angle (e.g. Ehrlich 1990; Sotirova 2004), and, in fact, there seems to be agreement that assigning point of view to unmarked passages is possible in general.¹³ Further discussion of this issue is certainly beyond our present scope. What I have tried to show is simply that a constructional, coercion-based analysis may provide some motivation for the highly marked, shift-amplifying aspect of the *Past progressive + now* pattern.

8.4 Mental space structure and the *Past + now* construction

Mental Space Theory has been applied to the analysis of FIS by Sanders and Redeker (1996), who observe that a wide range of perspectivization phenomena, including point of view in discourse, can be accounted for by the embedding of mental spaces and the transference of viewpoint from one to the other. In a narrative text, the reality of the narrator is the basic mental space, the Base Space, which is the starting point of the discourse representation. Every time the narrator lets the characters speak or present their thoughts, experiences, and so on, an embedded mental space is created within the Base Space. In this framework, they describe the three basic modes of representing speech and thought, namely, direct, indirect and free indirect, as well as a fourth mode, which they label “implicit perspective” (accounting for cases where without direct, indirect, or free indirect representation, there is still perspectivization of the presented events).

Two points in their analysis are important for the present discussion. First, they stipulate that the default interpretation for FIS is one in which viewpoint is located in the embedded (rather than the base) space, which is also the space in focus.¹⁴ Although this is in agreement with much of the literary and linguistic literature, given the conflicting clues in FIS (tense and choice of pronouns are determined by the Base Space, while deictics and expressive elements by the embedded space), the location of viewpoint should in principle be open and susceptible to either interpretation. In fact, as argued by Vandelanotte (2004), FIS involves two separate deictic centers, even if none is fully operational; nevertheless, “the sayer/cognizant-related features actually outnumber and outweigh the speaker-related ones” (Vandelanotte 2004: 493; see also Traugott and Pratt 1980), placing FIS, as a discourse mode, closer to direct than to indirect speech or thought. So both of these linguistic analyses converge on a treatment of FIS

as incorporating a default perspective based on the character rather than the narrator.¹⁵ For example, in (22) (Example [5] from Sanders and Redeker 1996),

- (22) He heard something and turned around. There were the three Englishmen again. Now, could they really be tourists?

the most natural interpretation is that the character referred to by *he* already knows the three Englishmen and is thus able to refer to them by the definite description.¹⁶

In the following discussion, I essentially adopt this analysis, which, after all, reflects the basic intuition that FIS, and, more generally, shifted narration, involves transference of subjectivity *from* narrator *to* character, and not the other way round. However, Sanders and Redeker's approach to the mental space structure of FIS is made without any reference to the specific formal marking that serves to set up the embedded space and trigger the shifted interpretation. What I suggest is that the shift in viewpoint associated with FIS as a whole is effected in different ways, depending on the particular formal cluster that triggers it. While the *Past + now* pattern, which has been the main focus here, is a central feature in the expression of FIS, and of shifted narration in general, there are, as noted, other formal indices of an FIS interpretation. Importantly, although all or some of these features may coexist in a given passage, the presence of one of them may also be enough to support the shifted reading. It stands to reason, therefore, that in each case meaning is constructed through a different route, involving different mental space configurations, vital relations, and cross-space links (Fauconnier and Turner 2002). Indeed, as shown in Sweetser (2005), similar (end-point) mental space structures may be prompted or built in quite different ways. Whatever the space structure may be for an example like (23), it obviously involves complex identity relations for the referring pronouns appearing in an inverted question, which is embedded in a non-direct speech context, thus precluding a Base Space interpretation.

- (23) The way to the Regent's Park Tube station – could they tell her the way to Regent's Park Tube station – Maisie Johnson wanted to know.

(Virginia Woolf, *Mrs Dalloway*)

But the mental space configuration of (23) cannot be the same as the mental space configuration set up by the *Past + now* construction – that is, while they both result in a shifted interpretation, they reach it through a different route. In the same line, Janzen (this volume) analyzes an ASL construction of viewpoint shifting, formally realized in terms of mentally rotated (as opposed to static) space. This fully grammaticalized construction is “prototypically used in the narrative past,” which renders it directly comparable to FIS in terms of its semantic import, although the interpretation relies on totally different formal

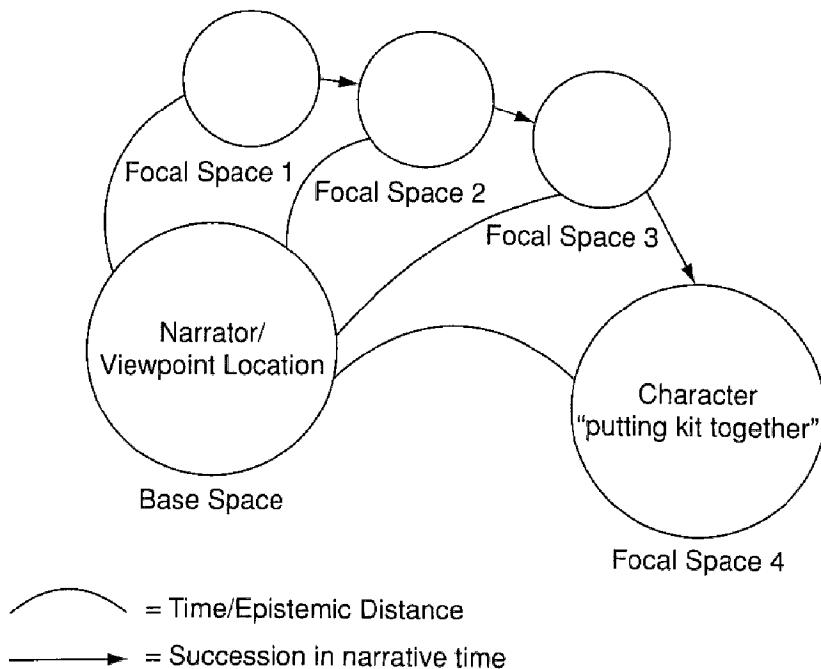


Figure 8.1

clues – for example, 180-degree spatial rotation rather than manipulation of a time relationship. I now turn to the *Past + now* construction, aiming to show that the conceptual integration network in this case relies directly – and more straightforwardly than for (23) – on the compression of time vital relations.¹⁷

The embedding context for *Past + now*, as shown in the previous sections, is the context (construction) of the past narrative. This entails a default construal in which the narrator is located in a different space from the character and the narrated events, viewing them from a distance. In the case of biographical narration (or of true deictic past), the viewpoint space and the narrated space are therefore separated by time, while in fictional narration and depending on the analysis assigned to the past, they are separated by metaphorical time or epistemic distance. As argued in Dancygier and Vandelanotte (2009), we are capable of conceptualizing two spaces at the same time, which may be separated by space, time, or epistemic distance, and which may often contain apparently contradictory representations (see also Fauconnier and Turner 2002: chapter 5). In ordinary (non-shifted) third person narration – and even in first person narration – the narrator's space (i.e. the viewpoint space) is backgrounded, the focal space (in Sanders and Redeker's terms) being the space of the narrated events.¹⁸ So the first part of the passage in (24) (up to the *Past + now*) prompts an interpretation that can be sketched as in Figure 8.1.

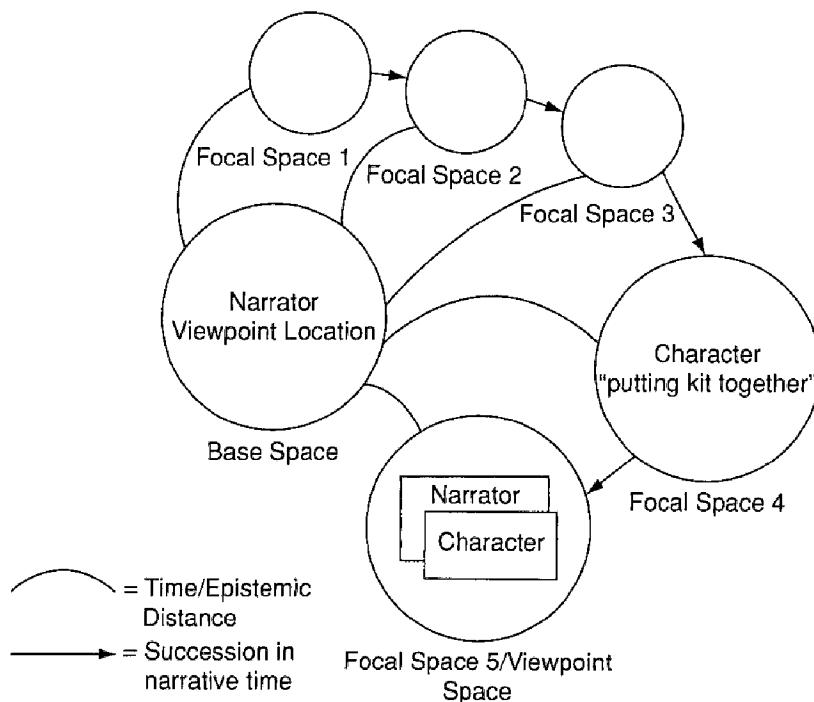


Figure 8.2

- (24) A few years ago I wrote a little book called *Discovering Backpacking*. Just before Christmas I was intrigued to get a letter in careful dictionary English from Kenya, saying that the sender had read my book and **was now putting** some kit together along the lines I had suggested. However, one piece of equipment was causing him a problem – a survival bag. Where could he obtain one, and how much would it cost?

(periodical, domain: leisure, unsigned)

But the occurrence of *now* in the context of past shifts the viewpoint to the space of the narrated event, collapsing the viewpoint of the narrator (external or separate up to this point) with that of the other consciousness, and in this way zooming in on the event (in this example, the putting together of the kit to which the narrator, and therefore the reader, is construed as a witness). This is achieved precisely by the present time deictic, which triggers in the blend the compression of the time (and, by extension, the space) distance separating the viewpoint from the focal space (Figure 8.2). The narrative continues then from this new blended space, which becomes the Viewpoint Space, at least for a while, as evidenced by the inverted question, whose content is clearly attributed to the sender of the letter.¹⁹

What I suggest, then, is that the coerced interpretation of the *Past + now* construction corresponds directly to a blend resulting from the compression of

a time vital relation; while all the other formal realizations of FIS may result in a very similar blended space with collapsed viewpoints (and, by implication, involving time compression of the input spaces as well), the constructional meaning of the pattern at hand is time compression or the abolishment of temporal (and, by extension, spatial and/or epistemic) distance. This is in accordance with the characterization of the *Past + now* cluster as prototypical among the formal realizations of FIS (Banfield 1982; Adamson 1995), on the basis of its occurrence in all represented speech, thought, and perception, and its necessarily triggering a shifted interpretation. Further, what is captured by the representation in Figure 8.2 is precisely that viewpoint itself becomes part of the focal space. This, I suggest, is precisely the *raison d'être* of the *Past + now* construction and, arguably, of this type of shifted narration, and, by extension, of FIS: to profile not only the narrated events, but also their narration from a particular point of view.²⁰ The fact that viewpoint itself is part of the designation or the profile of this discursal construction implies, of course, its markedness as a mode of discourse (as opposed to simple narration or conversation), and, perhaps, greater metalinguistic awareness in its processing. While this study does not extend in this direction, these points can be tested empirically and may open lines for further inquiry.

8.5 Conclusion

Discoursally motivated constructions – that is, conventional associations of form with a particular discourse type, may provide a principled account of formal combinations that deviate from the expected syntactic and semantic norms. *Past + now* as a formal cluster is only possible in a narrative context, defined as non-conversational and non-present; in this context, it evokes a shift to the perspective of a consciousness other than the narrator's (including, as shown, the narrator's past self); as shown in some examples, the shift may persist in the following discourse, even without any other formal marking, further justifying an analysis that can extend to chunks of language larger than the sentence.

As a literary style, FIS is characterized by the *Past + now* pattern, along with other constructions and expressions. As suggested earlier, we may even postulate a more extended or higher-level narrative construction, comprising all the formal reflexes associated with FIS, whose semantic-pragmatic import is to narrate from the point of view of the character. Although not all constructional theories would accommodate chunks larger than the sentence and include diverse formal realizations as a single construction, such an analysis seems plausible on intuitive grounds; as is evident from some of the preceding examples, the occurrence of one of these constructions or expressions in an earlier stretch primes certain expectations as to the occurrence of another later

on. In support of this, it is worth mentioning another conventional expectation in FIS, namely the frequent occurrence of epistemic adverbs – for example, *surely*, *perhaps*, or *of course* – whose expression of judgment is attributed to the character, as in (25).

- (25) When there are fifteen people sitting down to dinner, one cannot keep things waiting for ever. **She was now beginning** to feel annoyed with them for being so late; it was inconsiderate of them . . . Yet **of course** tonight, of all nights, out they went, and they came in late . . .

(Virginia Woolf, *To the Lighthouse*, p. 58, cited in Bosseaux 2004)

To the extent that these can be defined as a lexical set, or a kind of drop list from a particular semantic domain, they can be accommodated in the analysis of the large-scale construction as its *lexical* fillers. Since all versions of the theory allow for both grammatical *and* lexical fillers, Construction Grammar appears ideally suited for the analysis of such conventional discourse patterns, provided Construction Grammar extends in this direction.

Discoursally based constructions may be seen as originating in the concept of “idioms with a pragmatic point” (in Fillmore *et al.*’s [1988] classification of idiomatic expressions). Like *once upon a time* (a typical example of the category), *Past + now* evokes a specialized discourse context, and any motivational account of the contribution of its parts can be given only in reference to this context. The semantics of the past, the progressive, and *now* figure centrally in such an account; as argued here, the interpretation associated with the pattern can be analyzed in terms of constructional coercion, resolving the conflict with the deictic in the case of the simple past, and, additionally, with the punctual meaning in the case of the progressive. The coerced interpretation cues a particular kind of blended space, which captures the two most central features in the makeup of the construction. First, a collapsed viewpoint interpretation (as in Figure 8.2) presupposes and relies on the availability of a separate conceptualization of the narrator’s and the character’s viewpoint spaces (as in Figure 8.1); in this way, the notion of the necessary embedding context for the construction and, in general, for FIS is given actual pragmatic content. Second, viewpoint, as already stated, is part of the profile of the *Past + now* construction; this, in turn, is consistent with the instantly recognizable shift in perspective associated with this form.

References

- Adamson, Sylvia. 1995. From empathetic deixis to empathetic narrative: stylisation and (de)subjectivisation as processes of language change. In Dieter Stein and Susan Wright (eds.). *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge University Press, 195–224.

- Ariel, Mira. 1998. Discourse markers and form – function correlations. In Andreas H. Jucker and Yael Ziv (eds.), *Discourse Markers*. Amsterdam: John Benjamins, 223–59.
- Bally, Charles. 1912. Le style indirect libre en français moderne. *Germanisch-Romanische Monatsschrift* 4, 549–56, 597–606.
- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston: Routledge and Kegan Paul.
- Benveniste, Emile. 1966. *Problèmes de Linguistique Générale*. Paris: Gallimard.
- Berman, Ruth A. and Dan I. Slobin. 1994. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale NJ: Lawrence Erlbaum.
- Bosseax, Charlotte. 2004. Translation and narration: a corpus-based study of French translations of two novels by Virginia Woolf. Ph.D. dissertation, University College London.
- Bray, Joe. 2007. The “dual” voice of free indirect discourse: a reading experiment. *Language and Literature* 16:1, 37–52.
- Brinton, Laurel. 1995. Non-anaphoric reflexives in free indirect style: expressing the subjectivity of the non-speaker. In Dieter Stein and Susan Wright (eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge University Press, 173–94.
- Collins Cobuild English Language Dictionary*. 1987 (repr. 1993). Editor-in-Chief John Sinclair. London: Collins.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford University Press.
- Croft, William and Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge University Press.
- Cuyckens, Hubert, René Dirven, and John R. Taylor (eds.). 2003. *Cognitive Approaches to Lexical Semantics*. Berlin/New York: Mouton de Gruyter.
- Dancygier, Barbara. 2005. Blending and narrative viewpoint: Jonathan Raban’s travels through mental spaces. *Language and Literature* 14:2, 99–127.
- Dancygier, Barbara and Eve Sweetser. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Dancygier, Barbara and Lieven Vandelanotte. 2009. Judging distances: mental spaces, distance, and viewpoint in literary discourse. In Geert Brône and Joeren Vandaele (eds.), *Cognitive Poetics: Goals, Gains and Gaps*. Berlin/New York: Mouton de Gruyter, 319–69.
- De Swart, Henriette. 1998. Aspect shift and coercion. *Natural Language and Linguistic Theory* 16, 347–85.
- Ehrlich, Susan. 1990. *Point of View: A Linguistic Analysis of Literary Style*. London/New York: Routledge.
- Fauconnier, Gilles and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind’s Hidden Complexities*. New York: Basic Books.
- Fillmore, Charles J. 1981. Pragmatics and the description of discourse. In Peter Cole (ed.), *Radical Pragmatics*. New York: Academic Press, 143–66.
1990. Epistemic stance and grammatical form in English conditional sentences. *Chicago Linguistic Society* 26, 137–62.
- Fillmore, Charles J., Paul Kay, and M. Catherine O’ Connor. 1988. Regularity and idiomacticity in grammatical constructions: the case of *let alone*. *Language* 64, 501–38.

- Fludernik, Monika. 1993. *The Fictions of Language and the Languages of Fiction: The Linguistic Representation of Speech and Consciousness*. London/New York: Routledge.
- Fried, Mirjam and Jan-Ola Östman. 2004. Construction Grammar: a thumbnail sketch. In M. Fried and J.-O. Östman (eds.). *Construction Grammar in a Cross-Language Perspective*. Amsterdam/Philadelphia: John Benjamins, 11–86.
- Galbraith, Mary. 1995. Deictic shift theory and the poetics of involvement in narrative. In Judith F. Duchan, Gail A. Bruder, and Lynn E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, NJ: Lawrence Erlbaum, 19–59.
- Genette, Gérard. 1980. *Narrative Discourse*. Oxford: Blackwell.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.
2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.
- Kay, Paul and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the what's X doing Y? construction. *Language* 75, 1–33.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites*. 2 vols. Stanford University Press.
1991. *Foundations of Cognitive Grammar. Vol. II: Descriptive Application*. 2 vols. Stanford University Press.
2000. *Grammar and Conceptualization*. Berlin/New York: Mouton de Gruyter.
2002. Deixis and subjectivity. In Frank Brisard (ed.). *Grounding: The Epistemic Footing of Deixis and Reference*. Berlin/New York: Mouton de Gruyter, 1–28.
- Lyons, John. 1977. *Semantics*. Cambridge University Press.
1982. Deixis and subjectivity: *loquor, ergo sum?* In Robert Jarvella and Wolfgang Klein (eds.). *Speech, Place and Action: Studies in Deixis and Related Subjects*. New York: Wiley, 101–24.
- Michaelis, Laura 2005a. Construction Grammar. [spot.colorado.edu/~michaeli/Michaelis_ELL\(CG\).pdf](http://spot.colorado.edu/~michaeli/Michaelis_ELL(CG).pdf).
- 2005b. Entity and event coercion in a symbolic theory of syntax. In Jan-Ola Östman and Mirjam Fried (eds.). *Construction Grammars: Cognitive Grounding and Theoretical Extensions*. Amsterdam/Philadelphia: John Benjamins, 45–88.
- Nikiforidou, Kiki. 2010. Viewpoint and construction grammar: the case of *past + now*. *Language and Literature* 19:3, 265–84.
- Östman, Jan-Ola. 2005. Construction discourse: a prolegomenon. In Jan-Ola Östman and Mirjam Fried (eds.). *Construction Grammars: Cognitive Grounding and Theoretical Extensions*. Amsterdam/Philadelphia: John Benjamins, 121–44.
- Sanders, José and Gisela Redeker. 1996. Perspective and the representation of speech and thought in narrative discourse. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 290–317.
- Sotirova, Violetta. 2004. Connectives in free indirect style: continuity or shift? *Language and Literature* 13:3, 216–34.
- Sweetser, Eve. 2005. Putting the “same” meaning together from different pieces. In Sophia Marmaridou, Kiki Nikiforidou, and Eleni Antonopoulou (eds.). *Reviewing Linguistic Thought: Converging Trends for the 21st Century*. Berlin/New York: Mouton de Gruyter, 23–51.

- Sweetser, Eve and Gilles Fauconnier. 1996. Cognitive links and domains: basic aspects of Mental Space Theory. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 1–28.
- Traugott, Elizabeth Closs and Mary-Louise Pratt. 1980. *Linguistics for Students of Literature*. New York: Harcourt, Brace, Jovanovich.
- Vandelanotte, Lieven. 2004. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36, 489–520.
- Vološinov, Valentin N. 1973. *Marxism and the Philosophy of Language*. Trans. L. Matejka and I. R. Titunik. New York/London: Seminar Press.
- Wright, Susan. 1995. Subjectivity and experiential syntax. In Dieter Stein and Susan Wright (eds). *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge University Press, 151–72.

9 “Wait till you got started”

How to submerge another’s discourse in your own

Lieven Vandelanotte

9.1 Introduction

Research from various angles has amply demonstrated the fact that viewpoints of others are routinely embedded as the speech, thoughts, or emotion states of those others in various text types. In conversation, for instance, direct speech is often used to highlight dramatic peaks (e.g. Mayes 1990; Shuman 1993; Holt and Clift 2007); news reporting uses both explicit and implicit speech reporting modes to incorporate the viewpoints of different sources (e.g. Short 1988; Semino and Short 2004; Sanders 2010); and fiction makes extensive use of a range of direct and (free) indirect modes of discourse presentation, not just to move the action forward, but also to conjure up characters’ inner lives (e.g. Banfield 1982; Fludernik 1993; Vandelanotte 2009).

This chapter takes as its focus only cases where reported material is explicitly presented,¹ typically in the form of a reported clause, though sometimes the reported material particularly in direct speech or thought consists of just a minor clause or one-word utterance (*She said, “Yes”, I was like “Wow!”*). This excludes from consideration so-called “narrative reports” of speech or thought acts as described by Leech and Short (1981: chapter 10) and Semino and Short (2004), as in *he thought about his childhood* or *she talked on*. Discussion of forms that present the reported material usually centers on three main forms, direct, indirect, and free indirect speech or thought, which have been analyzed in cognitive linguistics as involving different forms of mental space embedding (Sanders and Redeker 1996), with the embedded space being accessed either directly as a new Base Space (in direct speech/thought) or indirectly, via the narrative Base Space, with different degrees of narrator’s influence (in [free] indirect speech/thought).

The first aim of this chapter is to argue that the variable landscape of speech and thought representation is better described if a further, fourth landmark is recognized in it, in addition to the three traditionally included. This type, called “distancing indirect speech or thought” (Vandelanotte 2004a, 2004b, 2009), differs from the types looked at by both Sanders and Redeker (1996) and Nikiforidou (this volume), in that the deictic viewpoint of the “narrator”

is not shifted wholly (as in direct speech/thought) nor even partly (as in [free] indirect speech/thought) to the character. Instead, another's discourse space is evoked from the narrator's deictic viewpoint, to serve the narrator's discourse purposes – for instance, to provide evidence or to express irony or sarcasm.

In thus extending the frame of reference for describing forms and functions of speech and thought representation, the claim is not that every instance encountered will fit seamlessly into one of four clear-cut categories. As is well known, grammatical categories show prototype structure as much as lexical items (Lakoff 1987; Taylor 2003 [1989]), so it should come as no surprise that mixed forms also occur in the realm of speech and thought representation, as witnessed, for instance, by the widespread journalistic usage of combining direct and indirect modes, as in Example (1), in which the first person pronoun *my* (along with the conventional use of quotation marks) indicates that a direct quote is incorporated into indirect speech.

- (1) Dr. Lopez said last night he wished “to express my gratitude to all the Stanford students for all the time spent” in the search for his small son.

(Stanford Daily, 20 February 1956, cited in Schuelke 1958: 90)

While mixed forms as well as ambiguous and vague instances should be recognized (cf Vandelanotte 2009: 171–5, 240–3), the claim here will be that separating out distancing indirect speech/thought, where the viewpoint remains with the narrator, from free indirect speech/thought, where viewpoint is located with the character, permits a more fine-grained analysis of reportative constructions.

The sense in which the various types of speech and thought representation can be called “constructions” is that described by Nikiforidou (this volume) and Dancygier (this volume), namely as discourse constructions that are not necessarily identifiable by a complete array of lower-level constructional forms. Instead, the presence of a particular salient aspect of form may metonymically afford or cue access to the whole constructional frame, with “contextual background [being] involved in setting off such cuing” (Dancygier and Sweetser 2005: 272).

In the next section, I will briefly survey some terminological preliminaries, and set out the main features of direct, indirect, and free indirect speech or thought. Against this background, section 9.3 then proposes to distinguish distancing indirect speech/thought (DIST) as a mode of speech and thought representation that incorporates another discourse without a shift in the deictic center from narrator to character, and section 9.4 illustrates some of its range of usage. Section 9.5 next considers the way in which the “distancing” characteristic affects the type of mental space blending involved in DIST, compared to the type of blend described for free indirect forms by Nikiforidou

(this volume). The role of DIST in the intersubjective management of viewpoints is considered in section 9.6, and conclusions are offered in section 9.7.

9.2 A brief sketch of direct and (free) indirect speech and thought constructions

A defining feature of all speech and thought representation is that it involves two events, with the *current* speech event serving as a frame from which a *represented* speech or thought event is accessed. Correspondingly, we can speak of the “current” speaker, who presents some material as the utterance or thought of the “represented” speaker; in narrative contexts, these two roles are often referred to respectively as the narrator and the character. The doubling of speech events also implies that other aspects making up the two situations usually differ, including time, place, and speech participants, which means we can distinguish current and represented addressees as well. The notion of deictic center can be used to refer to the total set of coordinates that require reference to a speaker for meaningful interpretation, including *I*, *you*, *here*, and *now*.

This rather limited set of primitives already prompts questions as to whose deictic viewpoint – the current or the represented speaker’s – determines such features as the use of first, second or third person pronouns, tense forms, and deictic adverbials. Apart from these concerns, syntactic considerations also play an important role in defining different types of speech and thought representation: the freer syntax of direct speech or thought, for instance, easily allows reported clauses starting with discourse markers and interjections, and can accommodate non-declarative syntax (question inversion, exclamations, and so on), whereas the more strongly dependent reported clause of indirect speech is limited in these regards. Let us look briefly at some of the resulting features of the three main types usually distinguished (for a fuller account, see Vandelanotte 2009).

An important feature of direct speech or thought (DST) is the full deictic and expressive shift (von Roncador 1980, 1988) from current to represented speaker: all manner of deictic forms (such as tense forms, personal pronouns, and spatiotemporal adverbials) in the reported clause of DST are viewed from the represented speaker’s deictic center. One obvious sign of this is the occurrence of two referents for the first person within the confines of a single sentence (Banfield 1982), as when *me* in Example (2) refers to the current speaker and *I* to the represented speaker.

- (2) Then he was introduced to Andrea. “*I* couldn’t get over how lovely she is,” she told *me*.

(Cobuild corpus)²

In addition to the shift in the deictic center, there is in direct speech or thought also a kind of syntactic “reset,” as the directly reported material can feature any one from the whole range of expressive features, such as interjections and discourse markers (*He exclaimed “Shit! I’m bored!”*, *She was like “Oh well, we’ll see”*), as well as non-declarative structures (compare *He asked her “Do you love me?”* to the declarative structure in indirect *He asked her whether she loved him*).

Indirect speech or thought (IST) is distinguished by the lack of the deictic shift to the represented speaker, particularly as regards pronouns, the grammatical number of which clearly is determined from the current speaker’s deictic viewpoint. Nevertheless, some forms do occur which deviate from the IST prototype in that they involve the kind of blend described by Nikiforidou (this volume) and Sweetser (this volume). In (3) and (4), for instance, the current speaker’s speaker-hearer referential system is blended with the represented speaker’s temporal deictic viewpoint.

- (3) Erm my little boy’s now seven and he will be one of the first to be tested and we went to a meeting today at school er for the head teacher to explain what was being done and he said that they’re not being tested as such but we were told to take it as an assessment.

(Cobuild corpus)

- (4) He told Soviet television he *now* had a mandate from the republics to take to the London talks.

(Cobuild corpus)

The present tense in (3) indicates that the view that the “assessments” are not really “tests” lies with the represented speaker (the head teacher), and not with the current speaker (the parent, who, judging by the second clause, does take these assessments to constitute tests; see Davidse and Vandelanotte 2011 for discussion). Similarly, the *now* in (4) refers to the represented speaker’s “original” *now*, and not to the moment of the current speaker’s report of the announcement.

A further indication of the possible viewpoint complexity of IST lies in the realm of referring expressions for which the reference can be resolved either from the current or from the represented speech event, as in the textbook example of Oedipus and his mother Jocasta:

- (5) Oedipus said that *his mother* was beautiful.

We know that in the legend Oedipus unwittingly married his mother, so the standard interpretation of (5) sees *his mother* as the current speaker’s designation, but in principle nothing precludes an interpretation whereby *my mother* is presumed to be the represented speaker Oedipus’s designation.

Thanks to its “free” syntax, free indirect speech or thought (FIST) can go further than IST in (re-)creating the represented speaker’s viewpoint: the reported material of FIST can include, for instance, interrogative and exclamative clauses, as in (6) and (7), incomplete sentences, discourse markers, and interjections (for many examples of such expressive devices used in FIST, see Fludernik 1993).

- (6) She opened her scissors, and said, *did he mind her just finishing what she was doing to her dress?*
(Virginia Woolf, *Mrs Dalloway*, cited in Banfield 1982: 281, note 7)
- (7) To-morrow was Monday, Monday, the beginning of another school week!
(D. H. Lawrence, *Women in Love*, cited in Oltean 2003: 173)

As Example (6) shows, I do not consider the absence of a reporting clause as a defining feature of FIST: in more typical cases like (7) none is present, and if one is present it is typically interposed or postposed, but if there is a kind of break, as suggested in (6) by the comma, even a preposed reporting clause is, less typically, possible. The main point is that the reported material of FIST allows “main clause” syntactic phenomena such as the interrogative clause structure in (6), and this distinguishes it from IST, which is more restricted in this regard.

In terms of viewpoint, the combination of having key deictic coordinates blended with the represented speaker’s coordinates (as with *To-morrow* in [7]), and of allowing the full range of expressive constructions to occur in its reported material, means that viewpoint in FIST becomes located in the represented speech event, with the represented speaker (or “character”), despite the fact that tense as well as the number of pronouns are seen from the current speaker’s deictic center (see, for example, the past tense in [6] and [7] and the third person *he* in [6]). These last two features effectively separate FIST from the direct mode, and the resulting blend has often been described in terms of a “dual voice” (e.g. Pascal 1977), on the assumption that both narrator and character “speak” at the same time. A different view is espoused by van der Voort (1986) and Vandelanotte (2009: chapter 7), among others, who argue that the blended deictic viewpoint of FIST does not mean that the current speaker or “narrator” overlays his or her own evaluation (for instance, ironic distance), but instead serves to silence, in van der Voort’s (1986) phrase, the “loud *I*” of DST often felt to be inappropriate for rendering a character’s typically unverbalized thoughts and feelings. When a current speaker’s subjective viewpoint does intervene beyond mere deictic alignment of pronouns and tense, I would argue that there are good grounds to treat such a case as an instance of distancing

indirect speech or thought (DIST) as a current-speaker oriented counterpart of FIST. It is to a description of this type that I turn in the next section.

9.3 DIST: evoking another's discourse from a constant deictic viewpoint

So far we have distinguished three prime landmarks in the landscape of speech and thought representation by combining a syntactic criterion – a freer vs a more dependent type of relation between reporting clause and reported material – with variation in terms of deictic shift. In DST, it was argued, there is a complete deictic shift to the represented speaker's deictic center. FIST is typically characterized by a partial deictic “displacement” (Sweetser and Nikiforidou, both in this volume), whereby spatiotemporal coordinates expressed by adverbs such as *now* and *here* are viewed from the represented speaker's deictic center. IST also allows this to some extent, though this seems less common. We can thus posit a cline from full deictic shift (DST) over partial (FIST) to a minimal or even absent shift (IST).

What this section argues is that DIST represents a further step on this cline, where typically the reported material is presented from the single deictic viewpoint of the current speaker in the syntactically freer kind of syntax shared with DST and FIST. Compared with FIST in particular, DIST then involves the current speaker structuring a “borrowed,” non-asserted thought or utterance from his or her own viewpoint, rather than locating viewpoint with the represented character as in FIST. The main grammatical symptoms promoting a reading as DIST are illustrated in this section.

In the realm of person deixis, what DIST uniquely allows is for first and second person pronouns to refer to the current speaker and addressee in the current speech situation. That this is not the case in the reported clause of DST will be clear: here, *I* and *you* refer to the represented, not the current speaker and addressee respectively. In FIST, *I* can occur only with first person reporting clause subjects to refer to the represented speaker, deictically remote from the current speaker (as in *Now my luck was coming to an end, I realized (then)*), and *you* only occurs in the infrequent context of second person narrative, as a kind of story-internal, “addressed” protagonist. Crucially, as Banfield (1982) has argued, no addressed *you* can occur in the reported clause of FIST.³

The examples in (8) and (9) below show how this possibility of incorporating another's discourse while keeping the viewpoint and deictic center constant has been exploited in narrative and drama. In (8), the I-narrator Max recalls a scene in which the topic of sin is discussed by a friar, referred to as “Foamfleck” by his students because he had “flecks of white stuff permanently at the corners of his lips”. (8a) and (8b) indicate a corresponding direct and free indirect rendering of the crucial part. The pronoun *we* in (8) refers to the current speaker (narrator)

Max, along with his peers; the direct “original” is *you* (8a), and opting for *they* would be consistent with a free indirect rendering in which the viewpoint is located with the friar in the represented speech event (8b). Instead, the current speaker Max “egocentrically” draws the report into his own viewpoint, echoing the friar’s discourse to serve Max’s current discourse purposes – for instance, the expression of a dissociative, possibly ironic attitude towards the friar’s “disquisition.”

- (8) I am recalling with especial clarity an enraptured disquisition he delivered to us one fine May morning on the sin of looking. Yes, looking. We had been instructed in the various categories of sin... but here it seemed was a new category: the passive sin. *Did we imagine, Fr Foamfleck scoffingly enquired*, pacing impetuously from door to window, from window to door, his cassock swishing and a star of light gleaming on his narrow, balding brow like a reflection of the divine effluvium itself, *did we imagine that sin must always involve the performance of an action? Looking with lust or envy or hate is lusting, envying, hating.*

(John Banville, *The Sea*)

- (8a) DST: *do you imagine...*
 (8b) FIST: *did they imagine...*

An example that illustrates the potential of DIST to feature second person pronouns referring to the current addressee is (9), in which the *you* directly addressed by Mrs. Zero is her husband, whom at the beginning of the play she elaborately condemns for not having lived up to the expectations he had raised when, as a young, ambitious man, he was her suitor. As shown in (9a), the husband’s role of current addressee is conflated with that of represented (“original”) speaker.

- (9) [Mrs. Zero:] I was a fool for marryin’ you... I wish I had it to do over again, I hope to tell you. *You* was goin’ to do wonders, *you* was! *You* wasn’t goin’ to be a bookkeeper long – oh no, not *you*. Wait till *you* got started – *you* was goin’ to show ‘em. There wasn’t a job in the store that was too big for *you*. Well, I’ve been waitin’ – waitin’ for you to get started – see? It’s been a good long wait too.

(Elmer Rice, *The Adding Machine*, cited in Fludernik 1993: 117–18)

- (9a) DST: *I’m going to do wonders, I am! I’m not going to be a bookkeeper long...*
 (9b) FIST: *He was going to do wonders, he was! He wasn’t going to be a bookkeeper long...*

The appearance of first and second person pronouns referring to current speaker and addressee is not the only striking characteristic flowing from DIST's deictic viewpoint singularity. Still in the realm of reference to people, the different viewpoint structure of FIST and DIST has implications for the choice of noun phrase type – pronouns vs fuller NPs such as proper names and descriptive NPs – when reference is made to represented speakers and addressees. In FIST, where viewpoint resides in the represented speech event, pronouns are used to refer to represented speakers and addressees because these speech participants are fully cognitively “activated” from the viewpoint of the represented speaker. Cognitively active referents are highly accessible and thus easily retrievable (Ariel 1990), which explains why pronouns can be used to refer to them. NPs such as proper names or descriptive NPs (e.g. “the minister of education”), on the other hand, mark low cognitive accessibility, and can be used to introduce referents not assumed to be known to an interlocutor, or to reintroduce such referents after some break in the flow of discourse (e.g. paragraph breaks, chapter breaks). In DIST, in which viewpoint remains in the current speech event (or base space), it is the current speaker who assesses the cognitive accessibility of characters, and who can thus choose to use proper names or descriptive NPs wherever judged necessary or appropriate.

As an example of this, consider the underlined NPs in (10) and their direct and free indirect counterparts in (10a) and (10b) respectively. The represented speaker, Sir William, and the represented addressee, Mrs. Warren Smith, are referred to pronominally in direct and free indirect speech, as the speaker and addressee in any exchange are inherently maximally accessible. When the entire obliquely rendered dialogue is drawn into the viewpoint of the current speaker (narrator) as in (10), however, it is the current speaker's prerogative to opt for proper names, if only to clearly distinguish the characters at every stage, but perhaps also to suggest mild mockery of the tone and style of the interaction by signaling the narrator's presence in the choice of NP type.

- (10) Shortly and kindly Sir William explained to her the state of the case.
He [Septimus] had threatened to kill himself. There was no alternative. It was a question of law. He would lie in bed in a beautiful house in the country. The nurses were admirable. Sir William would visit him once a week. If Mrs. Warren Smith was quite sure she had no more questions to ask – he never hurried his patients – they would return to her husband.

(Virginia Woolf, *Mrs Dalloway*, cited in Ehrlich 1990: 75)

- (10a) DST: *I will visit him once a week. If you are quite sure you have no more questions to ask...*
- (10b) FIST: *He would visit him once a week. If she was quite sure she had no more questions to ask...*

The way in which person deixis works in DIST thus suggests that the represented speaker's discourse space is aligned with the current speaker's viewpoint. In a sense this means that the represented speaker's discourse ends up submerged in that of the current speaker, who, by adjusting the former's utterance or thought deictically to his or her own viewpoint, effectively appropriates this "other" discourse. Such appropriation seems to bear resemblances to the system of mentally rotated space in American Sign Language as described by Janzen (this volume), in which the signer (like the current speaker in DIST) brings narrative interactants' views into alignment with his or her own view. This particular mode of conceptualizing space in ASL interaction requires identification of referents by means of full NPs (also more likely to be used in DIST), since "pronominal" (indexical) pointing to a space to identify a referent does not work within a 180-degree rotated space.

As an illustration of the fact that the description of DIST given so far extends to other domains besides person deixis, consider example (11), in which, apart from the occurrence of a first person referring to the current speaker (who, as shown in [11a], is the represented addressee, *you*), the occurrence of the adverb *there*, interpretable clearly from the current speaker's viewpoint, and not from that of the represented speaker, suggests a reading as DIST.

- (11) The moment I tried to speak of the business that had brought me to his house, he [Mr. Fairlie] shut his eyes and said I "upset" him . . . As to the settlements, if I would consult his niece, and afterwards dive as deeply as I pleased into my own knowledge of the family affairs, and get everything ready, and limit his share in the business, as guardian, to saying Yes, at the right moment – why, of course he would meet my views, and everybody else's views, with infinite pleasure. In the meantime, *there* I saw him, a helpless sufferer, confined to his room.

(Wilkie Collins, *The Woman in White*, "The Story Continued by Vincent Gilmore," I, cited in von Roncador 1988: 230)

- (11a) DST: *In the meantime, here you see me, a helpless sufferer, confined to my room.*
- (11b) FIST ($i \neq j$): *In the meantime, here he_i saw him_j, a helpless sufferer, confined to his_j room.*

Separating current-speaker-oriented DIST from represented-speaker-oriented FIST in the area beyond the direct and indirect prototypes on the basis of grammatical distributional criteria like those just sketched enables a more nuanced stylistic and narrative analysis, as it removes the difficulty in traditional models which claimed that FIST was sometimes empathetic, sometimes ironic (Leech and Short 1981; Semino and Short 2004; see Vandelanotte 2009: chapter 7 for discussion). Considering the radial structure of grammatical categories,

however, it is only to be expected that instances are found that do not completely correspond to the “idealized” kind of DIST that is deictically entirely current-speaker-construed. A case in point is (12), which represents a son’s grilling of his mother on the whereabouts of paintings he considered part of his inheritance, and the mother’s equally fierce riposte.

- (12) That did it. I shouted, I waved my fists, I stamped about stiff-legged, beside myself. Where were they, the pictures, I cried, what had she done with them? I *demanded* to know. They were mine, my inheritance, my future and my son’s future . . . Demand, did I? – I, who had gone off and abandoned my widowed mother, who had skipped off to America and married without even informing her . . . what right, she shrilled, what right had I to demand anything here?

(John Banville, *The Book of Evidence*)

The consistent use of *I* across the two sides of the dialogue, with first the son as speaker (*I demand to know*) and then the mother (*Demand, do you?*) shows that the I-narrator egocentrically presents the exchange from his deictic viewpoint (and so uses DIST), but the deictic *here* at the end of the excerpt cannot be understood as the narrator’s current location – the prison cell where he is rehearsing his defence for a theft which led to a murder. On account of this spatial coordinate located in the represented rather than the current speech event, (12) thus deviates slightly from the DIST prototype.

This section’s description of DIST as a current-speaker-oriented counterpart to FIST yields an interesting picture as far as discourse goals are concerned: what is it that the current speaker aims to “do” with the deictically adjusted discourse borrowed from the represented speaker? The next section argues that the current speaker’s communicative goals may vary, but that some kind of attitude is always expressed towards the appropriated or “borrowed” represented discourse.

9.4 Current speaker attitudes expressed in different uses of DIST

If in DIST the current speaker appropriates another’s discourse by adjusting it deictically to the current speech event, we can expect that there are current communicative purposes for doing this instead of shifting wholly or partly, as in direct or (especially free) indirect forms, to the represented speech event. As argued in Vandelanotte (2004b, 2009: chapter 7), these purposes may either be more associative or more dissociative. In terms of Relevance Theory (Sperber and Wilson 1995 [1986]), DIST can be considered an “echoic” type of language use, as it involves a type of meta-representation (like all reported speech), but combines this with an attitude expressed towards the meta-represented discourse.

More *associative* attitudes expressed in DIST can be found in contexts – for instance, in conversation, journalism or argumentative prose – where the current speaker associates with a viewpoint while at the same time indicating that the source of the viewpoint is ultimately another speaker. Reinhart's (1975: 136) example (13a) illustrates such usage, whereas (13b) is a FIST version with different effect.

- (13a) John will be late, he said. (current speaker's viewpoint)
- (13b) He would be late, John said. (represented speaker's, i.e. John's, viewpoint)

In this example set, the distinct status of DIST shows up in tense use,⁴ since associative or “evidential” DIST allows non-subordinated, “absolute” tense related directly to the current speaker's deictic center; FIST normally allows absolute tense only in the exceptional context of gnomic statements (see Vandelanotte 2009: chapter 6 for discussion). The relative (or “subordinated”) tense *would* in (13b) is future with respect to the represented speaker John's past moment of speaking, not to the current speaker's moment of telling. In contrast, in DIST (13a), the futurity *is* plotted precisely from the current speaker's deictic center, as it is she who is announcing John's late arrival, with reference to something John said to make her conclude he would be late (e.g. “I will be late,” “My car broke down so I'll have to take the bus”).

Attested examples that likewise show absolute tense in the reported clause are given in (14) and (15).

- (14) Indonesia has arrested the alleged operations chief of an al-Qaeda-linked terror group blamed for bombings on the resort island of Bali, the nation's police chief said Wednesday.
 (www.usatoday.com, December 4, 2002, cited in
 Vandelanotte 2004b: 560)
- (15) Prof. Nowé has taken a few days off, he says, so he won't be there.
 (attested from email, April 12, 2000, cited in
 Vandelanotte 2009: 209)

(14) is taken from an online news report and is more amenable to a DIST reading than a DST reading, on account of such elements as *Indonesia* and *alleged*, which are best understood as designations chosen by the current speaker (i.e. the journalist). The absolute tense *has arrested* strengthens the interpretation of (14) as “evidential” DIST, in which a news fact is reported with identification of the source in the reporting clause. (15) is a “conversational” example in which the reporting clause itself is already present tense, and the present perfect *has taken* is motivated by the relevance to the current speaker's moment of speaking (in the interaction with the current addressee): the currently relevant fact that

Prof. Nowé has taken a few days off implies that he will not be in his office, which in turn implies that there is not much reason for the current speaker (or addressee) to go up there.

Conversation can also be a source of strongly *dissociative* uses of DIST, particularly when another's negative comments about oneself are sarcastically echoed. If someone says *I am a moron, I can't do anything on my own*, thereby echoing what someone else said about him or her, the usual interpretation is as an ironic echo and not as an expression of low self-esteem. This DIST interpretation may be made more easily available if the hearer was present at the original exchange in which the insults were uttered, or if the current speaker accompanies his delivery with verbal, prosodic, or gestural cues.

In narrative examples, DIST often involves a milder attitude of mockery or some form of intellectual superiority – for instance, in representing, from a consistent current speaker's viewpoint, the petty concerns of, or the exchange of formalities between, different characters, as in (16) below. In this example, there are two brief stretches of direct speech, signaled typographically as well as by the use of *you* to refer to the represented, not the current addressee (*Very well, indeed, father. And you father?* and *Good afternoon, Mrs Sheehy*). The part in between those two lines obliquely represents the dialogue between Father Conmee and Mrs Sheehy in the form of DIST, as witnessed, among other things, by the use of full NPs to refer to the represented speaker Father Conmee and the represented addressee "the wife of Mr David Sheehy M.P." (compare direct *I am very glad to see you looking so well*).

- (16) He walked by the tree of sunnywinking leaves and towards him came the wife of Mr David Sheehy M.P.

– Very well, indeed, father. And you father?

Father Conmee was wonderfully well indeed. He would go to Buxton probably for the waters. And her boys, were they getting on well at Belvedere? Was that so? Father Conmee was very glad indeed to hear that. And Mr Sheehy himself? Still in London. The house was still sitting, to be sure it was. Beautiful weather it was, delightful indeed. Yes, it was very probable that Father Bernard Vaughan would come again to preach. O, yes: a very great success. A wonderful man really.

Father Conmee was very glad to see the wife of Mr David Sheehy M.P. looking so well and he begged to be remembered to Mr David Sheehy M.P. Yes, he would certainly call.

– Good afternoon, Mrs Sheehy.

(James Joyce, *Ulysses*, partly cited in Banfield 1982: 208)

The current speaker's appropriation of character discourse in (16) is shown not only in the choice of noun phrase type, but also in the deliberate choice to present only one side of the dialogue in the DIST part, that of Father Conmee, addressing Mrs. Sheehy and reacting to her replies, which we do not get (the part *Still in London* could be read as Mrs. Sheehy's reply, but seems more likely to be Father Conmee's consenting repetition of her reply). All of this, in combination with the rather pompous use of the lengthy NPs *the wife of Mr David Sheehy M.P.* and *Mr David Sheehy M.P.* in one and the same sentence, adds an air of comedy and mockery, though not irony as such, to the scene.

With this overview of some of the main characteristics (section 9.3) and usage range (section 9.4) of DIST in mind, we are in a position to consider more closely the way in which DIST involves distancing and blending.

9.5 DIST, discourse distance, and blending

In a general sense, all types of speech and thought representation are "non-assertive" or "non-commitment" constructions: they allow us to *say* things without *asserting* them ourselves. If DIST is called "distancing," then, it is clear that something more specific is meant than this general, non-assertive feature. The specific sense of distance involved in DIST results from the unusual combination of two characteristics:

- (i) The absence of deictic shifting to the represented speech event, and of syntactic limitations of the kind IST imposes on its reported clause, means that there are no clear deictic or syntactic indications to suggest that something other than an ordinary speech act is going on (to which the current speaker would be committed).
- (ii) In spite of this, nonetheless two speech events – current and represented – are involved, possibly signaled by the presence of explicit reporting clauses or by the presence of expressive elements originally linked to the represented speaker, or else contextually reconstrued.

A kind of contextual "doubling back" (Galbraith 1995: 40), whereby an initial interpretation of a stretch of discourse has to be altered on closer inspection, has also been discussed for FIST (e.g. Ehrlich 1990; Galbraith 1995; Mey 1999; Sotirova 2004; Bray 2007). However, in the case of DIST it can be argued that the "garden path" effect is more fundamental: because shifted deictics such as a *now* drawing its reference from the represented speech event will typically be absent, and because of the use of first and second person pronouns and full NPs described above, readers may initially be led up the garden path of a non-reportative reading as a straightforward current speaker's statement (or question, etc.).

Such "momentary processing difficulties" (Sperber and Wilson 1995 [1986]: 237–43) often posed by DIST seem close in kind to those described by Tobin and Israel (this volume) for irony. Like irony, DIST requires a "zooming out"

or “re-evaluation” to resolve a perceived incongruity, in the case of DIST between current speaker construal and the involvement of a represented speech event. Along the lines suggested by Tobin and Israel, the recognition of this apparent mismatch or incongruity in DIST can be analyzed as prompting the current addressee/reader to decompress the blend of represented speaker discourse with the current speaker’s deictic viewpoint, and to construe a “higher,” decompressed current speaker’s space as the one from which the blend is ultimately viewed. Unlike irony more generally, however, this zooming out in DIST need not always result in dissociative attitudes, and more associative uses can also be found (section 9.4 above).

While FIST can also be analyzed as a blend of the current and represented speaker’s spaces, in the manner described for the *was-now* pattern by Nikiforidou (this volume), in the FIST blend the viewpoint adopted is the represented speaker’s, with no decompression involved. In contrast, in DIST the blend adopts the current speaker’s deictic viewpoint and is decompressed to create the altered, attitudinally “distanced” view of the current speaker on the represented utterance or thought.

The fundamental sense in which DIST involves this attitudinal distance arrived at through decompression motivates the use of “distancing” in its name. Dancygier and Vandelanotte (2009) have specifically proposed the notion of *discourse distance* to refer to material being inconspicuously borrowed from another speaker by the current speaker to construct his or her own discourse, without a deictic shift and without subscribing to or asserting the borrowed thought or utterance. This utterance or thought need not necessarily really have occurred, but may be something that was inferred or contextually “around,” and may also be a non-verbal contextual prompt, as when someone responds to seeing a map of London by saying “The British Museum is near University College?” (Noh 2000: 148).

Discourse distance is similar to earlier concepts such as “echo” in Relevance Theory (e.g. Sperber and Wilson 1995 [1986]) or “mental space evocation” in cognitive linguistics (Dancygier and Sweetser 2000, 2005). It departs from the earlier concepts in attempting to analyze the phenomenon as an extension of the very basic underlying mental space structure of distance as a cognitive image schema, namely a set-up involving two separated spaces and the speaker’s deictic alignment with one of them. What is specific to discourse distance is that a separate discourse space, emanating from a separate speaker’s knowledge base, is used in the current speaker’s discourse, without the current speaker knowing or asserting the content of the “borrowed” discourse. This is different from epistemic distance (Fillmore 1990), which involves a “non-real” space distanced from the current speaker’s Base Space, and thus does not invoke a second speaker’s space. An example of epistemic distance is (17), in which the hire is known by the current speaker to be counter to fact. In contrast, in the distanced discourse example in (18), the current speaker constructs a reasoning

about the hire on the basis of information gleaned from another speaker (*you*), the truth of which the current speaker of (18) does not commit to.

- (17) If she had been hired, she wouldn't need our help anymore.
- (18) If (as you say) she was hired, she doesn't need our help anymore.

Other examples of discourse distance in the realm of grammar include metalinguistic negation (Horn 1985; e.g. *that analogy wasn't strained, it was irrelevant*), metalinguistic conditionals (Dancygier 1998; e.g. *all I need to do for the oral presentation is powerpointize, if that's a word*), and “cited predictions” (Dancygier 1993, 1998; e.g. *if (as you say) he'll be buying me a nice birthday present after all I won't get angry with him just yet*); broader applications in literary discourse are discussed in Dancygier and Vandelanotte (2009) and Vandelanotte (2010).

It might be objected at this point that, just as any type of speech and thought representation involves some manner of “non-assertion” and “non-commitment,” some form of appropriation and echoing of the represented speaker’s utterance or thought is always involved, also in other types than DIST. Sternberg’s (1982) “Proteus principle,” for instance, holds that the function of any stretch of represented speech or thought is determined wholly by the surrounding context, so that basically any function can be performed by any type. While I agree that the wider embedding of represented discourse influences the functions fulfilled in context, I do think it makes sense to attempt a description of typical functions corresponding to the typical grammatical features of reportative modes such as DIST. The sense in which DIST involves distancing or echoing is fairly specific, as it involves an attitude expressed towards an utterance or thought borrowed from a represented speaker, but deictically incorporated into the current speaker’s discourse. The way in which this distancing discourse representation contributes to the overall management of different subjectivities in a text forms the topic of the next section.

9.6 DIST and the mutual management of different viewpoints

As noted by Dancygier (this volume) and Ferrari and Sweetser (this volume), there are currently different, competing understandings of subjectivity and intersubjectivity in linguistics. Two prominent views on subjectivity, and the associated historical process of subjectification, are Langacker’s (1985) in terms of implicit construal of Ground-related meanings (see Sweetser, this volume), and Traugott’s (1989, 1995; Traugott and Dasher 2002) in terms of the encoding of speakers’ subjective belief states and attitudes. While Traugott (1995) and Langacker (2006) analyze the differences between these conceptions, both De Smet and Verstraete (2006) and Ferrari and Sweetser (this volume) suggest

ways in which they can be reconciled. De Smet and Verstraete (2006: 369–70) pin down the way in which the two conceptions can be seen as overlapping as follows: Langacker's distinction between implicit and explicit reference to the speaker (as in implicit *this man* vs explicit *the man next to me*) subdivides along a vertical axis Traugott's understanding of "subjective" as "speaker-related," situated on a horizontal axis on which non-subjective constructions and subjective ones are distinguished.

This horizontal dimension can be enriched, as proposed, for example, in Traugott and Dasher (2002) and Traugott (2010), with a further point on the cline to distinguish intersubjective constructions, which in Traugott's understanding encode the speaker's awareness of the addressee's attitudes and beliefs, specifically the latter's "face" needs or self-image, and may include, for instance, honorifics and expletives. As Traugott (2010) points out, this conception of intersubjectivity differs from another that has recently hit the scene,⁵ namely Verhagen's (2005) broad notion of intersubjectivity as the mutual management of cognitive states within the speech event or "ground" constituted by (essentially) speaker and hearer. While Dancygier (this volume) is optimistic about the compatibility of these views, Traugott (2010: 32) insists on the difference, suggesting that Verhagen's "mutual management" understanding captures aspects of context, whereas her own main concern is "not with this context, but with linguistic MARKERS and EXPRESSIONS that index subjectivity and intersubjectivity and how they arise."

Verhagen's notion of intersubjectivity as involving the coordination of cognitive states between two conceptualizers (typically speaker and addressee) in the Ground almost by definition does not work with a principled distinction between "subjectivity" and "intersubjectivity"; indeed at times the two seem to be used almost as notional variants (e.g. Verhagen 2005: 18; 133). The fact that the concept, like Langacker's notion of subjectivity, is mainly used "analytically" or "synchronously" constitutes a further difference with Traugott's approach, in which the dynamic, diachronic application ("[inter]subjectification") plays a central role, and the distinction between pragmatic (inter)subjectivity in context and the semanticization or codification thereof is insisted on (cf Traugott 2010). As a result of this, however, there seem to be only a few genuine examples of "intersubjectification" in the strict sense.

While Traugott's precisely delineated notions of subjectivity and intersubjectivity allow one to chart the development of specific constructions from non- or less subjective towards subjective, and possibly into intersubjective meanings, Verhagen's notion seems better equipped to deal with the broader discourse level in charting the negotiation of different viewpoints between participants. The topic at hand in this chapter, DIST, is a discourse construction in the sense discussed by Nikiforidou (this volume), in that it needs contextual information about the current and represented speech events to combine with salient

formal cues to prompt the appropriate constructional frame. Considering its complex intertwining of conceptualizations attributed to different conceptualizers (the current and the represented speakers), DIST can thus more fruitfully be analyzed as an extension of Verhagen's framework, bearing in mind Dancygier's suggestion (this volume) that "intersubjectivity" in speech and thought representation is less argumentative in nature than Verhagen's examples such as clausal negation, which in prompting an alternative space foregrounds the hearer's viewpoint (cf Verhagen's example, *Mary is not happy. On the contrary, she is feeling really depressed*, discussed by Dancygier, this volume).

The point of departure for such an application to DIST is formed by Verhagen's (2005: chapter 3) analysis of complementation constructions as invitations to the addressee to entertain the content of the complement in the way that the onstage conceptualizer mentioned in the complement-taking clause does. Verhagen's focus is on complementation structures as in IST, of the (simplified) form *S-V (that)-Complement*, with extensions to impersonal complementation constructions such as *it is important that*. Constructions such as FIST and DIST, which represent speech or thought indirectly in complements with so-called "main clause" syntax and which typically occur either without reporting clause, or with an inter- or postposed one, are not directly examined by Verhagen.

The structural difference between IST (which is covered by Verhagen's analysis of complementation constructions) and the syntactically freer forms FIST and DIST is relevant to the kinds of lower-level constructions that may or may not occur in their respective reported complements. As we have seen in section 9.2, IST is more limited in its range of lower-level constructions: it reduces different clause types (such as interrogatives and exclamatives) to declaratives, and it does not as easily accommodate interjections, discourse markers, and "incomplete" complements (single words or phrases, as in *She said "No way!"*).

In addition, these structural differences are highly relevant to the identification of the relevant conceptualizer for the content of these complements. Due to the tighter incorporation of reporting and reported clause in IST, IST always has a sentence-initial reporting clause, putting the conceptualizer of the complement – the represented speaker – "onstage" in its subject (e.g. *he* in *he said that*...). In contrast, such an onstage conceptualizer is often not present in FIST and DIST, as witnessed by Examples (7), (9), (10), (11), and the second stretch of DIST in (8) above. If an onstage conceptualizer is explicitly given, this is typically in an interposed or postposed reporting clause, as in the first piece of DIST in Example (8) above. Sentence-initial reporting clauses followed by a prosodic and typographical break can occasionally also introduce FIST or DIST (as in the FIST example in [6] above), but the text frequency of this constellation seems very low. In most cases of FIST and DIST, then, what

Verhagen calls the “onstage conceptualizer” remains *implicit*, which leaves more work to be done by the addressee/reader to figure out a coherent reading.

As suggested above, in FIST this work is facilitated by certain lower-level constructions, such as the *was-now* pattern, which indicate the involvement of two deictically distinct speech events, with the current speech event forming the ground for the past tense and the represented speech event for the *now*. DIST, on the other hand, was argued in the previous section to require more effort to additionally decompress the blend of discourse spaces and to construe the current speaker’s decompressed space as the “higher” one, so as to resolve the incongruity between deictic singularity and the borrowing nonetheless of material from another, represented speaker’s discourse space. Following Tobin and Israel’s (this volume) interpretation of the decompression and zooming out in irony, DIST’s distance-taking with respect to the evoked represented speaker’s thought or utterance can be viewed as a kind of *desubjectification*, as the current speaker invites the addressee/reader to coordinate with his view on the represented speaker’s viewpoint. As the examples in section 9.4 have demonstrated, the “views on a viewpoint” afforded by DIST need not, as with the different kinds of irony Tobin and Israel discuss, always be dissociative, but can also be more associative.

9.7 Conclusions

In this chapter, I have argued that a more fine-grained description of the range of represented speech and thought phenomena is possible if the whole area between direct and indirect speech or thought is not lumped together in one broad category of free indirect forms, but rather approached through the lens of two categories: a represented-speaker-oriented type FIST and a current-speaker-oriented type DIST. In DIST, the current speaker submerges the represented speaker’s discourse in his or her own discourse, by adjusting it to his or her current deictic viewpoint and judgment of the cognitive accessibility of referents. This allows the current speaker to use the submerged represented speaker’s thought or utterance for current discourse purposes, such as providing evidence or expressing irony or sarcasm. FIST, in contrast, locates viewpoint with the represented speaker, while avoiding the full deictic shift to the represented speaker’s deictic center characteristic of DST, for instance, because this is deemed stylistically inappropriate or inelegant for the representation of intimate, “preverbal” thoughts and feelings.

Next, I have tried to specify the sense in which the echoing of a represented speaker’s discourse space from the current speaker’s deictic viewpoint in DIST constitutes distancing, or, more specifically, “discourse distance.” In order to make sense of the apparent clash between the current speaker’s deictic viewpoint and the use of represented speaker discourse, decompression is required

to construct the current speaker's viewpoint as that from which the represented speaker's is viewed, whether in a more associative or more dissociative manner. While this complex interplay between viewpoints contributes in a general sense to the "intersubjectivity" of the discourse, the attitudinal distance it involves can be analyzed as an instance of "desubjectification" of the represented speaker's viewpoint.

References

- Ariel, Mira. 1990. *Accessing Noun Phrase Antecedents*. London: Routledge.
- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston: Routledge and Kegan Paul.
- Bray, Joe. 2007. The "dual voice" of free indirect discourse: a reading experiment. *Language and Literature* 16:1, 37–52.
- Dancygier, Barbara. 1993. Interpreting conditionals: time, knowledge and causation. *Journal of Pragmatics* 19:5, 403–34.
1998. *Conditionals and Prediction: Time, Knowledge and Causation in Conditional Constructions*. Cambridge University Press.
- Dancygier, Barbara and Eve Sweetser. 2000. Constructions with *if*, *since* and *because*: causality, epistemic stance, and clause order. In Elizabeth Couper-Kuhlen and Bernd Kortmann (eds.). *Cause – Condition – Concession – Contrast*. Berlin/New York: Mouton de Gruyter, 111–42.
2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Dancygier, Barbara and Lieven Vandelanotte. 2009. Judging distances: mental spaces, distance, and viewpoint in literary discourse. In Geert Brône and Jeroen Vandaele (eds.). *Cognitive Poetics: Goals, Gains and Gaps*. Berlin/New York: Mouton de Gruyter, 319–69.
- Davidse, Kristin and Lieven Vandelanotte. 2011. Tense use in direct and indirect speech. *Journal of Pragmatics* 43:1, 236–50.
- De Smet, Hendrik and Jean-Christophe Verstraete. 2006. Coming to terms with subjectivity. *Cognitive Linguistics* 17:3, 365–92.
- Ehrlich, Susan. 1990. *Point of View: A Linguistic Analysis of Literary Style*. London: Routledge.
- Fillmore, Charles J. 1990. Epistemic stance and grammatical form in English conditional sentences. *Chicago Linguistics Society* 26, 137–62.
- Fludernik, Monika. 1993. *The Fictions of Language and the Languages of Fiction: The Linguistic Representation of Speech and Consciousness*. London: Routledge.
- Galbraith, Mary. 1995. Deictic Shift Theory and the poetics of involvement in narrative. In Judith F. Duchan, Gail A. Bruder, and Lynne E. Hewitt (eds.). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, NJ: Lawrence Erlbaum, 19–59.
- Holt, Elizabeth and Rebecca Clift (eds.). 2007. *Reporting Talk: Reported Speech in Interaction*. Cambridge University Press.
- Horn, Laurence. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61:1, 121–74.

- Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.
- Langacker, Ronald W. 1985. Observations and speculations on subjectivity. In John Haiman (ed.). *Iconicity in Syntax*. Amsterdam/Philadelphia: John Benjamins, 109–50.
2006. Subjectification, grammaticalization, and conceptual archetypes. In Angeliki Athanasiadou, Costas Canakis, and Bert Cornillie (eds.). *Subjectification: Various Paths to Subjectivity*. Berlin/New York: Mouton de Gruyter, 17–40.
- Leech, Geoffrey N. and Michael H. Short. 1981. *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. London: Longman.
- Mayes, Patricia. 1990. Quotation in spoken English. *Studies in Language* 14:2, 325–63.
- Mey, Jacob L. 1999. *When Voices Clash: A Study in Literary Pragmatics*. Berlin/New York: Mouton de Gruyter.
- Noh, Eun-Ju. 2000. *Metarepresentation: A Relevance-Theory Approach*. Amsterdam/Philadelphia: John Benjamins.
- Nuyts, Jan. 2001. *Epistemic Modality, Language, and Conceptualization: A Cognitive-Pragmatic Perspective*. Amsterdam/Philadelphia: John Benjamins.
2005. Modality: overview and linguistic issues. In William Frawley (ed.). *The Expression of Modality*. Berlin/New York: Mouton de Gruyter, 1–26.
- Oltean, Stefan. 2003. On the bivocal nature of free indirect discourse. *Journal of Literary Semantics* 32:2, 167–76.
- Pascal, Roy. 1977. *The Dual Voice: Free Indirect Speech and its Functioning in the Nineteenth Century European Novel*. Manchester University Press.
- Reinhart, Tanya. 1975. Whose main clause? (Point of view in sentences with parentheticals). In Susumu Kuno (ed.). *Harvard Studies of Syntax and Semantics* 1. Cambridge, MA: Department of Linguistics, Harvard University, 127–71.
- Sanders, José. 2010. Intertwined voices: journalists' modes of representing source information in journalistic subgenres. *English Text Construction* 3:2, 226–49.
- Sanders, José and Gisela Redeker. 1996. Perspective and the representation of speech and thought in narrative discourse. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 290–317.
- Schuelke, Gertrude L. 1958. "Slipping" in indirect discourse. *American Speech* 33:2, 90–8.
- Semino, Elena and Mick Short. 2004. *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.
- Short, Michael. 1988. Speech presentation, the novel and the press. In Willie van Peer (ed.). *The Taming of the Text*. London: Routledge, 61–81.
- Shuman, Amy. 1993. "Get outa my face": entitlement and authoritative discourse. In Jane H. Hill and Judith T. Irvine (eds.). *Responsibility and Evidence in Oral Discourse*. Cambridge University Press, 135–60.
- Sotirova, Violeta. 2004. Connectives in free indirect style: continuity or shift? *Language and Literature* 13:3, 216–34.
- Sperber, Dan and Deirdre Wilson. 1995 [1986]. *Relevance: Communication and Cognition*. 2nd edn. Oxford: Blackwell.
- Sternberg, Meir. 1982. Proteus in quotation-land: mimesis and the forms of reported discourse. *Poetics Today* 3:2, 107–56.

- Taylor, John R. 2003 [1989]. *Linguistic categorization*. 3rd edn. Oxford University Press.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65:1, 31–55.
1995. Subjectification in grammaticalisation. In Dieter Stein and Susan Wright (eds.). *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge University Press, 37–54.
2010. (Inter)subjectivity and (inter)subjectification: a reassessment. In Kristin Davidse, Lieven Vandelanotte, and Hubert Cuyckens (eds.). *Subjectification, Intersubjectification and Grammaticalization*. Berlin/New York: Mouton de Gruyter, 29–71.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge University Press.
- Vandelanotte, Lieven. 2004a. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36:3, 489–520.
- 2004b. From representational to scopal “distancing indirect speech or thought”: a cline of subjectification. *Text* 24:4, 547–85.
2009. *Speech and Thought Representation in English: A Cognitive-Functional Approach*. Berlin/New York: Mouton de Gruyter.
2010. “Where am I, lurking in what place of vantage?”: the discourse of distance in John Banville’s fiction. *English Text Construction* 3:2, 203–25.
- Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford University Press.
- van der Voort, Cok. 1986. Hoe vrij is de vrije indirekte rede? *Forum der Letteren* 4, 241–55.
- von Roncador, Manfred. 1980. Gibt die Redewiedergabe Rede wieder? L.A.U.T., Series A, Paper no. 71. Trier.
1988. Zwischen direkter und indirekter Rede: Nichtwörtliche direkte Rede, erlebte Rede, logophorische Konstruktionen und Verwandtes (Linguistische Arbeiten 192). Tübingen: Niemeyer.

Conclusion: multiple viewpoints, multiple spaces

Barbara Dancygier

The establishment, maintenance, and negotiation of viewpoints in discourse are cognitive phenomena which take a number of forms, affecting various levels of linguistic structure. While specific expressions (like modals) or constructions (such as represented speech and thought) have received thorough treatment in a variety of linguistic approaches, it is not common for broader ranges of viewpoint constructions to come under investigation at the same time. The framework chosen throughout this volume, Mental Spaces Theory, provides a natural opportunity for such a bird's-eye view and invites some reflection on the general nature of viewpoint in discourse.

Viewpoint phenomena come under scrutiny mostly in the cases where the basic deictic arrangement is somehow extended or disrupted. By default, we expect at least two participants in any discourse – the speaker and the hearer. These participants have each other's communicative needs at the center of attention and they alternate their deictic roles with every new turn. Thus, even when we are analyzing a speaker's expression of her construal or speech act, we are tacitly assuming that it does more than put the speaker on record with regard to some belief, description, or any other kind of utterance, and that the production of the utterance is situated in the context in which the hearer will receive it.

In the mental spaces and blending framework, the roles of the speaker and the hearer have been described in varying degrees of detail and have been given differing degrees of salience. For example, Coulson and Oakley (2005) assign the subjectivities of the speaker and the hearer to the “grounding box” of the expression, alongside the remaining aspects of the context, as they are deemed relevant to the exchange. Verhagen (2005), on the other hand, accepts the role of the Ground as a default basic arrangement, but proceeds to distinguish a variety of ways in which the subjectivities of the speaker and the hearer are involved in the negotiation, rather than simply communication of construals. In what follows, I will assume, with Verhagen, that the participants directly involved in an exchange, as well as the subjectivities evoked by the participants, are all engaged in viewpoint-specific construals negotiated in discourse. I will thus attempt to distinguish various classes of viewpoint configurations prompted by

the linguistic choices of the participants, and the ways in which these choices engage participants in the Ground and beyond.

There are several distinctions we need to draw. The viewpoint structure of an expression may be associated with one participant/subjectivity or more, who may be present in the context of the exchange (speaker and hearer) or not. Then, viewpoint may structure one mental space (locally) or a network of spaces (globally), all emerging in the discourse of one participant. For instance, a stance expression such as *perhaps* elucidates the speaker's epistemic viewpoint on the content of the utterance, but a complex clause such as *He must know I'm away* is engaging the speaker's stance and another person's stance. Even more interestingly, a participant may engage other viewpoints – either associated with the same participant in different situations, or with other participants, all brought into focus by the speaker, for the purposes of either cross-participant coordination, or negotiation, potentially leading to viewpoint shift or maintenance. Furthermore, as in the cases of reports and narratives, one participant (the speaker/narrator) may engage viewpoints expressed in other discourses, either simply to represent them/align herself with them, or to mark a specific attitude (of acceptance, irony, distance, etc.). Finally, discourse may focus on blending different viewpoints at the constructional level, instead of clearly marking the different viewpoint alignment of different participants. All these options require different constructional choices, and, as the chapters in this volume show, engage elaborate discourse patterns, designed specifically to manipulate viewpoint across different spaces.

Viewpoint may affect various aspects of utterance construal, beyond the spatial arrangement of the Ground – that is, the speaker may align her utterance with various aspects of the space topology (such as epistemic stance or temporal location), reaching far beyond the spatial viewing arrangement. In the simplest cases, epistemic modal verbs (*can*, *may*, *must*, or *have to*) mark attitudes of speakers, while deontic ones mark stances of obligating or permitting agents (who may be identified with the speaker). Stance verbs (such as *believe*, *know*, *doubt*, *wish*, or *guess*) project epistemic and/or emotional stance into the space in their scope. Viewpoint expressions may thus add or highlight aspects of the topology of one space, as in the case of modals, but they may also project aspects of viewpoint into the spaces lower in the network (as in the case of *wish* or *know*). A very good example of such projections is Cutrer's analysis (1994) of time in narrative stretches of discourse,¹ where the speaker's alignment with one temporal space, such as the present, allows her to then view past situations from the present perspective, as in the standard uses of Present Perfect forms, while an alignment with the past event yields a Past Perfect form. For example, *I have already seen "Avatar"* presents the speaker as looking at the past film-viewing from the perspective of the present moment, while *I had seen "Avatar" before you did* describes the movie-watching event (possibly the same as the

one referred to in the previous sentence) from the compound viewpoint of the speaker's "now" and the interlocutor's visit to the movie theatre.² Such chains of temporal spaces are indeed good examples of viewpoint taking, but what the analyses in this volume suggest is that viewpoint-related constructional or grammatical choices rely on types of viewpoint much more complex than the temporal construals prompted by the grammar of tense and aspect, especially on attitudinal and narrative choices. They also reach far beyond Cutrer's treatment of temporal viewpoint spaces as linked into a coherent viewpoint network. As the chapters in this volume show abundantly, a whole range of constructional phenomena depend crucially on competing, blended, or distanced viewpoints.

Epistemic stance

Possibly the most widespread range of viewpoint expressions has to do with epistemic stance – the speaker's or another participant's attitude to the certainty of what is being communicated. At the lowest level, stance can be marked through the choice of verb forms. Fillmore's (1990a, 1990b) initial discussion of verb choices as signals of stance helps to distinguish the positive-stance sentences using present or future forms in main clauses (as in *I will finish the paper tomorrow*) from neutral-stance present tense uses in causal and conditional clauses (*When/if I finish this tomorrow*), and negative-stance past tense forms (*If I finished this tomorrow*). Characteristically, the past tense is not used here to express temporal reference, but to mark the speaker's stance as negative – that is, as an expression of uncertainty, doubt, or, in other cases, deference or politeness (cf Fleischman 1989).

However, the verb-related stance phenomena are also associated with constructional patterns (cf Fillmore 1990a, 1990b; Sweetser 1996; Dancygier 1998; Dancygier and Sweetser 2005), so that the negative stance of the conditional clause (*If I finished/had finished this*) is further projected into the ensuing (syntactically higher but spatially embedded) main clause (*I would start/have started another project*). Required maintenance of negative stance throughout the construction suggests that the viewpoint, once expressed, has to extend over all the discourse elements further projected in the scope of the initial stance-setting clause/space. In other words, the epistemically distanced conditional (*If I finished the project next month, I would start another one right away*) has to maintain distanced verb forms in both clauses of the construction, because the viewpoint set up by the conjunction and the first clause needs to be maintained throughout the network. Even a further continuation of the same train of thought (*It would be about adjectives, but I'm not sure it would be as interesting as the previous one*) will also be marked with the same stance. Such examples clearly suggest that viewpoint is a composite concept – past tense alone may mean simply past time, but can also mark epistemic stance

in a construction such as a conditional, set up with a conjunction and clearly referring to the future.

Stance marking via stance verbs (see Kärkkäinen 2003; Dancygier, this volume; Ferrari and Sweetser, this volume) requires a different constructional pattern (*I know that X, he thinks that Y*, etc.), but a similar projection pattern – the stance of the higher space (*know* or *think*) is projected into the lower space (the complement clause). It might then be argued that the viewpoint marked in the higher space of the network is projected down into its daughter space, but the direction is not correlated with syntactic structure – the subordinate conditional clause marked with *if* determines the stance in the main clause, but the stance marked by the main verb in *I know* is then projected to the complement clause. At the same time, the epistemic stance of *know* may affect a preceding subordinate clause, as in *He's going to win, I know*. In such a case, however, the “expectation” stance of the main clause is reinforced and strengthened with the added *I know* expression. The variety of viewpoint/stance configurations cannot be sufficiently described here, but it is important to note that viewpoint structure is correlated more closely with the online sequence of the spaces set up than with patterns of subordination of clauses.

One participant, many viewpoints

The above examples illustrate how grammatical forms and stance expressions can represent one participant’s viewpoint in a network of constructionally linked spaces. Apparently, different means are required when a single participant (e.g. a writer) attempts to represent multiple viewpoints, without adding discourse participants. Such viewpoint shifts through the choice of deictics were described in Rubba (1996). Recently, instances of shifts marked solely through the choice of lexical descriptors were described in Dancygier (2005) as decompression for viewpoint, where an otherwise unified concept is decompressed to create additional loci for the profiling of additional viewpoints.

- (1) The Mississippi was two rivers. They lay right beside each other, but flowed in opposite directions. The steam boats, the fancy Golden Age hotels, the scenic bluffs and gift shops were all going one way, while the river on the charts, with its tows, grain elevators . . . was going quite another. I had done my share of travelling on the first river, but it was a cute irrelevance compared with the deep, dangerous, epic power of the real Mississippi.

(Jonathan Raban, *Old Glory*)

In this example, the writer Jonathan Raban retains his single-handed grasp of narratorship, but presents the Mississippi river from two perspectives. His use of the expression *two rivers* suggests a decompression of a single concept into

two construals – each one representing a different viewpoint. The narrator's own attempt to do justice to the power of the river is contrasted with the glitz of the touristy side of it, which is the only construal accessible to those unwilling to look for the day-to-day reality. In such cases, where viewpoint is not associated with any grammatical forms and not connected to epistemic stance, the profiling of viewpoint is a matter of competing construals, adding depth to the narrative, but not requiring additional participants. What such cases share with the epistemic stance instances described above is the consistent construction of a mental space network from the perspective of one, primary participant.

Raban's travel narratives consistently maintain his viewpoint. Even when decompression signals the availability of other viewpoints, they are still construed from the point of view of the narrator. We can note that even in the cases of decompression, as in the Mississippi example, the "other" (tourist) viewpoint is not attributed to any participant of discourse context. Rather, it is still construed as the narrator's comment on various perspectives available. The two construals are thus representative of different viewpoints, as seen from the narrator's viewpoint; it is a matter of narrative construction, rather than of representation of discourse. Similar types of choices are apparently being made by the participants in the experiment conducted by Parrill (this volume), where gesture indicates varying alignment of the teller with either a participant in the story recounted or an observer. In the instances described by Parrill, all the gesturing modes seem natural by virtue of representing some identifiable aspect of the linguistic narrative's contents; but Raban's narrative choices are rather conspicuous because they rely entirely on the wording. This may suggest that the gesturing options naturally structure the multiple viewpoints available in the act of storytelling, but do not allow or require stylistic acrobatics of the kind found in Raban's texts.

Multiple participants

The examples discussed by Parrill are interesting in that the gestures aligned with the character evoke story participants rather than discourse participants. At the same time, other aspects of gesture move the teller from "inside" the story space to its "outside" – for example, a narrator may go from gestural representation of story events (climbing a ladder) to holding eye contact with the addressee and holding out a palm-up hand (possibly meaning "see, I've made my point") (McNeill 1992, 2005). However, viewpoint phenomena become significantly more complex with the explicit profiling of participants other than the speaker. One such case is a narrative report in which various subjectivities have to be profiled to represent the events of the story from a chosen narrative or reporting perspective. Here the choices are structured by at least two

major viewpoint options. In one, the teller maintains one coherent viewpoint, deictically connected to herself/himself, and incorporates other available participant viewpoints into the one, top-level viewpoint space. Examples of this option include many varieties of represented speech and thought, or the use of negation. Further structural complexity is added by varying degrees of distance across viewpoint spaces, to yield a variety of stylistic and cognitive effects. For example, the speaker may use negation to initiate a construal wherein something absent should be present (as in *We have no bread for lunch*), or also add a distancing stance expression (as in *I'm afraid we have no bread for lunch*, or *I think we have no bread for lunch*).

The second strategy is to use the profiling of various subjectivities in an attempt to either overtly coordinate viewpoints or to open them to negotiation – along the lines described by Verhagen (2005). In an example quoted in Verhagen, the speaker objects to the hearer's description of a person as "happy" by negating the utterance and proposing a different description, as in *Mary is not happy. On the contrary, she is quite upset*. Such an utterance relies on various constructional forms (negation, the expression *on the contrary*) to explicitly evoke the interlocutor's statement and object to it at the same time. The network thus profiles two viewpoints, not just the one of the speaker. Crucially, intersubjective construals are similarly structured in American Sign Language (ASL), as discussed in Janzen and Shaffer (2008). Each of the major strategies outlined involves various constructional choices, from negation, through adverbial constructions, to a broad range of constructions known as "represented speech and thought."³

There are many such strategies and options described throughout this volume, and the profiling of multiple subjectivities is dealt with in different ways by the various viewpoint constructions discussed. Even though the chapters focus on various communicative modalities (sign language, gesture, spontaneous discourse, or fiction narratives), there are mental space structure commonalities across various modes which shed some light on the nature of viewpoint phenomena in general.

Many of the constructions discussed throughout this volume have the speaker/writer/signer report events allocated to a specific reported or narrative space. The "reported" or "narrative" space (I will refer to it as "narrative," since most reports are mini-narratives as well) is thus separate, and often temporally and epistemically distanced from the "reporting" or "narrating" space. The crucial aspect of such constructions is thus the way in which the teller negotiates the difference between the two spaces – either treating them as distanced or as partially, or perhaps even fully, blended. The degree of integration depends crucially on whether the teller is placed in the narrative space, along with the other participants, or in the narrating space, separate from them. These varying viewpoint configurations yield different constructional forms.

For example, the integration of narrative and narrating viewpoint can take a fully embodied form in sign language. As the examples discussed by Shaffer (this volume) suggest, a signer may use her body to allocate the two viewpoints (narrating and narrated) in space – classically, the narrating viewpoint involves the signer facing forwards and making eye contact with the addressee, while a narrated viewpoint may involve a bodily rotation for role shift to a viewpoint character's space, and accompanying cessation of eye contact. The signer then uses bodily movement and signing to render the varying contributions of the representing and represented participants and maintain a coherent viewpoint structure within the narrated or narrating space. In the data discussed by Parrill, for comparison, even when the spoken text is third person narration (with no salient linguistic adoption of character viewpoint), accompanying gesture may have different effects. It may contribute to the distance between the spaces, by maintaining the narrator's perspective, or it may integrate them to some degree, by adopting a gesturing viewpoint of a participant inside the story being told. That duality of viewpoint is always relevant in reporting a sequence of events, but the particular ways in which discourse may highlight just one viewpoint or profile two is a matter of interest.

When the body position and gesture play a role independent of linguistic discourse (spoken or signed), the divergent or convergent allocation of viewpoint is achieved in different modalities. When viewpoint is negotiated within a single modality, such as narrative fiction, the contrast or convergence between the narrating space and the narrative space is achieved through interesting choices of linguistic forms. As Nikiforidou (this volume) shows in her chapter, narrative discourse may create constructions that are not found elsewhere and are specifically targeting the negotiation of viewpoint. Nikiforidou explains the occurrence of so-called *was–now* constructions, which combine the use of past tense and present time adverbials in one sentence, as in (2), where the narrator introduces his main character.

- (2) He now lived, for the most part, retired in the country, with one sister, for whom he had a very tender affection.

(Henry Fielding, *Tom Jones*)

Nikiforidou explains the constructional processes whereby a sentence marked with conflicting temporal signals is interpreted. Such a construction creates a viewpoint blend: the choice of tense maintains the selected temporal viewpoint affecting the narrative space (it is past with respect to the moment of telling), while the choice of time adverbial foregrounds the continued presence of the narrating space, where the (disembodied) narrator tells the story. In spite of the high degree of integration, then, the viewpoints of the two spaces are maintained without major shifts.

At the same time, it is not uncommon for the teller to yield the viewpoint to the participants, as in narrative fragments retaining pieces of dialogue. This

may be seen in speech as well as in writing, fiction or non-fiction, and allows the teller to step back and give the participant full voice, as well as focusing the viewpoint on them. There are also various “mixed” narrative modes (such as forms of free indirect discourse) which maintain both the teller’s and the participants’ viewpoints alongside each other. However, there is also a less explored viewpoint pattern wherein the teller’s space dominates throughout, at least as a shared deictic set-up. Interestingly, two chapters here explore the details of such a configuration, in different modalities. Vandelanotte’s analysis proposes a category of speech and thought representation, which, rather than blending the narrating space with the narrative one, evokes the discourse of participants, and subordinates it to the deictic viewpoint of the teller.

(3) “What time does your train go?”

“Heavens, what a question.” She glanced laughingly at John, sharing the joke. “You don’t come to meet me, and the first thing you ask is, *when am I going?* Half-past six, if you must know. What would you like to eat?” she added, as one of the ladies in print overalls came and stood by them.

(Philip Larkin, *Jill*; italics added)

The clause in italics is a rendition of the question asked in the first line. Characteristically for this reporting mode, the question form and the tense are repeated from the original question, but the speaker changes “you” into “I,” thus blending the report with her viewpoint represented throughout the response to the question. As a result, two viewpoints are aligned into the deictic viewpoint of one of the participants. In fact, such a viewpoint configuration is similar to the phenomenon of “mental space rotation,” identified by Janzen (this volume) in the discourse of ASL users, which allows the teller to adopt the viewpoint of a story participant while maintaining the narrating viewpoint at the same time.

Viewpoint negotiation and subjectification

There is also the question of whether viewpoint needs to be aligned with a participant (real or profiled for the purposes of the discourse, as an “omniscient narrator”). As Tobin and Israel (this volume) show in their analysis of irony,⁴ there is an important difference between shifting to the viewpoint of another participant and shifting to a viewpoint involving a different space or an entire network of spaces. In numerous studies of viewpoint, especially in narratological studies of focalization, viewpoint is understood as aligned with a participant, whose mental or visual response to the events is the source of construal of the narrative events. But in the examples studied by Tobin and Israel, viewpoint shifts to one of the spaces in the network, while several remain available. Tobin and Israel consider some rather complex examples of irony,

but even in the simplest instances the network is set up and used similarly. Irony is often illustrated with an example wherein the speaker has planned a picnic and wakes up to see rain outside; what she might say then is something like, *Oh, what a wonderful day!*, which is blatantly not descriptive of the actual situation. The speaker is thus relying on the space profiling rain, but additionally sets up a “nice day” space that was imagined when the picnic was planned. The ironic viewpoint consequently relies on the contrasted topology of both spaces and sets up a viewpoint space wherein the contrast can be seen and the true situation, as well as the speaker’s sentiments, is available. As Tobin and Israel thus show, viewpoint may mean alignment both with a mental space (e.g. the one that is really true) and with its topology (as part of a network involving other mental spaces profiling different situations). Thus a construal such as irony involves a very specific configuration of mental spaces and the conceptualizer’s choice of the specific space as the locus of viewpoint relative to the larger meaning network. The analysis also opens many questions as to the mechanisms of viewpoint shift in other staples of pragmatic analysis, such as implicature. Importantly, Tobin and Israel’s return to the line of mental space analysis of pragmatic phenomena, started by Fauconnier in his discussion of presupposition (1985 [1994]), opens a new avenue in viewpoint studies.

Most of the studies in this volume make it clear that multiplicity of viewpoints can properly be explained through postulating an added layer of viewpoint – a “bird’s-eye view” space from which the other profiled viewpoints can be evaluated and construed. Contrary to the common understanding of viewpoint, tied to a “viewer,” aligning viewpoints with spaces allows us to reveal viewpoint hierarchies and relationships, regardless of the number of participants. Tobin and Israel are perhaps more explicit on this issue, but the need to consider viewpoint phenomena in terms of networks of spaces rather than viewers permeates the approaches presented here.

While Tobin and Israel stress the importance of considering viewpoint configurations set up by a participant, Dancygier (this volume) stresses the constructional aspects of the mechanisms of viewpoint negotiation. The examples of negation discussed in the chapter rely on the situations where the speaker is not only including the hearer’s stance in the network, but is actively rejecting it, as in (4).

(4) A: What do you expect Tim to do now?

B: I don’t expect anything. He has never helped me with anything.

Speaker B overtly disagrees with A as to the stance to be taken with regard to the subject matter under consideration. There are many complex constructional options available in such cases, involving stance and negation among others, which reveal subtle differences in viewpoint-taking among the participants

profiled. Interestingly, the chapter by Ferrari and Sweetser reveals the increasingly subjectified meaning structures of stance constructions, often involving meaning historically incorporated from context that was not originally conventional or constructional at all. Crucially, the two chapters jointly show various degrees of entrenchment of stance and viewpoint, also uncovering the range of lexical and constructional means of representing stance and viewpoint patterns. Also, they raise further questions regarding the expressions of intersubjectivity and the contrast between constructional and subjectified expressions of stance.

Moreover, as Narayan shows in her chapter, viewpoint negotiation may involve a complex exchange carried out at many levels of construal, starting with the orientational and directional perception of one's own body. Most crucially to Narayan's argument, negotiation of viewpoint requires that in order for visual viewpoint to emerge, the language users' bodies need to be relied on as the basic resource. The embodied viewpoint thus provides the ground for the visual viewpoint, which in turn provides the opportunity for the construal of the whole scene and the events within it. In a way, what Narayan's analysis is suggesting is that in order to come up with a narrative explaining the scene we attempt to construe, we need to begin with positioning the body mentally in a way that provides affordances for visual construal to emerge, thus allowing more complex aspects of the construal to enter the scene. Such a progression, from the body, through vision, to the narrative is further confirmed in the ASL data analyzed by Shaffer and Janzen, where the signer's body posture in telling a story essentially reflects the visual and embodied features of the scene.

Further directions

What future research should probably address in more detail is the nature of mechanisms yielding the configurations of viewpoint described throughout this volume. For example, earlier work by Dancygier (2005, 2007, 2008, 2011) has outlined the mechanism of viewpoint compression, which allows two or more viewpoints available in the network to be compressed to the higher space. One such example can be a third person narrative that represents a character's discourse or thought without using direct discourse. Thus the narrative passage in (5) represents the character's words, which could be rendered as in (6).

- (5) Mr. Pomfret didn't mention references. His sole concern was the nature of her past duties. Had she typed, had she filed, taken shorthand? . . . Also she was expected to brew the coffee; he hoped that wasn't a problem.

(Ann Tyler, *The Ladder of Years*)

- (6) Have you typed? have you filed, taken shorthand? . . . Also, you'll be expected to brew the coffee – I hope it's not a problem.

The tense and pronoun choices in (5) suggest an overall, text-wide viewpoint space (either profiling a teller or not), from the perspective of which all the embedded narrative spaces are viewed – hence past tense and third person. Crucially, all the reports of discourse taking place in the lower narrative spaces are structured from the perspective of the highest viewpoint space, but the grammatical integration of the lower space, as roughly represented in (6), also involves a compression of the entire space hierarchy, so that what was said is aligned with a local narrative viewpoint and becomes a part of the flow of the story. This requires that the lower space viewpoint is compressed into the higher-level viewpoint. Compression thus enables the continuity and viewpoint coherence of an extended narrative, while maintaining the (compressed, but recoverable) viewpoints of the lower spaces.

Compression may account for a number of viewpoint configurations described in this collection. It is useful in explaining both the subjectified cases discussed by Ferrari and Sweetser, where the compressed construal is now lexicalized or grammaticalized to the point where contributing spaces are no longer overtly available. *Good-bye* historically involved two mental spaces, that of a wish *God be with you*, and the canonical discourse context in which that wish was embedded, namely at a social parting of speaker and addressee, but only the discourse context mental space structure is now referred to. But compression also explains the way in which instances of irony or viewpoint negotiation allow a coherent understanding of the entire network, including the choice of the highest viewpoint space, from which the participating spaces, even the contrasting ones, need to be viewed. In other words, as Tobin and Israel argue, irony could not be processed successfully as a viewpoint taken by one space or affecting just one space. Instead the emerging viewpoint needs to have the entire network in its scope and blend the inconsistencies to create the compressed ironic viewpoint.

Compression, as exemplified here, may depend on decompression. In Example (1) above, the river is decompressed based on viewpoint, but both spaces thus created are compressed up to the overall story viewpoint, where it becomes clear that the writer is constructing the dual viewpoint for the reader, rather than bringing it from some independent discourse. Compression also seems to take part in all the narrative cases discussed in this collection. Pointing out the compression patterns represented by each of the analyses would exceed the limits of this discussion, but let me just mention the case of ASL “mental space rotation” discussed by Janzen. Janzen describes signers as simply assuming a character’s viewpoint (without the partial bodily rotation of canonical role shift), and interacting gesturally – pointing, for example – from that character’s viewpoint; the narrator can switch characters and must thus rotate the imagined construal of the signer’s physical space to represent a different character’s physical viewpoint. If two imagined characters are facing

each other, the result is a 180-degree virtual shift in interpretation of the signer's body; when one character points to the right, the other character would point left to the same imagined object or location. Note that compression is crucial here. These shifts between locally conflicting gestural patterns can only be seen as "rotation," rather than complete shift, if we can understand them as different character's views of a single scene or sequence of events – that is, conflicting viewpoints become coherent because the discourse maintains a higher narrative space where the viewpoint of the signer is located, and allows other viewpoints to be compressed temporarily with it.

Narayan's chapter provides an interesting related example of cross-speaker or intersubjective gestural viewpoint maintenance. In discussing and gesturing about a picture that only one of the participants can see, the speaker who lacks direct visual access initially uses gestural viewpoint parasitic on that of the speaker who can see the picture, showing in gestural viewpoint the degree to which his cognitive access is embedded in hers. But as he discerns something about the depicted events that she had failed to figure out, his gestures switch to a model as if he were looking directly at an imagined picture, reflecting gesturally his diminished cognitive dependence on her mental space.

We will, of course, be searching for further mechanisms that can explain the ways in which an overwhelming variety of viewpoint configurations and discourse goals can be smoothly processed. What this volume has definitely done is to show the rich variety of viewpoint patterns, in various modalities, but also to stress the usefulness of the mental spaces framework in bringing these different choices into a shared focus. We will no doubt find more interesting and unexpected viewpoint patterns, but we can use the common ground established by the discussion in this volume in arriving at an understanding of the phenomenon of viewpoint in its entirety.

References

- Coulson, Seana and Todd Oakley. 2005. Blending and coded meaning: literal and figurative meaning in cognitive semantics. *Journal of Pragmatics* 37, 1510–36.
- Cutrer, Michelle. 1994. Time and tense in narratives and everyday language. Ph.D. dissertation. University of California, San Diego.
- Dancygier, Barbara. 1998. *Conditionals and Prediction: Time, Knowledge and Causation in Conditional Constructions*. Cambridge University Press.
2005. Blending and narrative viewpoint: Jonathan Raban's travels through mental spaces. *Language and Literature* 14:2, 99–127.
2007. Narrative anchors and the processes of story construction: the case of Margaret Atwood's *The Blind Assassin*. *Style* 41:2, 133–52.
2008. The text and the story: levels of blending in fictional narratives. In Todd Oakley and Anders Hougaard (eds.). *Mental Spaces in Discourse and Interaction*. Amsterdam and Philadelphia: John Benjamins, 51–78.
2011. *The Language of Stories: A Cognitive Approach*. Cambridge University Press.

- Dancygier, Barbara and Eve Sweetser. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Fauconnier, Gilles. 1985 [1994]. *Mental Spaces*. 2nd edn. Cambridge University Press.
1997. *Mappings in Thought and Language*. Cambridge University Press.
- Fillmore, Charles J. 1990a. Epistemic stance and grammatical form in English conditional sentences. *Comparative Literature Studies* 26, 137–62.
- 1990b. The contribution of linguistics to language understanding. In Aura Bocaz (ed.), *Proceedings of the 1st Symposium on Cognition, Language, and Culture*. University of Chile. 109–28.
- Fleischman, Suzanne. 1989. Temporal distance: a basic linguistic metaphor. *Studies in Language* 13, 1–50.
- Janzen, Terry and Barbara Shaffer. 2008. Intersubjectivity in interpreted interactions: the interpreter's role in co-constructing meaning. In Jordan Zlatev, Timothy P. Racine, Chris Sinha, and Esa Itkonen (eds.). *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam/Philadelphia: John Benjamins, 333–55.
- Kärkkäinen, Elise. 2003. *Epistemic Stance in English Conversation*. Amsterdam: John Benjamins.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.
2005. *Gesture and Thought*. University of Chicago Press.
- Rubba, Johanna. 1996. Alternate grounds in the interpretation of deictic expressions. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 227–61.
- Sanders, José and Gisela Redeker. 1996. Perspective and representation in speech and thought in narrative discourse. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 290–317.
- Semino, Elena and Mick Short. 2004. *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.
- Sperber, Dan and Deirdre Wilson. 1991 [1981]. Irony and the use-mention distinction. In Steven Davis (ed.). *Pragmatics: A Reader*. Oxford University Press, 550–63. (Originally in Peter Cole [ed.]. *Radical Pragmatics*. New York: Academic Press, 295–318.)
- 1995 [1986]. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Sweetser, Eve. 1996. Mental spaces and the grammar of conditional constructions. In Gilles Fauconnier and Eve Sweetser (eds.). *Spaces, Worlds, and Grammar*. University of Chicago Press, 318–33.
- Vandelanotte, Lieven. 2009. *Speech and Thought Representation in English*. Berlin/New York: Mouton de Gruyter.
- Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford University Press.

Notes

INTRODUCTION: VIEWPOINT AND PERSPECTIVE IN LANGUAGE AND GESTURE, FROM THE GROUND DOWN

- 1 By *unidirectional*, it is not meant that all meaning change is towards greater subjectivity. When a word for “feather” comes to mean “pen” (like English *pen*, from Latin *pinna*), there is no change in the referential relation to the Ground. The claim is that semantic change – particularly when associated with grammaticalization – often results in increased subjectivity, but *does not* result in decreased subjectivity.

1 IRONY AS A VIEWPOINT PHENOMENON

- 1 Indeed, in some cases an ironic interpretation may be built directly into the meaning of a linguistic construction – for example, the exclamation *big deal!* normally expresses a judgment of insignificance (i.e. that something is not, in fact, a “big deal”), and the double affirmation *yeah, yeah* is normally used to express snide disagreement. Israel (2006) treats examples like these as involving a sort of short-circuited implicature that allows a complex mental space configuration to be associated directly with a linguistic construction.

2 SUBJECTIVITY AND UPWARDS PROJECTION IN MENTAL SPACE STRUCTURE

- 1 Actually this is far too simple an analysis, since deictic meanings are not purely spatial, but are extended to social uses and other abstract meanings (Hanks 1990 provides a particularly interesting study of this; see also Rubba 1996); but we concur with Levinson (1983) and others that one can treat the spatial deictic meanings as central. Some of the complexities of deixis are discussed in the Introduction to this volume.

3 NEGATION, STANCE VERBS, AND INTERSUBJECTIVITY

- 1 For a discussion of the degree of subjectification of stance expressions such as *don't know*, see Bybee and Scheibman 2007.

5 *MAYBE WHAT IT MEANS IS HE ACTUALLY GOT THE SPOT: PHYSICAL AND COGNITIVE VIEWPOINT IN A GESTURE STUDY*

- 1 This study was supported by a Doctoral Dissertation Research Improvement Grant from the National Science Foundation.
- 2 The keyboard was on a table within Speaker's reach.
- 3 I follow the form in Coulson (2001) by laying out these blends as tables; every column represents a Mental Space, and items that share a row have mappings between them. This format lets one lay out the complex, rich structure of spaces involved in the blend more clearly than is possible with the traditional bubbles and lines; it also highlights the frames structuring spaces. As such, it is more suited to the current endeavor.
- 4 Since Speaker and Listener are discussing a two-dimensional image, a 180-degree rotation is equivalent to horizontal mirroring. I am not making any claim about which of these Listener does; the data do not disambiguate.
- 5 Or a transformation of our frame of reference – but that commonly applies to three-dimensional objects that can be construed as having their own inherent frame of reference, and is not relevant to the current data.
- 6 Which is consistent with the orientation of the car in the picture, if not with the motion trail and the word SCREECH.

6 REPORTED SPEECH AS AN EVIDENTIALITY STRATEGY IN AMERICAN SIGN LANGUAGE

- 1 A video relay service (VRS) is a telecommunication service that allows people who are deaf or hard of hearing to communicate over the telephone with hearing people using video and a signed language interpreter.

7 TWO WAYS OF CONCEPTUALIZING SPACE: MOTIVATING THE USE OF STATIC AND ROTATED VANTAGE POINT SPACE IN ASL DISCOURSE

- 1 We might note, though, that a 180-degree shift in viewpoint may not always be the case. If it were so, we might suggest that the spatial scene is simply reversed rather than rotated. It appears, however, that suggesting such a reversal is insufficient. Signers describe situations and tell about interactions between people who are spatially located anywhere within their viewable field. In one narrative, for example, the signer depicts an interaction between one person and others positioned behind him over his left shoulder, and several other described interactions are between people who are side by side. It is true that in these cases the depicted signers must still look at each other, but their spatial orientation is not directly across from one another; thus the shift in viewpoint reflects their relative positioning in space.
- 2 A detailed description of this narrative can be found in Janzen (2004, 2005), the specifics of which will not be dealt with here.
- 3 At times the construal seems to be that of a generic plural reference – that is, of all four people who are in the car. There are no grammatical markers to signal this, but there is no reason to think that they would not all be watching and experiencing

the same thing; and in terms of grammar, pronouns in ASL are often unexpressed if sufficiently topical (see Janzen 1998; Slobin 2006).

- 4 Lillo-Martin (1995) goes so far as to say that the physical shift in space operates as the verb in the construction, which indicates that it is recognized by researchers as an overt marker with at least morphological status, but this analysis will not be debated here.
- 5 Although note that signers use eye gaze in complex ways – for example, signers direct eye gaze to elements in their real world articulation space in that they often glance at their hands as they position them in space; but also, in descriptions of conceptualized (non-present) scenes and in narratives involving conceptualized actions and events, signers look around these mental scenes. Despite this, signers maintain a close visual connection with their addressees. Altogether, eye gaze tells us much about the mental spaces being evoked, and, further, about blended spaces (Dudis 2004).
- 6 Note, however, that in Janzen (2008) I give evidence that such iconic mapping is a mapping not of actual real-world features, but of conceptualized form determined by subjective construal.

8 THE CONSTRUCTIONAL UNDERPINNINGS OF VIEWPOINT BLENDS: THE *PAST + NOW* IN LANGUAGE AND LITERATURE

- 1 The term was originally coined by Bally (1912) as “style indirect libre.”
- 2 In addition to the original Construction Grammar (Fillmore *et al.* 1988; Kay and Fillmore 1999), Cognitive Grammar (Langacker 1987, 1991, 2000), Cognitive Construction Grammar (Goldberg 1995, 2006), and Radical Construction Grammar (Croft 2001) are prominent versions of the constructionist paradigm.
- 3 From these, at least non-anaphoric reflexives also occur in non-literary language with similar semantic-pragmatic effects. This would therefore be another construction shared by FIS and non-literary discourse.
- 4 See, for example, Sanders and Redeker 1996; Sweetser and Fauconnier 1996; Dancygier 2005; Dancygier and Sweetser 2005; Dancygier and Vandelanotte 2009.
- 5 The research reported in sections 8.2 and 8.3 of this chapter also appears in different form in Nikiforidou 2010. Readers may refer to that article for an extended corpus analysis complementing the blending analysis given here.
- 6 In fact, as evident in some of the examples, the verb and the deictic can be separated from each other regardless of their relative order and with the same semantic effect.
- 7 The interpretation is not necessarily predictable, however, since profiling might be restricted in principle to the starting phase of the process.
- 8 As is clear from all the examples of *Past + now* discussed, of interest are only those meanings of the deictic that can be characterized as temporal (even in an extended sense). *Now* also has a discourse function (cf Berman and Slobin 1994; Ariel 1998), which does not concern us here (an instance of this discourse marker *now* appears for instance in Example [22]).
- 9 In Fillmore *et al.* (1988) classification of idioms, the category of formal (or schematic) constructions refers to the patterns that can be filled by a relatively free range of syntactically and semantically appropriate words or phrases. Such expressions, while semi-regular or productive, are crucially associated with syntactic, semantic, and often pragmatic (or, in this case, discoursal) properties, which cannot be predicted

from the general syntactic or semantic rules of the language, or from an all-purpose pragmatic component; as such, they must be directly assigned to the construction as a whole.

- 10 Eve Sweetser (personal communication) has in fact offered a motivation for this particular extension. The narrator of a fictional tale is adopting the imagined viewpoint of a fictional narrator for whom the narrated events are past – naturally past, since it is hard to imagine a real or a fictional person who regularly sees the future (see also the discussion before on first person narratives). But views to the effect that the absence of any linguistic indicators of a narrating agent (apart from person and tense) cause the past to lose its deictic (past-to-me) designation and become a marker of narrativity or serial ordering have been also voiced (e.g. Adamson 1995: 219). A homonymy treatment appears, however, counterintuitive given the pragmatic and functional relatedness of examples like (15) or (16) with an example like (17).
- 11 One of the reviewers pointed out that “coercion” may be too strong a term for this kind of semantic accommodation. However, in the constructional literature, different kinds of syntactic-semantic discrepancies have been treated as instances of coercion (e.g. both the adjustment of noun reference in the plural construction, as in *two beds*, and the more pronounced conflict in *some blanket*). So, although “coercion” as a term may imply strong conflict, in essence it is used broadly for different types of discrepancies and constructional unification (cf De Swart 1998; Michaelis 2005a).
- 12 Mirjam Fried (personal communication) suggested that the sudden drop in numbers could conceivably be due to changes in the text-type structure of the magazine, with fewer narrative texts after the 1960s. Although this is certainly something that should be checked out, the difference between the 1940s and the neighboring decades still remains significant.
- 13 Bray (2007) tests experimentally the bias for a shifted reading in the succeeding co-text of an FIS sentence.
- 14 They use “focus” and “focalization” in a technical sense (following Genette 1980) to refer to the story as the “vision” of a character in it (as opposed to the voice of the narrator).
- 15 It is important to note that the viewpoint issue is different from the “dual voice” discussion surrounding FIS (Vandelanotte, this volume). The latter refers to whether FIS expresses simultaneously the voice of the author and the character (“speaking at the same time,” as suggested by Vološinov [1973]), or whether “the so called merging of a narrator with a character can be more adequately described as the absence of a narrator” (Galbraith 1995; also Banfield 1982). Both these views are consistent with viewpoint being in the embedded space (regardless of whether the narrator actively adopts it or is completely effaced), while a linguistic analysis proper cannot, in my view, contribute significantly to their discrimination (but see Bray 2007 for an experimental attempt).
- 16 However, according to the authors, a marked/unlikely reading is also possible, where the implied narrator uses the referential terms as his/her own labels.
- 17 Fauconnier and Turner (2002: chapter 6) discuss in detail some of the most important vital relations, which, besides time, include cause–effect, change, space, part–whole, representation, role, analogy and disanalogy, similarity, category, intentionality, and uniqueness.

- 18 In Langacker's (1991, 2000, 2002) framework, we would say that the space of the narrator receives a maximally subjective construal (being offstage and backgrounded), while the space of the narrated events, which are on stage and profiled, receives a maximally objective one (with minimal interference from a narrating agent).
- 19 As Ferrari and Sweetser (this volume) show, the viewpoint space can be modeled as one of the spaces in the Basic Communicative Space Network, which includes further a speech act space, the Ground or Real Space and an epistemic space or spaces. The BCSN model allows for an explicit representation of FIS, which requires, as argued here, the collapse of two formerly distinct spaces. For a detailed analysis of viewpoint compression in narratives, see also Dancygier 2005.
- 20 In Langacker's terms, then (see note 16), we might say that the construction appears to realign the subjective-objective construal. The narrated events now receive a subjective interpretation (approached from the vantage point of someone who is close or even participating in them), while a/the narrator, by virtue of being aligned with the character, gets "onstage," thus receiving an objective construal.

9 “WAIT TILL YOU GOT STARTED”: HOW TO SUBMERGE ANOTHER’S DISCOURSE IN YOUR OWN

- 1 As is well known, what is *presented as* “reported” need not actually have been uttered before; cases in point where no prior “original” is involved include negated or modalized uses, as in *I didn’t say X* or *I might have said Y (but I didn’t)*.
- 2 Examples from the Cobuild corpus were obtained via remote log-in from the University of Leuven, where the author is a research fellow in the English linguistics section, and are reprinted here with the kind permission of HarperCollins publishers.
- 3 As an illustration of this point, consider the constructed example below of a direct thought involving direct address (*you*), and its corresponding renderings in FIST in third and first person narrative contexts. The central difference, and the reason why *you* becomes *her* in the free indirect renderings, pertains to the shift to the represented speech event (in which the woman is present and hence addressable) in direct thought, vs the lack of such a complete shift (resulting in the absence and inaddressability of the woman).

Direct thought: I don’t love you anymore (he thought as he kissed her).

Corresponding free indirect thoughts:

Third person: He didn’t love her anymore (he thought).

First person: I didn’t love her anymore (I thought).

- 4 In more narrative, typically past tense contexts, the situation is less clear-cut, since the tense then tends to be adapted to the past tense narrative context. Past tenses in stretches of DIST in narrative contexts, as in examples (8)–(12) above and (16) below, can be treated on a par with the tense of the main narrative.
- 5 As recognized by both Verhagen (2005: 5, note 4) and Traugott (2010: 33–4), yet another use of the term is that of Nuyts (2001, 2005), for whom an “evaluation is . . . intersubjective if [the issuer] indicates that (s)he shares it with a wider group

of people, possibly including the hearer” (Nuyts 2005: 14); this contrasts with subjective evaluations, which are strictly the issuer’s own responsibility.

CONCLUSION: MULTIPLE VIEWPOINTS, MULTIPLE SPACES

- 1 Also resumed in Fauconnier 1997.
- 2 Such an analysis elaborates the distinctions made earlier in Reichenbach’s model of English tense usage. However, that model is primarily focused on temporal relations, while a mental spaces analysis extends the same mechanisms into the representation of epistemic stance through tense forms.
- 3 Represented speech and thought constructions were analyzed in terms of mental spaces and viewpoint in the pioneering article by Sanders and Reckter (1996). They have been discussed in a substantial body of literature, which cannot be fully reported here, but the most recent exhaustive descriptions can be found in Semino and Short (2004) and Vandelanotte (2009).
- 4 This chapter leaves no room for extensive comparisons, but the interpretation of irony discussed here substantially changes the generally accepted view of irony, as proposed by Sperber and Wilson (1991 [1981], 1995 [1986]).

Index

- absolute spatial language, 13–14
access
 cognitive, 6, 7, 52, 69, 79, 205, 215, 230
 perceptual, 6, 7, 10, 141
Adamson, Sylvia, 177, 178, 179, 183, 185, 193
alienation, 34, 42
alternative spaces/alternativity
 in conditionals, 74
 in negation, 69–72, 79, 82–7, 214
American Sign Language, 14–16, 124, 139, 174, 190–1, 206, 224, 226, 228, 229–30
rotation (mentally rotated space), 16, 123–4, 156–8, 160–9, 171–2, 190–1, 206, 226, 229–30
static space, 156, 157–8, 159–60, 161, 162, 164, 165, 166–9, 171–2
argumentation, 73, 77, 87
articulation space, 156, 159, 160, 161, 162, 164, 165, 166, 167, 168, 169, 170, 171, 172, 234
attention flow, attention shifts, 32

Banfield, Ann, 7, 177, 179, 184, 185, 193, 198, 200, 202, 203, 209
Base Space, 31, 34, 39, 40, 55, 61, 74, 101, 102–5, 119, 123, 127, 189, 198, 205, 211
Basic Communicative Space Network, 9, 18, 48–51, 55, 64, 65, 66, 236
Benveniste, Emile, 184, 185
blends, blending, 4, 6, 9–11, 13–14, 31, 37, 38–9, 40, 65, 73, 97, 101–2, 113–35, 139–54, 177–94, 201, 202–3, 210–11, 220, 224, 225, 226, 229
Conceptual Integration Theory, 139, 141
Real Space blends, 14–16, 102–5, 114, 123–4, 127, 130–2, 139, 141, 146, 159, 160
 reified blend, 123–5
body shift, 156, 157–8, 159–60, 161, 162, 165, 166–8, 169, 171–2
Brinton, Laurel, 75, 178, 179, 183
Bybee, Joan, 58, 59, 65
coercion, 177, 179, 185, 186, 188, 189, 194, 235
Cognitive Grammar, 27, 81, 139, 234
comparative discourse, 157, 160, 169
Conceptual Integration Theory. *See* blends, blending
conceptualized space, 156, 160, 161, 164, 168, 170–1
construal, 1, 2, 4, 5, 7, 8, 9, 10, 31, 35, 38, 42, 43, 48, 51, 65, 73, 79, 86, 89, 90, 114, 119, 132, 134, 160, 187, 191, 211, 212, 219–21, 222–3, 224, 226–7, 228, 229, 233, 234, 236
constructed discourse, 141, 149
Construction Grammar, 66, 177, 178, 179, 186, 194
constructional compositionality, 88–9
Content Space, 48–51, 53, 56, 59–60, 62, 66
counterfactuality, 35, 75
Croft, William, 82, 178
Cutrer, Michele, 5, 31, 48, 54, 220–1

Dancygier, Barbara, 5, 6, 34, 47, 48–50, 59, 71, 72, 74, 75, 76, 79, 80, 85, 89, 178, 191, 199, 211, 212, 221, 222, 228
Dasher, Richard, 8, 12, 47, 73, 154, 212, 213
Davidse, Kristin, 201
de Haan, Ferdinand, 143, 144, 152, 153
decompression, 35, 38, 211, 215, 222, 223, 229
definite article, 6, 51–2, 178
deixis, 4–5, 9–12, 14, 152–3
 deictic center, 9, 11, 13, 34, 179, 189, 200–1, 202, 203, 208, 215
 deictic displacement, 9
 deictic field, 9–11
 deictic shift, 200–1, 203, 210, 211, 215
 spatial deixis (social uses of), 12
 temporal deixis/deictic, 4–5, 177, 184, 201
desubjectification, 44, 215, 216
determiners, 5

- direct speech and thought, 100, 101, 105, 198–9, 200–1, 209
discourse genre, 157, 158, 162, 169
discourse status, 106–9, 110
distance, 50, 59–64, 75, 98–9, 100, 143–4, 147, 179, 185–6, 191, 192–3, 215, 220
discourse distance, distanced discourse, 76, 210–12, 215–16
distance in mental space networks, 192, 224, 225
epistemic distance, 7, 79, 191, 211
ironic distance, 25, 38–9, 202
spatial distance, 7, 154, 164, 193
distancing indirect speech and thought, 38, 76, 166, 198–9, 202–16
dual voice, 202, 235
Dudis, Paul, 13, 102, 114, 116, 119, 139, 141, 146, 149, 153, 160, 163
echo, echoic mention, 26, 29, 72, 85, 209, 211
embedded viewpoint, 10
embodied, embodiment, 1, 3–4, 6, 13, 17, 114, 134, 171, 225, 228
Epistemic Space, 9, 48, 50, 51, 54, 55, 59–61, 63, 64, 65
Event Space, 31, 33, 42, 114, 146, 149
evidentiality, evidentials, 5, 6, 139–41, 142–54, 160, 208
reported speech evidentials, 140, 144–53
external perspective, 99, 100–1

Fauconnier, Gilles, 3–4, 5, 13, 17, 27, 31, 35, 48, 54–5, 66, 69, 101, 114, 116, 139, 141, 146, 190, 191, 227
figure (rhetorical), 28
Fillmore, Charles, 4, 5, 6, 9, 75, 116, 177, 178, 194, 211, 221
Fleischman, Suzanne, 7, 58, 59, 62, 64, 153, 221
Focus Space, 32, 35, 48, 50, 51, 52–5, 58–9, 61, 64, 134
force dynamics, 2, 116
frame (semantic), framing, 4, 5, 18, 38, 52, 56–7, 85, 116–18, 121–2, 127, 128, 144, 157–8, 160, 162, 167–9, 170–1, 199, 214
frame of reference, 113–14, 121, 123–31, 132, 133–5, 199, 200
free indirect speech and thought, 198–9, 202–3, 205, 206, 208, 210, 211, 214–15
free indirect style, 6, 65, 177, 178–83, 187, 189–91, 192–4
future tense
 GONNA future, 9, 50, 59–60, 61, 82
 WILL future, 59, 60–2
generic structure, 117
gesture
 character viewpoint gesture, 14, 16, 100, 104–5, 106–9
 dual viewpoint gesture, 105
 gestural deixis, 13
 gestural viewpoint, 13–14, 98, 100–1, 102–6, 107–10, 134, 230
 observer viewpoint gesture, 100–1, 103–4, 107–9
grammaticalization, 50, 51–2, 59–60, 64, 65, 80, 90, 154
Ground, 1, 7–9, 37, 39, 42, 47–66, 70, 73, 90
 Ground Base Space, 48, 51
 implicit reference to Ground, 8, 48, 51–2, 62–4
handshape, 167
Hanks, William, 9, 12
honorifics, 5, 12, 213
Hopper, Paul, 47, 51, 58, 59
Horn, Laurence, 71–2, 82, 212
iconic, iconicity, 16, 114, 119, 131–2, 134, 152, 163, 168, 171
if, 5, 74, 80, 89, 212, 221–2
image schema, 123, 124, 125, 128, 130, 131, 211
implied author, 36–9
indexical, indexicality, 11, 13, 206
indirect speech and thought, 34, 104, 189, 198–9, 200, 201, 215
inferencing, 34, 55, 56–7, 58–9, 60–1, 62–3, 144
inferential transfer, 58, 64
internal speech and thought, 203, 214
intersubjectivity, intersubjective, 6, 8, 11, 12, 51, 52, 53, 70, 72–3, 75–7, 84–5, 86–7, 88–9, 90, 122, 212–14, 216, 224, 228, 230
irony, 4, 44, 72, 199, 210–11, 215, 220, 226–7, 229
 cosmic, 27, 35, 36
 dramatic, 27, 28, 36–7
 echoic mention theory of, 26, 29–31
 Gricean theories of, 30–1
 indirect negation theory of, 30–1
 literary accounts of, 26
 pretense theory of, 26, 30
 situational, 27, 28, 31, 36, 42–3
 stable, 36–9, 40
 structural, 31, 36–9
 unstable, 26, 40–3
verbal, 25, 27, 28–31, 33, 35–6
victims of, 25–6, 34, 36, 38

- Janzen, Terry, 13, 139, 146, 156, 160, 224, 226
 joint attention, 13, 32
 jokes, 25–6, 36, 40–2
- Kay, Paul, 177, 194
know, 52–4, 65, 77–8, 80, 83, 220, 222
- Langacker, Ronald, 7–8, 11, 27, 34, 47–8, 49, 50, 64–5, 73, 90, 101, 154, 185–6, 187, 212–13
- Lazard, Gilbert, 143–4
- Levinson, Stephen, 9, 13
- Levy, Elena, 106
- Liddell, Scott, 13, 14, 48, 51, 102, 114, 119, 127, 139, 141, 146, 160, 164
- linguistic viewpoint, 4–7, 12, 98, 99–101, 105–6, 109–10
- locus, loci, 10, 14–15, 99, 100, 103, 105, 109, 144, 156, 159, 162, 164, 165, 167, 169–70, 222, 227
- Lyons, John, 187
- material anchor, 13
- McNeill, David, 17, 100–1, 102, 103, 104–5, 106, 109, 223
- meaning change. *See semantic change*
- meaning construction, 26, 115, 119, 133
- mental model, 101, 119
- mental spaces
- mental space embedding, 198
 - mental space evocation, 34, 35, 211
 - mental space networks, 2, 9, 50, 52, 59–64, 70, 80, 121, 179, 223
- Mental Spaces Theory, 27, 31, 35, 48, 66, 114, 119, 219
- metalinguistic space, 49, 50
- metaphor, 1–2, 7, 97, 102–3, 134, 191
- mirror neurons, 2–3, 10
- modal verbs, 8, 50, 58–9, 73, 74, 100, 220
- multimodal language, 17, 97–110
- multiple viewpoints, 2, 27–8, 77, 90, 222–3
- narrative, 11, 16, 31, 38, 90, 100, 106, 141, 158, 161–2, 164, 165–9, 170, 171, 177–9, 184, 185–6, 187, 189, 192, 203, 206, 209, 220, 223, 224, 225–6, 228, 229
- first person, 182–3, 191
 - third person, 6, 182–3, 191, 225, 228
- narrator viewpoint, 36–9, 102–3, 225–6
- narrators, 6, 10–11, 13, 14, 28, 36, 37–9, 42–3, 48, 65, 71, 98, 102–3, 105, 107–9, 156, 161, 177, 181, 183, 185, 189, 190, 191–2, 193, 194, 198–9, 200, 202–4, 205, 207, 220, 222–3, 225, 226, 229
- negation, 4, 6, 69–90, 224
- and intersubjectivity, 70, 72–4, 76, 84–90, 224
- and positive polarity, 87
- and stance, 70–1, 74, 76, 77–9, 80–81–90, 227–8
- in grammatical constructions, 70, 71, 72–3, 74, 79, 82–90
- indirect, 30–1
- metalinguistic, 71–2, 76, 88, 212
- sentential vs morphological, 73–4
- Neg-Raising, 81–3
- Núñez, Rafael, 14
- O'Connor, Mary Catherine, 177, 194
- O'Dea, Barbara, 146
- or*, 71
- Pagliuca, William, 58, 59
- Parrill, Fey, 101, 102, 103, 105, 106, 114, 134
- partition, partitioning
- body partitioning, 139, 141–2, 144–7, 153–4, 162–4
 - space partitioning, 116, 121
- past tense, 6, 7, 48, 55, 64, 66, 74, 75, 80, 87, 177, 179, 184, 186, 202, 215, 221, 225, 229
- Perfect, 62–4, 65, 220
- Perkins, Revere, 58, 59
- perspective, 1, 13, 16, 28, 38–9, 97–8, 104, 118, 156, 157, 161, 190, 220–1, 222–3, 225, 229
- implicit perspective, 101–2, 189
 - internal perspective, 101, 187
 - perspective shift, 159–60, 162, 166–7, 169–70, 171–2, 177, 180–1, 183, 184, 188, 193, 194
 - viewer's perspective vs deictic center, 11
- poetry, poem, 1, 71
- pointing, 12–14, 152, 161, 162, 165, 167, 229
- deferred/displaced pointing, 13–14, 152–3
 - pronominal point, 156, 165, 166, 206
- Pratt, Mary Louise, 189
- present tense, 201, 208, 221
- presuppose, presupposition, 3, 5, 6, 52–4, 57, 65, 124, 135, 181, 187, 194, 227
- Presupposition Float, 55
- preterite, 64, 65
- Primary Metaphors, 7
- Primary Scenes, 2
- profiling (profiled, unprofiled), 10, 32, 47–8, 49, 63, 64, 65, 146, 154, 163–4, 186, 187, 222–3, 224, 227, 228, 229
- progressive aspect, 180, 183, 187–9, 194

- pronouns, 5, 7, 8, 9, 11, 14, 48, 55, 70, 75, 88, 99, 100, 106, 153, 166, 169, 178, 182, 189, 190, 199, 200, 201, 202–5, 229
proper names, 205
prototypes, 183, 186, 199, 201, 206, 207
quotative, 100, 101, 140
Real Space, 13–16, 17, 48, 51–2, 59, 102–5, 114, 119–21, 123–4, 127, 130–2, 141, 146, 160, 164, 165
reconstrual, 28, 32, 42
recursion, 43
Redeker, Gisela, 101–2, 104, 105, 189–90, 191, 198
referent tracking, 161, 162, 167–9, 170, 172
refutation, 84, 85–6, 88, 90
Relevance Theory, 82, 207, 211
reported speech, 139–54, 159, 207
represented perception, 99, 177, 179, 181, 185, 193
represented speech and thought, 4, 75–6, 179, 181, 193, 198–216, 219, 224. *See also* direct speech and thought, distancing
indirect speech and thought, free indirect speech and thought, indirect speech and thought, free indirect style
role shift, 14–16, 225, 229
Romanticism, 31
Sanders, José, 8, 18, 47, 48–50, 65, 101–2, 104, 105, 189–90, 191, 198
Sanders, Ted, 8, 18, 47, 48–50, 65
sarcasm, 26, 28, 30, 35, 199, 215
satire, 29, 38
semantic change, 8, 47, 56–64, 73, 74
Shaffer, Barbara, 139, 140, 146, 224
shared knowledge, 106–9
signed language, 13, 14, 100, 101, 139–74
simulation, 3, 17, 98–9, 101
 simulation-based semantics, 17
simultaneity, 165
Speech Act Space, 9, 49, 50–1, 58
speech acts, 7–8, 25, 33, 51, 52, 57, 64, 77, 152, 219
Sperber, Dan, 26, 29, 30, 34, 85, 207, 210, 211
stance, 6, 42, 53, 140, 141, 220, 227–8
 and mental spaces, 70, 80–1, 83–7
 and verb forms, 74, 75, 221
 epistemic, 5, 70–1, 75–83, 90, 220, 221–2
 in conditionals, 74, 80, 89, 221–2
 stacking, 87, 89
 verbs, 6, 75–83, 88, 220, 222
Stein, Dieter, 8
Story Space, 101–5, 223
subjectification, 8, 47–8, 50–66, 74, 75, 80, 89, 90, 154, 212, 213, 227–8
subjectivity, 8, 44, 47, 50–1, 64, 65, 73, 74, 75–6, 81, 84, 177, 183, 190, 212–13, 220
Sweetser, Eve, 5, 6, 8, 14, 16, 18, 34, 47, 48–50, 58, 59, 65, 69, 71, 74, 75, 79, 89, 102, 103, 114, 116, 134, 178, 190, 199, 211, 221
Taub, Sarah, 14, 134
tense, 4, 5, 8, 10, 47–8, 54, 59, 61, 64, 65, 80, 152, 185–6, 189, 200, 202, 208, 225, 226, 229
think, 4, 53–4, 75–8, 81–4, 89–90, 99, 222
Traugott, Elizabeth Closs, 8, 47, 50, 51, 58, 59, 64, 73, 75, 154, 189, 212–13
Turner, Mark, 13, 35, 66, 101, 114, 116, 139, 141, 146, 190, 191
van der Voort, Cok, 202
Vandelanotte, Lieven, 34–5, 75, 76, 85, 90, 189, 191, 198, 199, 200, 201, 202, 206, 207, 208, 211, 212
Verhagen, Arie, 6, 8, 11, 12, 53, 72–4, 75–7, 81, 88, 90, 213–15, 219, 224
viewpoint
 camera angle, 97
 cognitive, 113–14, 119–35, 171
 compression, 35, 38–9, 192–3, 228–30
 conceptual, 98–101, 102, 103, 109–10
 negotiation, 6, 8, 73–4, 219–20, 224, 225, 226–8, 229
 physical, 4, 5, 11, 14, 97, 113, 119, 125, 134, 159, 160–2, 229
 shift, 7, 14–16, 32, 39, 99, 101–2, 105, 109, 113–14, 119–22, 124, 129–31, 133–4, 156, 172, 177–94, 220, 222, 225, 226–7
 social, 11–12
 vantage point, 97
 viewing arrangement, 97
 visual, 2, 97, 113–14, 129–31, 133, 228
Viewpoint Space, 28, 31–2, 38, 42, 48, 54–5, 101–2, 103, 113, 119, 122, 127, 130, 134, 191–2, 194, 221, 223–4, 227, 229
was-now construction, 10–11, 55, 177–94, 225
Wierzbicka, Anna, 75
Wilson, Deirdre, 26, 29, 34, 85, 207, 210, 211
Wright, Susan, 8, 179, 185, 187
Wulf, Alyssa, 141
zooming out, 28, 31, 34, 35, 37, 40–2, 210–11, 215
zoom-out configuration, 40