In S. Strömqvist & L. Verhoeven (Eds.) (2004)
Relating events in narrative: Typological and contextual perspectives, pp. 457-476.
Mahwah, NJ: Lawrence Erlbaum Associates.

Variations on a Ranarian Theme

Herbert H. Clark

1. INTRODUCTION

Telling stories is hard. First you need something to tell – a genuine *story*. You must be acquainted with a happening, or series of events, about which you can make a point. Then you need an *occasion* to tell the story. Someone has just brought up a topic of conversation, and you have a story that is relevant (Jefferson 1978). Or you are participating in a round of storytelling, and it is now your turn (Kirshenblatt-Gimblett 1974). Or someone has asked for a particular story. Even before you start, you need an *overall plan* – where to begin, what events and evaluations to include or emphasize, and where to end. In the actual telling, you need to *formulate* utterances one by one. For each utterance, you must select the right words and gestures from the repertoire available to you – your language and culture – and you must do so in a timely fashion. Your selections must be both true to the events in the story and understandable to your audience. It is no wonder that telling stories is such an art – how many good storytellers do you know? – and that it takes children so long to acquire.

Many of these issues have been addressed in research on Mercer Mayer's (1969) story *Frog, where are you?* In these studies, people are asked to look through Mayer's wordless picture book (see Appendix I) and then, while paging through the book again, retell the story to someone else. The landmark collection of these studies is *Relating events in narrative: A cross-linguistic developmental study*, edited by Ruth Berman and Dan Slobin (1994). It set out the basic issues for the development of storytelling, and showed how children tell the frog story in many languages and cultures. The current collection is a follow-up with an even wider range of languages and cultures. Let me call the two volumes *Frog I* and *II*.

My aim is to take up three variations on this ranarian theme: story versus telling, diegesis versus mimesis, and conception versus description. If you listen

closely, you will also hear a leitmotiv that recurs throughout the collection itself – Slobin's (1996b) notion of thinking-for-speaking.

2. VARIATION 1: STORY VERSUS NARRATIVE

There is a traditional distinction in storytelling between the story someone tells and the telling of that story (see Berman & Slobin 1994). As Chatman (1978) put it:

> [E]ach narrative has two parts: a story (*histoire*), the content or chain of events (actions, happenings), plus what may be called the existents (characters, items of setting); and a discourse (*discours*), that is, the expression, the means by which the content is communicated. In simple terms, the story is the *what* in a narrative that is depicted, discourse the *how*.

With Chatman, I will distinguish between the same two parts, calling them the *story* and the *narrative* (or *telling*). Most work on the frog stories has focused on the narrative. And rightly so, for that is what changes most from language to language, and from childhood to adulthood. Always the contrarian, however, I want to take a closer look at the story.

2.1 Story types

One of the progenitors of the modern work on narratives was William Labov's work on *narratives of personal experience* (Labov 1972; Labov & Waletzky 1967). In one study, Philadelphia pre-adolescents, teenagers, and adults were asked such questions as (Labov 1972:354): "Were you ever in a situation where you were in serious danger of being killed, where you said to yourself 'This is it'?" or "Were you ever in a fight with a guy bigger than you?" When the respondents said "yes," they were asked "What happened?" and out came a narrative. The stories behind these narratives were experiences the narrators had lived, and telling the stories required them to reconstruct those experiences from memory. Indeed, Labov spoke of the narrators becoming "deeply involved in rehearsing or even reliving events of [their] past," and the narratives he cites suggest that the narrators got into them with heart and soul.

The frog stories are very different from the Philadelphia stories. Some of the differences make storytelling easier, but others make it harder.

i. *Factual versus fictional stories*. Labov's narrators were telling stories of fact, of events that actually happened. The frog narrators were telling stories that were pure fiction, events they knew hadn't actually happened. Now, fact and fiction place different requirements on storytellers. In

factual accounts you are constrained by what actually happened, and your story cannot include impossible events or non-existent people or settings. With fiction, you have license to make things up and to create people and settings in any way you like. A simple example is proper names. In the frog stories, the boy was given such invented names as Piita, Mikel, Adoni, Aitor, Torsten, Ali, Cookie, Pelle, Bert, Måns, Pepe, Pedro, Dani, and Tomasito, and the dog, Jensi, Zikin, Txuri, Zikina, Lasse, Bonkie, and Plutt. These names were created out of the language and culture of the narrator. The names used in the Philadelphia stories were presumably the characters' actual names, which the narrators tried to get right.

ii. *First- versus third-person stories.* The frog stories were told in the third person, where the narrator was distinct from the main protagonist. Here is an example from a nine-year-old narrating in English: "And he starts running. And he tips him off over a cliff into the water. And he lands."[1] The Philadelphia stories were told, instead, in the first person, where the narrator was also the main protagonist. Example: "An' then, three weeks ago I had a fight with this other dude outside. He got mad 'cause I wouldn't give him a cigarette."

First- and third-person perspectives require very different modes of thinking and linguistic resources (see, e.g., Segal et al. 1997). It may be easier to formulate certain utterances – say, about what protagonists are thinking, feeling, seeing, or trying to say – from the perspective of the protagonist than from the perspective of an observer, that is, in the first person rather than the third person. And different aspects of the story may get emphasized in the first person – excitement, fear, first impressions – than in the third person – the protagonist's appearance, location, and surroundings. Yet, other features of a story may be easier to formulate from the perspective of an observer, who can take in the broad sweep of a scene and not worry about the perspectives of different characters.

iii. *Creating a story versus retelling a story*. The Philadelphia narrators had to *create* both the stories and the narratives from scratch. They had to decide how to orient their audience to the who, what, where, why, and when of the story, which events to treat as the complicating action and which to treat as the resolution, what background commentary to include, and more. In Labov's analysis, the narratives they created emerged in six main sections: abstract, orientation, complicating action, evaluation, result or resolution, and coda.

The frog narrators, in contrast, had the stories handed to them in a book, and they had only the narratives to create. Their job was to *retell* the story as they understood it. Retelling a story has its own burdens. You are not free to make up this or that episode. Instead, you have to get the episodes of the story right, including the order of events and who is doing what. You must also be accurate on the point of the original narrator – why he or she was telling that story. This was a bit easier for the frog stories because the picture book gave the narrators leeway in the description of each event and the construal of Mayer's point. And the narrators were also reminded page by page, episode by episode, of how the story went. One result was that these narrators didn't create the sections that were characteristic of the Philadelphia stories. Creating sections was another burden lifted from their shoulders.

Plainly, then, storytelling is not all alike, a point that has been made by many others (e.g., Berman, this volume; Hickmann, this volume). It just isn't the same to tell a factual versus a fictional story, to tell a story from a first- versus a third-person perspective, or to create versus retell a story. Surprisingly little is known about these differences. And yet, if we are interested in the skill of storytelling, or the age at which it is mastered, these are issues we must understand.

2.2 The same story?

The research on the frog stories has relied on a tried-and-true experimental method: Have everyone retell the *same* story – as represented in Mayer's *Frog, where are you?* –, then attribute any *differences* in the narratives to the language, culture, age, or skill of the narrator. But is everyone really retelling the *same* story? The contributions to *Frog II* suggest that the answer is "no." First, the youngest children didn't seem to understand the story fully, or had alternative interpretations of the same pictures. And second, what the story was taken to be changed from culture to culture.

Where does the frog story come from? The answer, of course, is Mayer's *Frog, where are you?* But does everyone understand the story as Mayer intended? Clearly not. Many of the youngest narrators described what was depicted in each picture, and didn't seem to connect successive pictures. They might have recognized the boy, dog, and frog as the same in successive pictures, but they often failed to see a causal or temporal relation between the pictures. This point was made in a number of the contributions to *Frog I* and *II*. If this is true, what are we to make of the problems that children have in narrating the frog story? Is it that the children are unable to *interpret* the picture-book as intended, or is it that they are immature in their ability to *describe* cause, temporal relations, motion events, and other features of the basic story?

Indeed, Mayer's picture book isn't all that easy to interpret. In Picture 1, we meet a frog I will call Snag, and it leaves the house in Picture 2. In the penultimate

picture, the boy discovers a family of frogs – a father frog, a mother frog, and baby frogs –, and in the final picture, he leaves with one of the baby frogs, giving us the resolution of the story. But what has happened to Snag? I assumed that Snag was the father frog, and the boy was leaving with one of his babies. But as a friend pointed out, how could Snag have found a mate and sired a family overnight? Why didn't I assume that Snag was the mother who had gone back home just in time to deliver her babies? A number of the tellings suggest that I wasn't the only reader with problems.

It is no surprise that children and adults read *Frog, where are you?* against the common ground of their own culture. As an American, Mercer Mayer probably intended his depictions to be interpreted against American presuppositions about boys, dogs, and frogs, but narrators from other cultures clearly read it differently. This point is illustrated by how people identified the several features of the story – the gopher, the owl, the bees, the deer, and the body of water:

- the gopher was sometimes seen as a groundhog (Tzeltal), a mole, chipmunk, or squirrel (Japanese), or a squirrel or mouse (Spanish);
- the owl was sometimes a butterfly (West-Greenlandic) or bat (Japanese);
- the bees were sometimes flies (Icelandic), mosquitoes, or wasps (Spanish);
- the deer was sometimes a reindeer (West-Greenlandic) or elk (English), a gazelle or goat (Turkish), and sometimes even a horse (Spanish – do horses ever have antlers?);
- the body of water was a puddle or lake (Tzeltal), a river, lake, or sea (West-Greenlandic), a sea (Icelandic), a river or pond (Basque), or a stream or little puddle (American Sign Language);
- the frog was considered a pet in most cultures, but prey in Arrernte;
- the boy and dog were always a boy and a dog, as Mayer presumably intended.

These categories, surely, reflect cultural and not linguistic differences. The West-Greenlandic or Icelandic choices of "sea" for the body of water are dictated not by the syntax of Greenlandic or Icelandic, but by the narrators' familiarity with bodies of water. And the Turkish children may be more familiar with gazelles and goats than with deer – at least compared with American children. Ultimately,

familiarity is a cultural, not a linguistic issue. And with these simple categorical differences, we have barely scratched the surface of the cultural influence on the interpretation of Mayer's story.

One more point. It seems only common sense that narrators start with a story and end with the narrative. After all, there can be no telling of a story without a story to tell. But this is much too simple a conception of storytelling, a point demonstrated again and again in *Frog I* and *II*. Narrators don't construct a narrative *simply* to fit a story, selecting their words, phrases, and rhetorical devices to express the elements of a pre-determined conceptualization of events. They also pick and choose among the *potential* elements of a story to find ones they are capable of expressing. That, of course, is one of the ideas behind thinking-for-speaking. You see a guy leaving a room in a hurry. In English, you might speak about him *running* out of a room (*The man is running out of the room*), whereas in Spanish, you would speak about him *going* out of a room (*El hombre salió de la sala*) – with no mention of speed (see, e.g., Slobin, this volume). Viewed this way, English and Spanish encourage you – perhaps even require you – to tell different *stories* of what you saw. Each of these choices may seem inconsequential, but in the aggregate they can lead to very different stories.

In brief, if we are to understand the process of narrating, it is essential to understand the story being narrated – how it was created, where it came from, what its features are. It makes a difference whether the story is fact or fiction, third-person or first-person, created from scratch or retold. And it makes a difference what common ground – largely cultural – the narrator presupposes in narrating the story.

3. VARIATION 2: MIMESIS VERSUS DIEGESIS

Another distinction in storytelling, dating back to Plato, is between *diegesis* and *mimesis*. With diegesis, narrators *describe* things, and with mimesis, they *show* things, as in direct quotation. Both devices are useful in narratives. As David Lodge (1990:144) put it, "Roughly speaking, mimesis gives us the sense of reality in fiction, the illusion of access to the reality of personal experience, and diegesis conveys the contextualising information and framework of values which provide thematic unity and coherence." And both devices are *common* in narrative. As Lodge went on, "[The] alternation of authorial description and characters' verbal interaction remains the woof and warp of literary narration to this day". The frog stories offer a striking example of Lodge's observations. Mercer's picture book has only one character who can talk, so the frog stories should contain no conversation at all. And yet they are filled with mimesis. But how and why?

3.1 What is mimesis?

Showing, in its everyday sense, takes two main forms. The first is *indicating*. I can show you my car by pointing at it – by *indicating it as my car*. "There it is," I say, using the direction of my finger as an index to the car. The second form of showing is *demonstrating*. I can show you how my brother walks by doing a limp across the floor. "Here's how my brother walks with his sprained ankle," I say, while limping. By doing the limp, I *demonstrate something for you*, namely how my brother moves. What Plato and Lodge mean by mimesis, pretty clearly, is demonstrating and not indicating.

Indicating and demonstrating deploy what the philosopher Charles Sanders Peirce called indexes and icons. Indicating is a method of signaling by which a person means something by creating an index for an addressee, and demonstrating is a method by which a person means something by creating an icon for an addressee (see Clark 1996, Chapter 6). For Peirce, indexes and icons contrast with symbols (e.g., words and emblematic gestures), which are used in pure descriptions. That is, indicating and demonstrating contrast with *describing-as*. As Peirce himself noted, most signals rely on a combination of these methods. It is rare to find a pure indication, demonstration, or description.

How people demonstrate is often treated as self-evident, but it is hardly that. I will illustrate with the analysis of quotations as a species of demonstration (Clark & Gerrig 1990).

Demonstrations, Gerrig and I argued, are *selective depictions* that are intended to "enable others to experience [in part] what it is like to perceive the things depicted" (Clark & Gerrig 1990:765). Take this excerpt from a frog story by a Turkish adult I will call Eda, as translated into English (Aksu-Koç & Tekdemir, this volume):

> Ali started climbing on the rock to escape the owl. But what did he see? Actually he didn't see anything. He thought what he saw was a tree and holding onto a branch he started yelling, "Little frog! Little frog!"

Suppose that in producing "Little frog! Little frog!" Eda cupped her hands around her mouth and spoke slowly and loudly ("Li-i-tt-le fro-o-g! Li-i-tt-le fro-o-g!"). With this entire performance, she intended to create a selective depiction of what the boy, Ali, did. She tried to depict not only his words (*little frog, little frog*), but also the loudness and intonation of his voice, and the motion of his hands. And yet she was *not* trying to depict, for example, his gender or the direction in which he yelled.

Demonstrations therefore partition into four separable aspects:

i. *Depictive aspects*. Only some aspects of a demonstration are intended to depict features of the referent. Eda's demonstration was intended to depict the words, intonation, loudness, and hand formation of Ali's original. These constitute the *demonstration proper*.

ii. *Supportive aspects*. Other aspects of a demonstration are not themselves depictive, but are necessary as support in the performance of the depictive aspects. The gender and direction of Eda's voice were needed for her demonstration, but she didn't intend them to be depictive.

iii. *Annotative aspects*. Still other aspects of a demonstration are added as commentary on what is being depicted. These might include concurrent laughter or sneers, or exaggerated intonation.

iv. *Incidental aspects*. The remaining aspects of a demonstration are incidental, ones the speaker has no specific intentions about. These might include the way Eda sat, blinked her eyes, or pursed her lips.

Eda intends only part of her full demonstration to be depictive. Her addressees would have misunderstood her if they had construed her supportive or incidental aspects as depictive too. In general, people intend different aspects of their demonstrations to be depictive, supportive, and annotative. Gerrig and I called this the *decoupling principle*. The issue is how speaker and audience coordinate on this decoupling.

Producing and understanding demonstrations relies on imagination and pretense. Eda must begin by imagining Ali calling the frog and ask, "How can I get my audience to imagine his actions?" One way is by describing (diegesis), "… he started yelling for his frog," but that leaves all the work of imagining to the audience. Another way is to create a selective depiction that enables the audience to experience, in part, what it is like to see, hear, and feel Ali calling the frog. That experience should make it easier to imagine the right scene (see Clark & van der Wege 2001).

Another element needed for interpreting demonstrations is joint pretense. Eda's audience must join with her in the belief that, in producing "Li-i-tt-le fro-o-g! Li-i-tt-le fro-o-g!" with cupped hands, she *herself* isn't calling a frog. She is *pretending* to be the boy calling the frog, a type of play-acting, and she expects her audience to recognize the pretense. Like all fiction, quotations are to be understood against a *joint* pretense by speaker and addressee (Clark 1996, Chapter 12). Eda's thinking-for-speaking, then, includes imagining the right scene and determining which aspects to depict and how. It also includes establishing the right joint pretense.

3.2 Mimesis in the frog stories

The frog narratives contain a remarkable variety of mimesis, or demonstrations. It is even more remarkable how often the different varieties were used in combination.

Pictures

Mayer's picture book is itself entirely mimesis. It is a series of drawn depictions, and it takes skill to read these as a story. Indeed, the very title of Mayer's book is a bit of mimesis, the quotation "Frog, where are you?"

Understanding depictions isn't easy. To interpret Vincent van Gogh's painting *The potato-eaters*, according to Walton (1990), we must use the oil shapes on the canvas to help us imagine a fictional world in which there are two men and three women, poor Dutch farmers, sitting around a table eating potatoes and drinking tea. We must imagine ourselves in that world because we are viewing this tableau from within the room over the left shoulder of one of the women. We must also attribute the appropriate motivations to these farmers. We might assume, for example, that they are eating and drinking because they are hungry and thirsty. But are we to infer that they are starving, or that they have nothing to eat but potatoes? In interpreting this painting, then, *a*) we see the shapes on canvas as depictions of things; *b*) we imagine a coherent world in which those things reside; *c*) we understand that that world is fictional; and *d*) we infer states and attributes that are not directly depicted.

Interpreting *Frog, where are you?* requires even more. When we view Picture 1, we imagine a fictional world with a boy, a dog, a frog in a jar, and a bed, all in a bedroom at night. But when we view Picture 2, we don't create a new fictional world: we simply add a next scene to the fictional world of Picture 1. We see the boy, dog, frog, jar, bed, and bedroom of Picture 2 as identical to those of Picture 1, but at a later time. For other pictures, we are supposed to add attributes and motivations – curiosity, fear, happiness – that are not directly depicted in the drawings. So in interpreting the picture book, *e*) we see the successive pictures as successive states of the *same* world; and *f*) we see them as representing a story – a complicating action followed by its resolution. Mayer's pictures enable us, in short, to "experience [in part] what it is like to perceive the things depicted" – the story of the boy and his frog.

Narrators of the frog stories often used the pictures as props for their storytelling as well. As many authors in this volume noted, narrators often pointed at the pictures. They *indicated* things for their audience, showing them in the indicative sense of *show*. These acts piggy-back on the pictorial demonstrations. Suppose Eda points at the owl as she says, "Ali started climbing on the rock to escape the owl." Eda isn't pointing at the referent of *the owl*, which is fictional and nowhere to be found. She is pointing at a *depiction* of the owl, which is to be interpreted against the joint pretense that the thing depicted is part of an actual world.

Quotation

Quotations appear in one frog narrative after another in the examples cited in this collection – in Tzeltal, Icelandic, Swedish, West-Greenlandic, Spanish, Turkish, Japanese, Basque, and American Sign Language. What is more, many of these quotations are particularly sophisticated, as these examples illustrate (in English translation):

i. *Saying versus thinking*. A Tzeltal-speaking adult (Brown) used quotations not only for speech, but for thoughts: "Maybe he says to himself …" and "they are thinking …" An Icelandic nine-year-old (Ragnarsdóttir & Strömqvist) did the same: " 'Where has it gone?' she thinks."

ii. *Manner of speaking*. Swedish and English narrators (Strömqvist, Nordqvist & Wengelin) embedded direct quotations in a great variety of verbs, including *ask*, *mutter*, *exclaim*, *whisper*, and *call*. A West-Greenlandic narrator said: "Then he shouts 'Mummy.' " Swedish narrators also used a variety of adverbs with their verbs of speaking (e.g., "asked confusedly").

iii. *Animal quotations*. An American Sign Language narrator (Galvan & Taub) quoted the dog speaking (or rather signing): "The concerned dog barks up at the boy and says, 'Get down, get down.' " So did a Tzeltal narrator (Brown), " 'There's something here. There's something squatting here,' he says." A seven-year-old Spanish narrator, in contrast, quoted the dog barking ("The dog was au au au au au au!") as well as the bees buzzing ("and the bees were bbbbzzzzzzzz"). So the narrators sometimes depicted the dog's message and other times the dog's sounds. These reflect two distinct reasons for demonstrating.

iv. *Sound quotations*. A West-Greenlandic narrator said "Shh maybe they/ some are there," quoting not only the boy's words but his saying of "shh."

v. *Prosody*. Although one set of Swedish narrators were congenitally deaf (Strömqvist, Nordqvist & Wengelin), they used quotations of speech in their typewritten stories, complete with orthographic representations of prosodic effects. They typed "grodaaaaaa" to represent a prolonged "froooooog," and "H-A-L-L-Å" and "H-J-Ä-L-P" to represent the shouting of *hello* and *help*.

All of these features are common for spoken and written quotations (Clark & Gerrig 1990). There are probably many more such features in a larger corpus of the languages represented here.

Mimetics

Most languages have linguistic techniques for representing the sounds and other features of animals, things, and events. In English, for example, James Joyce used *pitapat* in *Ulysses*, "And while she gazed her heart went pitapat," and a newspaper columnist used *boom* in a story, "So we have these Trident 2 missiles fall down and go boom." Words like these are variously called ideophones, onomatopoeia, mimetics, and sound symbolism, and they take different forms in different languages. Many of these words are conventional, such as *pitapat*, *boom*, *cockadoodle doo*, *clunk*, and *ticktock* in English. But narrators also make up their own words, as in the following English examples (Clark & Gerrig 1990):

> The pounding of the cylinders increased: ta-pocket-pocketa-pocketa-pocketa-pocketa. (James Thurber, *The secret life of Walter Mitty*)

> The room reeked of camphor. "Ugf, ahfg," choked Briggs … (James Thurber, *The night the bed fell*)

> Everywhere you went you could hear the awful *brrrpppbrrrpppbrrrppp* of their saws. (Charles McGrath, *Id*)

Among the languages represented in this volume, it is Japanese and Basque that make the most of these expressions. Mimetics are common in Japanese, as in these English translations "The bees flew *buzzing*," "A bat flew out, so [the boy] fell down *thud*," and "The frog sneaked out *hop*" (Küntay & Nakamura). In the Basque stories, there were mimetics for *quickly* and *suddenly* in "And *'quickly'* [they] jump onto the log," and "And *'suddenly'* a mouse [comes] out" (Ibarretxe-Antuñano). These mimetics work something like *splat* in "The tomato hit him *splat* in the face," in which *splat* describes the hit as done "with a smacking or splashing noise."

Mimetics are also a type of demonstration (Clark & Gerrig 1990:788). When Joyce writes "her heart went pitapat," he doesn't mean that her heart produced literally these sounds. He intends only selected aspects of *pitapat* to be depictive – namely the rhythm. He could just as well have said *dum-de-dum*, *ratatat*, or even *Peter Pan*, and if he were speaking, he could have said it quickly or slowly to depict the heart's speed. The other features of *pitapat* are supportive aspects. In English, *pitapat* happens to be a conventional word for demonstrating a heartbeat, just as there are conventional words for demonstrating other sounds. So, like any demonstration, mimetics are selective depictions of their referents.

Iconic gestures

Narrators often use iconic gestures, and although gestures weren't recorded in most languages of this volume, they were in Basque (Ibarretxe-Antuñano) and American Sign Language (Galvan & Taub). Consider an example from a Basque adult: "And, 'suddenly,' a bird like this, em, an owl came [flew] out to him." Over the word *flew*, the narrator spread out his arms and flapped them like wings. We don't take the narrator to be trying to fly. We take him to be trying to depict an owl flying, and to get his audience to use that selective depiction to imagine the owl flying. The Basque narrators used many such gestures.

According to Galvan and Taub, signers use not only conventional signs, but iconic gestures or modifications that depict selective aspects of the referent. As they note, "Even lexical verbs will sometimes be spatially modulated to show specific paths and movements." And in a phenomenon called *role shift* (or *referential shift*) in sign language, "the signer's own face and body are taken to represent the face and body of some entity being described; the signer takes on that entity's 'persona' and demonstrates the entity's actions or experiences." These are demonstrations, and instances of joint pretense, for depicting entities. As Galvan and Taub noted, citing work by Engberg-Pedersen, signers have a narrative ideal: "In a signed story, wherever possible, narrators should 'show' the information using iconic forms, rather than 'describe' it using lexical forms." It is as if these narrators had read David Lodge's advice on how to write good fiction.

Mimetics are often accompanied by iconic gestures – as if the two were of a piece. The Japanese speakers recorded by Kita (1997) all gestured whenever they used mimetics. Much the same is found in Basque (Ibarretxe-Antuñano, this volume). When one Basque adult says, "And there goes the dog, very fast, running running running," he gestures a running motion in rhythm over the repeated word *running*. The repeated word is itself a type of sound demonstration (see Clark & Gerrig 1990:789), and the narrator fills it out with a second type of demonstration – an iconic gesture. Kita's and Ibarretxe-Antuñano's evidence suggests that the type of thinking needed for speaking here is a perceptual, imaginal representation. Without that, it would be impossible to formulate these composite auditory and visual demonstrations.

3.3 Imagining-for-speaking

Demonstrations add a notable wrinkle to Slobin's thinking-for-speaking. First, demonstrations are ubiquitous in narratives. They take the form of illustrations, quotations (both direct and free indirect quotations), mimetics of various types, iconic gestures, and forms I haven't even mentioned. To narrate, as Lodge noted, is both to tell and to show. Second, demonstrations are a central component of effective narratives. In an unpublished study of my own, I had 30 students listen

to a brief story and then retell it to a partner. Later, I asked other students to rate the quality of these 30 retellings. Rated quality rose with the amount of quotation used: the percentage of words that were in quotations accounted for about 25% of the variance in the judged quality of the narratives. And, third, child narrators use demonstrations – quotations, iconic gestures, mimetics – from the very start. (In a diary study by Eve Clark (unpublished), one child used his first quotation at age 2.) Oddly enough, the use of quotations and mimetics in narratives decreases over age in Japanese and Turkish (Küntay & Nakamura). These cultures appear to place a value on diegesis as the proper form of narration, as do some English-speaking cultures. It is as if too much mimesis is bad. Still, children are fluent in its use.

Demonstrations, or mimesis, require a radically different form of thinking-for-speaking, or so I have argued (Clark 1996, Chapter 6). To formulate a garden-variety quotation, you must imagine a person speaking and gesturing – the person's *appearance*, or what he or she looks and sounds like – and then select which aspects of that scene you are going to depict. You must select aspects not merely that you *wish* to depict, but that you are *able* to depict with devices you have at hand for demonstrating – tone of voice, intonation, concurrent gestures, the ability to lisp, slur words, or speak in dialect. Only then can you formulate the quotation itself. The selection process is no different, really, from the selection process behind describing motion events. There, too, you must select those features – path, manner, tense – not merely that you wish to express, but that your language – Tzeltal, Basque, Spanish – allows you to express.

4. VARIATION 3: CONCEPTION VERSUS DESCRIPTION

Motion events can be described in many ways. In English, I can describe one and the same *physical* event in these two ways:

[1] Ken ran out of the room.

[2] Ken left the room at a run.

Intuitively, these differ in how I conceptualize the event. In [1], I think about Ken trotting, sprinting, moving fast, and I think about this motion *in relation to* its direction – out of the room. In [2], I think about Ken in relation to the room – that he is no longer in it – and I think about *that* change of state in relation to the manner in which it happened – at a run. According to these intuitions, in formulating the descriptions in [1] and [2], I conceptualize the event in two distinct ways. What I have just described is, of course, an example of Slobin's thinking-for-speaking.

Here I take up four issues on this topic: *a*) no matter what the language, speakers must analyze scenes in order to describe them; *b*) formulating descriptions is a matter of distribution and detail from these analyses; *c*) analyzing scenes depends on community expertise; and *d*) analyzing scenes is incompatible with certain views of embodied cognition.

4.1 Scene analysis

Every description requires an *analysis* of the scene being described. Suppose I am looking at five goats. In order to describe them in a single phrase, I must analyze the scene into certain features, dimensions, aspects, parts, relations. For English, I might identify three elements: *a*) the individual animals; *b*) the classification of each individual as a goat; and *c*) the number of those individuals. The result: *five goats*. For Mandarin, I would analyze the scene into four elements: *a*) goatiness; *b*) the individuals; *c*) the classification of each individual as an animal; and *d*) number. The result (in its English equivalent): *five head of goats*. If I know the goats by name, I might analyze the scene into the five individuals: *Billy, Gruff, Beardo, Wilma, and Speedy*. It would be strange, though possible, to analyze the scene as follows: *a mammal, a thing, two goats, and Billy*. People can *conceptualize* the same scene in perhaps an unlimited number of ways.

What are the components, parts, dimensions, aspects, or relations of such an analysis – of such a process of abstraction? One of the only ways to answer the question is to look at what is required for descriptions. Here are a few examples for spatial descriptions (see, e.g., Talmy 2000a, 2000b):

Dimensions
We cannot describe a door as tall, wide, or thick, or a soup as thick or thin, or a town as to the north, without analyzing the door, soup, and local area into physical dimensions such as height, width, thickness, and north-southness (Bierwisch 1967).

Directions
Nor can we describe a person as tall versus short, or a lake as deep versus shallow, without knowing the positive and negative directions on these dimensions. *Tall* and *deep* are measures in the positive direction, and *short* and *shallow* in the negative direction, presumably for perceptual reasons (Clark 1973).

Trajector and landmark
To describe a cup as on a table, we must conceptualize the cup in relation to the table, and not vice versa. We must see the cup as trajector and the table as landmark – as thing related versus relatum (or as Talmy's Figure versus Ground).

Some relations seem easier to see as trajector and landmark than vice versa. These include:

i. small objects in relation to large objects, as in *the cup is on the table*;
ii. moving objects in relation to static ones, as in *the man ran by that chair*;
iii. objects in the focus of attention in relation to objects not in the focus of attention.

All spatial prepositions presuppose trajector and landmark, and so do many spatial verbs, adjectives, and even nouns.

Spatial relations
In English, we cannot select among *at* or *on* or *in* without analyzing a scene into such spatial relations as abutment, support, and containment, notions that themselves may need to be analyzed further.

With-respect-to relations
The relation of trajector to landmark is merely the prototype of a more general relation I will call *with-respect-to*, or *wrt*. We ordinarily see:

i. parts with respect to wholes, as in *the handle of the cup*;
ii. dimensions with respect to wholes, as in *width of the door*;
iii. directions with respect to origins or goals, as in *north of London*, *northerly*;
iv. attributes with respect to what they are attributes of, as in *the flavor of your coffee*, or *the color of that cloth*;
v. actions with respect to actors, as in *the singing of the contralto*.

Chains of relations.
With-respect-to relations can be combined into chains. In describing a room, for example, I see a desk chair next to a desk, and the desk in the corner. I don't see the desk chair in the corner as such (see Shannon 1984). And I see the brown color of the chair as belonging to the chair, which is next to the desk in the corner of the room. The brown doesn't belong to the desk or to the corner of the room. All of this is common sense, but only because we are so used to such analyses.

It is often assumed that features of the scene itself, or our perception of those features, dictate how we conceptualize it, but that isn't entirely so. We are readily able, for example, to see a table with respect to a cup – as supporting, or under, the cup – although that may take more time (Clark & Chase 1974). We can also see objects with respect to their parts, dimensions, directions, attributes, and actions,

even if that, too, may take more time. The with-respect-to relation appears to be universal – although this is always a dangerous claim. To describe the relation between two things, we are usually forced to conceptualize one of them with respect to the other. Symmetrical relations are rare.

4.2 Distribution and detail

Speakers must analyze motion events in order to describe them. According to Talmy (2000a, 2000b), the main elements of such an analysis are Motion, Path, and Manner, and they get combined in characteristic ways. Verb-framed languages prefer expressions like *Ken left the room running*, in which Motion and Path are combined in the verb (*left*), and Manner is elsewhere (*running*). Satellite-framed languages prefer expressions like *Ken ran out of the room*, in which Motion and Manner are combined in the verb (*run*) and Path is elsewhere (*out of the room*). Other languages keep all three components separate, as in *Ken went running out of the room* (see Slobin, this volume). These patterns may only be preferences. English, for example, allows all three major patterns.

But motion descriptions often consist of much more than Motion, Path, and Manner, a point demonstrated in many of the contributions to *Frog I* and *II*. Motion descriptions are always to be understood against presuppositions about the cultural situation – against people's common ground. And although some elements of these events are expressed in the description, other pieces are to be inferred. That is, languages differ in their distribution of information among: *a*) basic elements such as verb, object, and satellites to express Motion, Path, and Manner; *b*) other content of these expressions; *c*) auxiliary expressions; and *d*) inferences from common ground.

Let me illustrate these points with the expression *on the sofa*. If I were to ask a friend what it means to be "on the sofa," she might reply, "Is this a trick question?" The answer seems obvious. But consider: *the old woman is on the sofa*; *the couple was making love on the sofa*; *I looked for the manufacturer's tag on the sofa*; *she put antimacassars on the sofa*; *there are oak legs on the sofa*; *the sun is on the sofa*; *we put new upholstery on the sofa* (cf. Searle 1978). We cannot locate the woman, the couple, the tag, or the upholstery simply by understanding the "locative" *on the sofa*. Information about their location is distributed among several elements: *a*) the meanings of *on*; *b*) presuppositions we make in our culture about sofas; *c*) presuppositions we make in our culture about women, couples, tags, and upholstery (e.g., where manufacturer's tags are found and where antimacassars go); and *d*) the *on*-relations afforded by sofas. And when we need to, when *on the sofa* wouldn't be enough, we add auxiliary information, as in *there is a smudge on the left arm of the sofa*.

Talmy's schema, then, represents only part of the information we normally use in describing motion events. For example, individual languages have devices

for enriching these descriptions, as illustrated again and again in this collection. Here are three such devices.

Tzeltal would be classified as a verb-framed language, and indeed, narrators use a lot of verbs such as *enter*, *exit*, and *arrive*. But that doesn't do justice to their descriptions of motion events. As Brown documents, Tzeltal has a large inventory of *positional* verb roots, such as *lie–face-upward–spread-eagled*, *low-crouch*, and *insert-tightly*, and narrators use them to imply manner. When a narrator says "He [dog] looks like he's low-crouching walking [= he's limping]," he describes the position of the dog in a low-crouch, which implies a manner of walking. Tzeltal-speakers exploit their large inventory of positional verbs for work that narrators of other verb-framed languages (such as Spanish) might neglect or that narrators of satellite-framed languages like English would do with motion verbs (such as *limp*).

Japanese is also a verb-framed language, but that doesn't explain the common use of mimetics in describing motion events. Recall the four- and five-year-olds who used examples such as "The bees flew *buzzing*" and "The frog sneaked out *with a hop*," where the italicized words are mimetics analogous to English *thud* or *splat*. Mimetics were also common in Basque. In contrast, Turkish-speakers availed themselves of all of the same resources as Japanese-speakers – except for mimetics, which they almost never used (see Küntay & Nakamura).

American Sign Language, as I noted earlier, makes liberal use of another type of demonstration called *role shift* or *referential shift*. In this device, "the signer's own face and body are taken to represent the face and body of some entity being described; the signer takes on that entity's 'persona' and demonstrates the entity's actions or experiences." Signers used these shifts to enrich their motion event descriptions beyond the bare bones of Talmy's Motion–Path–Manner schema.

4.3 Community expertise

One of the oldest, and best documented, observations about language is that vocabulary size reflects shared community expertise (see Clark 1998). Lawyers have a large vocabulary of legal terms, doctors a large vocabulary of medical terms, and car enthusiasts a large vocabulary of cars and car parts. If a culture deals in chili peppers or rice, it will develop specialized terms for types and subtypes of chili peppers and rice, where other cultures will not (Berlin 1972; Burling 1970). Note that legal, medical, and car terms aren't part of the vocabulary of the general English-speaker, even though they are English words. It is the shared expertise of the community that counts.

It takes a special community expertise to learn and maintain some of the enrichment devices just mentioned. The large number of positionals in Tzeltal, for example, suggests that Tzeltal-speakers are experts on the shapes and orientations of natural objects and artifacts in daily life. Indeed, as Berlin (1968) documented,

Tzeltal also has over 500 nominal classifiers. These are words comparable to English *sheet* and *head* in *two sheets of paper* and *five head of goats*. They classify natural objects and artifacts mainly on the basis of animacy, shape, and orientation – much as the positionals do.

Other enrichment devices, in contrast, seem unrelated to community expertise. Japanese maintains a large number of mimetics, but are the Japanese in general any more expert on the referents of these mimetics than, say, the Turks? According to Slobin (this volume), English maintains a larger vocabulary of manner verbs (e.g., *limp*, *amble*) than Spanish, but is this a reflection of differences in cultural expertise? The obvious alternative is that the Japanese language maintains mimetics, and the English language manner verbs, precisely for enriching descriptions of motion events in Japanese and English.

The conclusion might be this: In descriptions of motion events, narrators enrich the skeletal Motion–Path–Manner schema with devices that are unique to their language. Some of these reflect an expertise that is special to that community, but others reflect a communicative expertise that is maintained simply for its utility in communication.

4.4 Embodied cognition

There is much talk nowadays of *embodied cognition*, the idea that people think about entities as physical objects – things that are seen, felt, and experienced holistically. Theories of embodied cognition come in two main varieties. One variety assumes that people think about dogs, owls, bees, boys, running, falling, and flying as wholes and do not analyze them into parts, features, components, aspects, or relations (e.g., Glenberg 1997). The second variety, which I myself have championed (see Clark 1973; Clark & Chase 1972), assumes that even if people can and do think about objects and events as wholes for some purposes, they must also analyze them into aspects, features, dimensions, parts, and relations for other purposes (see, e.g., Barsalou 2003). Let me call these two views the *holistic* and *analytic* versions of embodied cognition.

If we have learned one thing from *Frog I* and *II*, it is that the holistic version of embodied cognition is incomplete. For people to describe a motion event, they must, at some point, analyze that event into parts, features, components, aspects, relations – at least the elements of Motion, Path, and Manner. They need other elements if they are ever to formulate descriptions that include dimensions (e.g., *the large deer*), directions (*off over the cliff*), trajectors and landmarks (*owl in the hole*), as well as other parts, features, and attributes of objects (*startled dog*). We may well be able to view a scene holistically. But in preparation for speaking, we cannot stop there. We must abstract out all those elements that are needed in formulating effective descriptions in our particular language.

One argument often offered for the holistic version is that iconic gestures are based on holistic representations of objects and actions. But imagine June moving her right hand in small circles along a downward trajectory as a way of describing a ball rolling down a hill. Are June's gestures based on a holistic representation of the rolling ball? Not at all. Her gestures are a *selective* depiction – a depiction of only selected elements of the rolling-ball scene – and it takes an analysis of the scene to select those elements. Iconic gestures require just as much abstraction of elements as descriptions do. Recall the Basque narrator who demonstrated the flying owl. He depicted only *selected* aspects of the owl – that it was an animated being with projections at its side that went up and down. The Basque audience couldn't have understood his gesture without decoupling those aspects from the rest of his performance and using them to help imagine the scene the narrator was trying to depict. The same goes for quotations, mimetics, and drawings. Mimesis, like diegesis, is based on elements abstracted from the scene being described.

5. SUMMING UP

At a party, I told a friend that I was reviewing two volumes of research on kids telling stories about frogs. Two volumes? About frogs? And this is research? I realized too late how absurd it all must have sounded. But, in fact, the project has been a remarkable success. It is a model of how to carry out comparative research on language use. The recipe is simple. Find a rich, but circumscribed, event that people can describe in any language. Record people describing that event in many languages and at many ages. Then compare. Some years ago, Wallace Chafe followed this recipe with a dialogue-free film about pear pickers, and the results have been influential (see Chafe 1980). Mayer's picture book is not only more portable than Chafe's film, but also more accessible to children. Already, it has elicited a remarkable line of research, and its future looks even rosier.

I have considered three variations on our ranarian theme. What is the *story* in contrast to the *telling* of the story? What is *mimesis*, or demonstration, in contrast to *diegesis*, or description, and how is this distinction played out in the development of storytelling? And how do we conceptualize a motion event on our way to describing it? These questions have brought home – for me at least – an important point: Thinking-for-speaking is more than preparing for syntax, morphology, and lexicon. It is preparing for a factual versus fictional story, for a first- versus third-person story, and for a story told from scratch versus a retelling. It is also preparing to demonstrate versus describe – and I haven't even mentioned how we might prepare to indicate things. As it happens, demonstrating, describing, and indicating are normally found in combination – in composite signals – so preparing to describe, demonstrate, and indicate will not be simple.

For any piece of a story, narrators must select which of the three methods to use and how to combine them.

This research reflects an odd paradox about speaking and thinking. We seem to know much more about speaking than we do about thinking. The problem is that much of what we *do* know about thinking has come, explicitly or implicitly, from what we have discovered about thinking-for-speaking. How will we ever break this cycle?

NOTES

1. Narrators could have told the frog story in the first person – say, from the boy's point of view – but as far as I can tell, none did.