

Psychological Review

Depicting as a Method of Communication

Herbert H. Clark

Online First Publication, February 8, 2016. <http://dx.doi.org/10.1037/rev0000026>

CITATION

Clark, H. H. (2016, February 8). Depicting as a Method of Communication. *Psychological Review*. Advance online publication. <http://dx.doi.org/10.1037/rev0000026>

Depicting as a Method of Communication

Herbert H. Clark
Stanford University

In everyday discourse, people describe and point at things, but they also depict things with their hands, arms, head, face, eyes, voice, and body, with and without props. Examples are iconic gestures, facial gestures, quotations of many kinds, full-scale demonstrations, and make-believe play. Depicting, it is argued, is a basic method of communication. It is on a par with describing and pointing, but it works by different principles. The proposal here, called *staging theory*, is that depictions are physical scenes that people stage for others to use in imagining the scenes they are depicting. Staging a scene is the same type of act that is used by children in make-believe play and by the cast and crew in stage plays. This theory accounts for a diverse set of features of everyday depictions. Although depictions are integral parts of everyday utterances, they are absent from standard models of language processing. To be complete, these models will have to account for depicting as well as describing and pointing.

Keywords: depictions, gestures, quotations, make-believe play, staging scenes

This article is about depicting as a basic method of communication. When we think of depictions, we usually think of artifacts created at one place and time and exhibited at another—paintings, sketches, blueprints, maps, statues, movies, animated cartoons, TV sitcoms, or radio plays. These are *exhibited* depictions. In everyday discourse, however, people depict things with their hands, arms, head, face, eyes, voice, and body, with and without props. Examples include iconic gestures, facial gestures, quotations of all kinds, full-scale demonstrations, and make-believe play. These are depictions that people create *and* display with a single set of actions at a single place and time. These I will call *performed* depictions (often simply *depictions*).

The goal of this article is to offer a general theory of performed depictions—what they are and, in outline, how they are used and interpreted. We need such a theory for at least three reasons.

First, depicting is a basic method of communicating on a par with describing and indicating (Peirce, 1932, 1974; see Atkin, 2010; Clark, 1996). In describing, people use arbitrary symbols (e.g., words, phrases, nods, and thumbs-up) to denote things categorically, and in indicating, they use pointing, placing, and other indexes to locate things in time and space. In depicting, people create one physical scene to represent another. Although there are rich, comprehensive theories for describing and indicating, there is nothing comparable for depicting.

Second, depicting contrasts most directly with describing, but is based on a fundamentally different mode of thinking. Consider an interview, reported in the *New Yorker* (August 1, 2011), in which

Hollywood director WG told correspondent TF about having to stop filming in New York because of some falcons nesting on the ledge of a building:

“In L.A., they would have—” He leveled a finger at some imaginary nestlings and made a gun-cocking sound.

Although WG could have continued with the description “shot those falcons,” he continued with a depiction—a visible, audible model of shooting. In interpreting “shot those falcons,” TF would have imagined a scene based on the denotations of *shoot*, *those*, and *falcons*. For the depiction, he would have imagined a scene based instead on the scene WG created—the shape, orientation, and time course of his hand movements and tongue-click.

The contrast seems clear. With descriptions, people rely mainly on their knowledge of the vocabulary, syntax, and semantics of, say, English or Japanese and their ability to categorize. With depictions, they rely mainly on their visual, auditory, tactile, and proprioceptive knowledge of physical scenes and on their ability to use one scene in imagining another. Describing depends on knowledge of a language or code. Depicting does not.

Third, depicting is a common part of everyday language. In different studies of storytelling, narrators were found to produce iconic gestures 20 times per minute (Bavelas, Gerwing, & Healing, 2014), facial gestures in overlap with 30% of their words (Bavelas et al., 2014), and direct quotations 5.4 times per minute (Stec, Huiskes, & Redeker, 2015) or containing 17% of their words (Wade & Clark, 1993). And, as noted later, depictions are the core of children’s make-believe play.

Despite its prevalence, depicting is absent from standard accounts of language processing (see, e.g., Gleitman & Liberman, 1995; Goldrick, Ferreira, & Miozzo, 2014; Harley, 2013; Spivey, McRae, & Joanisse, 2012). Why? The assumption seems to be that depictions are not part of the syntactic or semantic composition of an utterance, so they can be set aside or ignored. And yet WG used his depiction of falcon shooting as if it were the verb *V* in “In L.A. they would have *V*.” Depictions like his—and they come in many forms—are genuine parts of everyday utterances and, therefore,

For helpful, often essential, suggestions on earlier versions of this work, I am indebted to Adrian Bangarter, Janet B. Bavelas, Philippe de Brabant, Richard Gerrig, Leila Takayama, Rolf Zwaan, three anonymous reviewers, and especially Eve V. Clark.

Correspondence concerning this article should be addressed to Herbert H. Clark, Jordan Hall, Building 420, Stanford University, Stanford CA 94305-2130. E-mail: clark@stanford.edu

ought to be included in accounts of language processing. *How* they are to be included is beyond the scope of this article.

What, then, *are* depictions? In the theory proposed here, called the *staging theory*, they are physical scenes that people stage for others to use in imagining the scenes depicted. WG staged the falcon-shooting scene so that TF would imagine the scene of someone in L.A. actually shooting falcons. Staging a scene for others is the same type of act that is performed by children in make-believe play and by the cast and crew in the theater.

Staging theory has several origins. The most immediate is the idea that quotations are demonstrations, the equivalent here of performed depictions (Clark & Gerrig, 1990; Wade & Clark, 1993; see also Clark, 1996; Holt, 2009; Tannen, 1989). That account was itself based in part on Goffman's (1974) notion of framing. A second origin is Walton's (1990) proposal that people interpret works of art by construing them as props in games of make-believe. A third origin is the rich literature on children's make-believe play.

The article is divided into four parts. First, I show that depictions are compositional parts of everyday utterances. Next, I argue that depictions are physical analogs of the scenes they depict and that they are scenes people stage for others. Third, I review evidence for staging theory from research on iconic gestures, facial gestures, quotations, full-scale demonstrations, and make-believe play. Finally, I return to a comparison of depicting and describing and characterize the problems depictions pose for accounts of language processing.

Depictions and Utterances

Why are depictions important? One reason is that they are integral parts of everyday utterances. Functionally, they may be *embedded* in utterances, *indexed* by linguistic expressions in utterances, or *adjoined* to linguistic expressions in utterances. They may also be used on their own as *independent* contributions to the discourse.

As evidence in this article, I will cite spontaneous depictions from a variety of sources. Many come from a lecture by MB, an eminent pianist, to a small audience on the art of playing the piano (Bilson, 2005); others come from tutoring sessions by three other eminent musicians (Duke & Simmons, 2006). In each example, the spoken words are in roman print; the visible and audible actions

are described in italics; and their overlap is underlined. Each example is labeled for easy reference.

Embedded Depictions

Embedded depictions are ones that function as parts of utterances—as if they were words, phrases, or other segments. Examples are shown in Table 1. In the one labeled *Booming*, MB depicted the actions of a former colleague named Smith, a depiction that functioned as the direct object of “hear.” Examples like this are traditionally called direct quotations or reported speech. In *Dee-duh-dum* and *Bartok*, the depictions were direct objects of “write” and “play.” All three functioned as if they were noun phrases. In *Oboist* and *Falcons*, the depictions functioned as if they were a noun (in “Now, you’ve got a *N*”) and a verb phrase (in “they would have *VP*”). All five assertions would have been incomplete without the depictions.

Indexed Depictions

Indexed depictions are ones that are introduced by expressions such as “this long,” “like this,” or “here” (see Fillmore, 1997; Klipple & Gurney, 2002). Two examples are listed in Table 2. In *Mozart*, MB depicted a passage from Mozart, which he indexed with “this.” In *Raised-voice*, he depicted Smith’s dramatic style of speaking by *fusing* the depiction with his utterance, and he indexed the depiction with “like this.”

Indexing a depiction is different from other deictic references. In *Mozart*, MB said “This is what I want to write” while playing a passage on a piano. Later he said “This is a late eighteenth century Viennese piano” while pointing at a piano. The referent of *this* in *Mozart* was a depiction that MB’s created for the occasion. The referent of *this* in the second utterance was a piano he did *not* create for the occasion. In *Mozart*, he used his hands and body to depict something; for the Viennese piano, he used his hands and body to locate something. Both were communicative acts, and both were composite parts of the utterances. Depicting things, however, is different from locating things (Atkin, 2010; Clark, 1996; Peirce, 1932).

Adjunct Depictions

Adjunct depictions are ones that are adjoined to parts of utterances as nonrestrictive modifiers. Two examples are shown in

Table 1
Examples of Embedded Depictions

Label	Example
<i>Booming</i>	[MB was speaking about a former academic colleague named Smith.] And you would go by his office and hear, “(MB turns his head and gaze 20 degrees left, looks at an imaginary person, and speaks in a booming voice with dramatic gestures) <u>Well, I can see you Wednesday at three o’clock.</u> ”
<i>Dee-duh-dum</i>	[MB was discussing a measure in a Mozart sonata.] But then he writes “(gazing at audience and singing) <u>dee-duh dum.</u> ” That is very expressive.
<i>Oboist</i>	[Oboist RK was speaking to an oboe student.] Now, you have got a (singing in rhythm) “ <u>Dee dah dee, (speaking) oom, (singing) bah duh dah.</u> ” You have got to feel that.
<i>Bartok</i>	[MB was discussing a piece played by Bela Bartok.] He does not play [demonstrates four measures on the piano while singing], but rather he plays—and he does it better than I do—[demonstrates the same four measures while singing, but differently].
<i>Falcons</i>	[Film director WK was speaking to reporter TF] “We were going to shoot on top of the MetLife Building . . . but falcons had nested there, so they shut it down. In L.A., they would have—” He leveled a finger at some imaginary nestlings and made a gun-cocking sound. (<i>New Yorker</i>)

Table 2
Examples of Indexed and Adjunct Depictions

Label	Examples from MB
<i>Mozart</i>	I am now Mozart. Forgive my immodesty. (<i>Sits down at piano.</i>) (<i>Plays two measures from Mozart sonata</i>) <u>This is what I want to write (2.8 s).</u> (<i>Stands.</i>) (<i>1.6 s</i>) How can I do it?
<i>Raised voice</i>	At the University of Illinois where I used to teach, there was a voice (<i>raising voice with RH gesturing dramatically toward audience</i>) <u>professor who said you should always speak like this.</u> [This was followed by <i>Booming.</i>]
<i>Tapering</i>	It means simply (<i>tracing a 30 cm long sideways V in the air with R thumb and forefinger near his R shoulder</i>) <u>what we would call the tapering.</u>
<i>Disgust</i>	[MB had just played an excerpt from Beethoven on the piano with the wrong phrasing.] (<i>showing disgust on his face and in his voice</i>) <u>It could not be that. I wouldn't think so.</u>

Table 2. In *Tapering*, MB's depiction was timed to overlap with the phrase "what we would call the tapering," and in *Disgust*, the depiction was timed to overlap with "It couldn't be that. I wouldn't think so." The expressions they are adjoined to are called *affiliates* (Schegloff, 1984). The depictions elaborate on their affiliates as if they were *non-restrictive relative clauses*—hence the term *adjunct*.

Adjunct depictions are different from indexed depictions. The carrier utterance in *Mozart* ("This is what I want to write") would not have been complete without the depiction, but the carrier utterance in *Tapering* ("It means simply what we would call the tapering") would have. Indexed depictions are mandatory, but adjunct depictions are not. And yet, adjunct depictions, like all nonrestrictive modifiers, add content to the discourse (see De Ruiter & Wilkins, 1998).

A critical difference between indexed and adjunct depictions is in what indexes what. In *Mozart*, the word "this" indexed MB's demonstration on the piano. The direction of indexing was *words-to-depiction*. In *Tapering*, it was the timing of the depiction that indexed the phrase "what we would call the tapering"; its timing, or temporal placement, was used as a *temporal index* (see Clark, 1996, 2005; Clark & Fox Tree, 2002). Here the direction was *depiction-to-words*.

Independent Depictions

Other depictions make independent contributions to the discourse. Two examples are shown in Table 3. In *Audition*, Morgan used the first quotation as a separate contribution to her story. In a corpus of oral narratives studied by Tannen (1986), 26% of the quotations were independent depictions (as in *Audition*), and the

Table 3
Examples of Independent Depictions

Label	Example
<i>Audition</i>	[Morgan was answering an interviewer's question about how she happened to audition for a role on MTV.] My friend Gary really talked me into it. "Come on. We gotta audition together. We gotta be on it together." And I was kind of like, "Whatever," you know. And um . . . (from <i>Laguna Beach</i> , MTV)
<i>Manifest</i>	[NT was asking a student about her interpretation of a musical phrase.] Pianist NT How does it manifest itself? Student (<i>plays a brief passage on the piano</i>) Pianist NT Oh! Okay.

rest were embedded depictions (as in *Booming*). And in *Manifest*, the piano student answered LT's question with a demonstration on the piano, so it, too, was an independent depiction (see Clark, 2012).

Content and Function

Every linguistic expression has both content and a function. Take the bracketed unit in MB's "I will now illustrate [a passage from Mozart]." Its content was a noun phrase formed from *a*, *passage*, *from*, and *Mozart*. Its function was its use as the direct object X_1 in "I will illustrate X_1 ." The function of the entire clause, in turn, was its use as a contribution to MB's lecture. With utterances like this, it is descriptions all the way down.

Whenever there are depictions, it is *not* descriptions all the way down. Compare "I will illustrate X_1 " with "He doesn't play X_2 " in *Bartok*. The units X_1 and X_2 have roughly the same function—both are direct objects of transitive verbs—but their contents are patently different. X_1 is a description, and X_2 is a depiction. Table 4 lists the content and function of eight of the examples cited so far.

Here, then, is the first hint of difficulty for models of language processing. None of them are capable of processing constituents that consist of playing the piano, singing an oboe score, or pretending to shoot falcons. Nor can they process adjunct depictions that are overlapped or fused with descriptions. These models will remain incomplete until they deal with all four types of depictions.

Foundations of Depicting

Depicting has been distinguished from describing in literary studies for over two millennia. In Plato's *Republic*, depicting, or *mimesis*, was contrasted with *diegesis*, or describing (see Bakhtin, 1981; Lodge, 1990). According to Lodge, *mimesis* is used in modern novels for "direct speech of characters," and *diegesis* for "authorial report, description, summary, commentary." For Plato, the epitome of *mimesis* was the theater. In comedies and tragedies, the cast and crew create nonactual scenes as a way of depicting scenes that are or could be actual.

Forms of Depicting

Depicting has also been distinguished from describing in scientific studies, though in five quite separate traditions:

1. *Iconic gestures*, as in *Tapering*, have been studied as a type of manual gesture (e.g., Cienki & Müller, 2008; Enfield, 2009; Kendon, 1980, 1994, 2004; McNeill,

Table 4
Content and Function of Eight Depictions

Example	Content	Function
<i>Booming</i>	Enactment of Smith speaking to student	NP in “You would hear NP”
<i>Bartok</i>	Performance of Bartok on piano	NP in “He doesn’t play NP”
<i>Mozart</i>	Performance of Mozart on piano	<i>Indicatum</i> of “this”
<i>Raised voice</i>	Enactment of a raised voice and dramatic gestures	<i>Indicatum</i> of “like this”
<i>Tapering</i>	Realization of a sideways V	<i>Adjunct</i> of “what we call the tapering”
<i>Disgust</i>	Enactment of a person grimacing	<i>Adjunct</i> of “It couldn’t be that. I wouldn’t think so.”
<i>Audition</i>	Enactment of Gary speaking	<i>Independent contribution</i>
<i>Manifest</i>	Performance of a phrase on piano	<i>Independent contribution</i>

1992, 2000, 2008; Streeck, 2008; Streeck, Groethues, & Villanueva, 2009). In their purest form, they contrast with (a) emblems, conventional gestures such as thumbs-up and crossed fingers, (b) deictic gestures, such as pointing or nodding at things, and (c) beat gestures, which mark the timing of certain speech events (see McNeill, 1992).

2. *Facial gestures*, as in *Disgust*, are often treated as expressions of emotion to be read off the face (e.g., Ekman, 1993; Ekman & Friesen, 1971). Many of these displays are really gestures, genuine acts of communication (Bavelas & Chovil, 2000; Chovil, 1991a,b).
3. *Spoken quotations*, as in *Booming*, have long been of interest to linguists, philosophers, and psychologists because of the problems they pose for theories of syntax, semantics, logic, and language processing (Cappelen & Lepore, 1997, 2003, 2012; Clark & Gerrig, 1990; Davidson, 1979; De Brabanter, 2005; Goodman, 1974; Partee, 1973; Potts, 2007; Recanati, 2001, 2008; Saka, 1998; Wade & Clark, 1993). In this literature, direct quotations are contrasted with indirect quotations. The first are depictions and the second are descriptions.
4. *Full-scale demonstrations*, as in *Bartok*, have been studied by psychologists and educators as a method of teaching (e.g., Csibra & Gergely, 2009; Gergely, Egyed, Király, 2007; Shakhshiri, 1989; Sokoloff & Thornton, 2004). Demonstrations like these contrast with narrative descriptions on the same topic.
5. *Make-believe play* has been studied mainly by developmental psychologists (e.g., Bretherton, 1989; Farver, 1992; Fein, 1981; Garvey, 1990; Garvey & Kramer, 1989; Harris & Kavanaugh, 1993; Leslie, 1987; Lowe, 1975; Mitchell, 2002). Children 24 months old are able to perform both independent depictions, such as “I want my teddy bear [whining in a petulant voice],” pretending to be a baby, and embedded depictions, such as “Say, ‘go to sleep now’” (Garvey, 1990, pp. 83–84).

Dividing depictions into these five classes, however, has led to parochial and often incompatible accounts. Theories that take quotations to be linguistic objects, for example, do not extend to iconic gestures, facial gestures, or make-believe play. And theories that take iconic gestures to be “visible actions” (see Kendon, 2004)

do not extend to quotations, demonstrations, or make-believe play. Depictions deserve a single, comprehensive account.

Depictions as Physical Analogs

Depictions, I argue, are *physical analogs* of what they represent. They are visible, audible, tactile, or proprioceptive models of things that one could actually see, hear, touch, or feel. The gesture in *Tapering* was a physical analog of a diminuendo sign, and the facial gesture in *Disgust* was a physical analog of a person showing disgust.

Performed depictions, like their exhibited cousins, are subject to two well-established principles:

1. *Pas-une-pipe principle*. A depiction is *not* what it depicts.
2. *Double-reality principle*. A depiction has two realities: its base, or raw execution; and its appearance, the features that are intended to be depictive.

The first principle is named for Magritte’s painting of a briarwood pipe, “Ceci n’est pas une pipe,” or “This is not a pipe.” The second is based on Gregory’s (1968, 1970) analysis of pictures into two realities: “A picture has a curious double reality, for we see its shapes, colors and textures as more-or-less familiar objects, while at the same time, we know we are seeing blobs of paint” (Gregory, 2005, p. 1233; see also Gombrich, 1960; Maynard, 1994; Walton, 1990). By these principles, depictions entail three scenes—a *base scene*, a *proximal scene*, and a *distal scene*—with mappings from one scene to the next.

Proximal and Distal Scenes

According to the *pas-une-pipe* principle, a depiction is different from what it depicts. In *Booming*, the depiction was a scene in which MB produced “(in a loud voice and with rhetorical gestures) Well, I can see you Wednesday at three o’clock.” What he depicted, however, was a scene in which Smith made an appointment with a student. These are the proximal and distal scenes:

Proximal scene (*the depiction*) ⇒ distal scene (*what is depicted*)

The proximal scene is in the here-and-now, and the distal scene in a there-and-then—in a world that is displaced in place, time, or reality.

Both of these are *physical* scenes—configurations in space and time of physical elements that we could, in principle, perceive with our senses. These elements may include people (e.g., MB or Bartok), physical objects and substances (e.g., a piano), people's actions (MB singing, Smith speaking), other physical events (a metronome ticking), or physical states (the size of the room).

The proximal scene is a physical analog of the distal scene. When physicist Ernest Rutherford first characterized the atom, he chose the analogy “The atom is like the solar system” (Gentner, 1983): The solar system is a physical analog of the atom. The mapping from proximal to distal scenes, denoted by double arrows (\Rightarrow), is shown in Table 5. Performed depictions are no different. In *Booming*, MB intended the proximal scene, his dramatization, to be mapped into the distal scene, Smith making an appointment.

All analogies are selective—often highly selective (see Gentner, 1983; Holyoak & Thagard, 1989). The proximal scene represents only part of the distal scene; for example, there are no elements in the solar system that map into protons or neutrons. Furthermore, only some aspects of the proximal scene map into the distal scene, where an aspect is a feature that holds for an entire scene. The aspect “3D geometry,” for example, maps from the solar system to the atom, but “size” does not. Rutherford took full advantage of this selectivity. The analogy he chose did what he wanted it to do and nothing more. People choose depictions for the same reasons.

Base and Proximal Scenes

By the double-reality principle, every proximal scene is built on a base scene, as illustrated for the sketch of a hydrogen atom in Figure 1:

1. The *base scene* is the raw, observable, yet-to-be-interpreted physical features of the depiction as executed. It is what the depiction *actually* is. In my sketch, it includes the paper, pencil marks, and smudges. Without explanation, these could mean anything.
2. The *proximal scene* is a construction, one that is based on the elements, aspects, and features of the base scene that are depictive. It is what the depiction *appears* to be. In my sketch, the proximal scene includes—in their idealized form—the large disk, small spot, and circle as props for the hydrogen nucleus, electron, and electron's orbit. It does not include the border, the initials, or the aspect “black-and-white.”

The mapping from base to proximal scenes is *functional*: It specifies how the parts of the base scene are to function in the

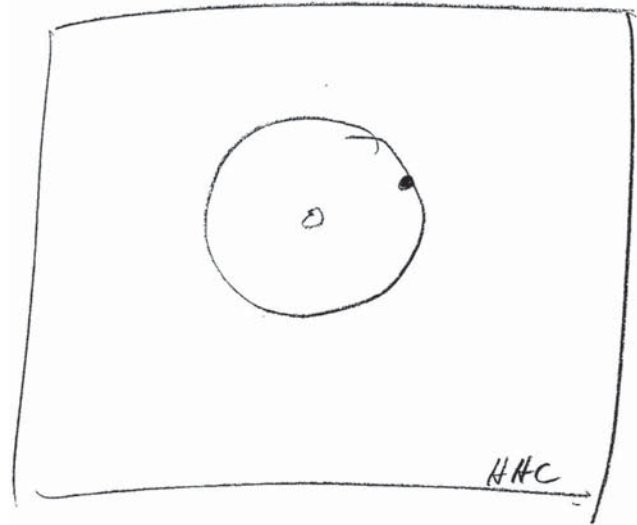


Figure 1. Sketch of hydrogen atom.

proximal scene. Here functional mappings are denoted with single arrows (\rightarrow) to distinguish them from analog mappings (\Rightarrow). Table 6 shows a functional mapping for the sketch in Figure 1.

Every base scene is construed according to an *interpretive framework*—the assumptions the creator intended to be used in the interpretation of the depiction. I established an interpretive framework for the sketch in Figure 1 by labeling it a “hydrogen atom,” implying the mapping in Table 6. If I had labeled it a “bicycle wheel,” I would have implied these rules instead:

Large disk \rightarrow prop for axle

Small spot \rightarrow prop for tire valve

Curved line \rightarrow prop for tire

There is no such thing as a depiction *simpliciter*. One cannot know what a base scene depicts without knowing or inferring what its creator *intended* it to depict—the interpretive framework (Bloom, 2010; Gelman & Bloom, 2000; Gregory, 1968, 2005; Preissler & Bloom, 2008).

Imagining What Is Depicted

The functional and analog mappings of a depiction (illustrated in Figure 2) are really a *system* of mappings: Each one constraints

Table 5
Analog Mapping for “The Atom Is Like the Solar System”

Proximal element	\Rightarrow	Distal element
Sun	\Rightarrow	Nucleus
Planet	\Rightarrow	Electron
Attracts(sun, planet)	\Rightarrow	Attracts(nucleus, electron)
Orbits(planet, sun)	\Rightarrow	Orbits(electron, nucleus)
\emptyset	\Rightarrow	Proton
Aspect(3D-geometry)	\Rightarrow	Aspect(3D-geometry)
Aspect(size)	\Rightarrow	\emptyset

Table 6
Functional Mapping for the Sketch of Hydrogen Atom in Figure 1

Base element	\rightarrow	Proximal element
Large disk	\rightarrow	Prop for nucleus
Small spot	\rightarrow	Prop for electron
Curved line	\rightarrow	Prop for orbit (electron, nucleus)
Rectangular frame	\rightarrow	\emptyset
“HHC”	\rightarrow	\emptyset
Aspect(black-and-white)	\rightarrow	\emptyset

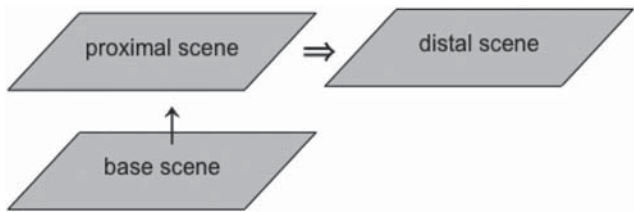


Figure 2. Relations among the base, proximal, and distal scenes of a depiction.

the other. Establishing this system takes both perceptual and conceptual imagination.

In perceptual imagining, we *imagine seeing, hearing, touching, or feeling* things (Walton, 1990). We can imagine seeing friends, listening to music, swallowing oysters, running, falling out of a window, feeling tense, and feeling sad. These are quasi-perceptual experiences (Shepard, 1984), and they activate the corresponding perceptual systems in the brain (Ganis, Thompson, & Kosslyn, 2004; Kosslyn, Ganis, & Thompson, 2001; Halpern & Zatorre, 1999; Zatorre & Halpern, 2005). In *Tapering*, the physical base of MB’s depiction helped his audience imagine seeing a large “V” on its side.

In conceptual imagining, on the other hand, we *imagine that* certain propositions are true (see also Gendler, 2011). We might imagine that today is the first of July 1620 or that Churchill was a secret Baptist. In *Tapering*, it was the interpretive framework that led MB’s audience to imagine that the sideways V was a prop for a diminuendo sign. Evidence has long shown that perceptual imagining is constrained by conceptual imagining, and vice versa (e.g., Bruner, Busiek, & Minturn, 1952; Carmichael, Hogan, & Walter, 1932), though there is much debate about how.

Perceptual imagining is often bound to physical stimuli in a process I will call *percept-bound imagining*. In reading tea leaves, fortune tellers “allow the imagination to play around the shapes suggested by them” (Melton, 1996). Once they discover a wedge shape (a base scene), they might construe it as a prop for Cupid’s arrow (the proximal scene) and predict romance. If, instead, they discover a circle, they might construe it as a prop for a wedding ring and predict marriage. It would be difficult, perhaps impossible, to construe the wedge as a prop for a wedding ring, or the circle as a prop for an arrow. This is one way conceptual imagination is constrained by perceptual imagination, and vice versa.

Depictions rely on percept-bound imagination. In *Tapering*, MB designed his gesture (the base scene) to bind his audience into imagining a sideways V. He could not have done that with an up-and-down or circular gesture. At the same time, he relied on the interpretive framework to get his audience to construe the V-ness of his gesture as a prop for a diminuendo sign.

Depictions as Staged Scenes

Depictions may be physical analogs, but this is only a necessary condition. A depiction is an act of communicating the appearance of a distal scene, and people do that by staging a proximal scene that is a physical analog of the distal scene. In *Booming*, MB played the role of Smith and enacted him facing a student and booming out “I can see you Wednesday at three o’clock.” In

Tapering, he took the role of prop-manager and created a visual prop for a diminuendo sign. Why did he stage these scenes? In *Booming*, it was to get his audience to imagine seeing and hearing Smith make an appointment, and in *Tapering*, it was to get them to imagine seeing a diminuendo sign. The staging theory of depicting is this: To perform a depiction is to stage a scene for recipients to use in imagining the scene depicted.

Performed depictions are nonactual scenes. In *Booming*, MB did not “really or actually or literally” make an appointment (Goffman, 1974; see Bateson, 1955). He created the mere appearance of someone doing so. These scenes range from rich, vivid, and detailed, as when MB demonstrated Bartok on the piano, to sparse, incomplete, and skeletal, as when MB sang “dee-duh dum” for three notes from Mozart. As with any communicative act, MB intended his audience and him to mutually recognize his intention in staging these scenes (Grice, 1957).

Staging Scenes in Make-Believe Play

Staging a scene should be a thoroughly familiar notion because it is what children do in make-believe play. Consider the episode called *Go-to-sleep* in Table 7 (from Garvey, 1990, pp. 83–84). Sally began by pretending that she was Mother, the doll was Baby, and Mother was putting Baby to bed. She then pretended that Alex was Father, and Father was telling Baby to go to sleep. And she pretended that she was Baby asking Father “Why?” Alex joined her and pretended that he was Father speaking to Baby.

What Sally and Alex did was stage a scene—a brief play about Mother and Father putting Baby to bed. As stage manager, Sally assigned the roles of Mother, Baby, and Father to Alex and herself, and introduced the doll and cradle as props. As script writer, she invented the dialog, and later, as actor, she played Mommy putting Baby to bed and Baby whining “Why?” Alex accepted Sally’s stage directions and, as actor, played Father. The children engaged in two types of activities: (a) as stage-managers *outside* the scene, they established what went into the scene; and (b) as actors *within* the scene, they played the roles of Mother, Father, and Baby.

The heart of the episode was really an elaborate depiction. The proximal scene was a series of actions in which Sally played Mother and Baby and Alex played Father, and it mapped into a distal scene in which Mother put Baby to bed, Baby whined, and Father spoke. The episode was planned and produced with individual depictions. Sally performed embedded depictions in “Say, ‘go to sleep now’” and “No, say ‘Because (*emphatically*).’” She also performed independent depictions of Mother putting Baby to

Table 7
Go-To-Sleep (an Episode of Make-Believe Play)

[Sally (3;3) was holding a baby doll and Alex (2;9) was watching her put it to bed.]	
1. Sally	Say, “Go to sleep now.”
2. Alex	Go sleep now.
3. Sally	[<i>in role of baby, whining</i>] Why?
...	
6. Alex	Because.
7. Sally	No, say “Because.” (<i>emphatically</i>)
8. Alex	Because. (<i>emphatically</i>)
...	

bed and Baby whining “Why?” As Garvey (1990, p. 86) put it, “Carrying out . . . make-believe is largely a matter of communication.”

Make-Believe Play as Theater

What, then, is make-believe play? In most accounts, it is taken to be a type of theater. It was described as “socio-drama” by Piaget (1962), “dramatic play” by El’konin (1969), and enactments of social roles by Vygotsky (1978), three pioneers in the field. It has since been characterized with a range of concepts from the theater, as in these examples (variously from Bretherton, 1984, 1989; Dunn & Dale, 1984; Garvey, 1990; Giffin, 1984; Haight & Miller, 1993; Harris, 2000; Lillard, Pinkham, & Smith, 2011; Miller & Garvey, 1984; Nelson & Seidman, 1984; Sawyer, 1997; Strömqvist, 1984; Taylor, 1999; Taylor & Mannering, 2006):

1. *Roles*: Children are said to take the “roles” of such “characters” as mothers and cowboys, Santa and Cookie Monster, drivers and passengers.
2. *Acting*: Children are said to “enact” roles or do “pretend actions of play characters.” They might “play” doctor or fireman.
3. *Props*: Children are said to use “props” such as dolls, toy beds, and blocks, objects that stand in for real objects.
4. *Scripts*: Children are described as establishing “story lines” or “scripts,” such as “playing house” or “playing cops and robbers,” which result in “improvised theater.”
5. *Staged versus real actions*: Children distinguish play from nonplay. They treat play as a “pretend world,” “subjunctive world,” or “alternative world,” as “non-literal,” or a “suspension of objective truth.”
6. *Acting versus stage managing*: Children distinguish “acting” (e.g., Sally’s whining “Why?”) from “stage directions” (e.g., Sally’s “Say, ‘Go to sleep now’”).

Make-believe play changes as children grow up (see, e.g., Garvey, 1990). The earliest forms are primitive, and are initiated and sustained by caregivers. Later forms are elaborate, self-initiated, and self-sustained. Children and young adults take part in ever more sophisticated and institutionalized forms of play—ritualized games, games with rules (Garvey, 1990), computer games (Goldstein, 1994), and, of course, theater (Goffman, 1974). The assumption is that the theater is really a highly evolved, institutionalized form of make-believe play.

The theater, like make-believe play, is a family of forms. There are stage plays, silent mime shows, puppet shows, radio plays, play readings, improv theater, street theater, and skits. All of these depict events. There are also *tableaux vivants*, an obsolete type of theater, that depict states, or instantaneous time-slices of scenes. Theater pieces may have many actors, one actor, or none, a large stage crew or none, many props or none, a script or no script, and they may be fiction or nonfiction. The theater represents not only what is common to staged scenes, but how they vary.

Logic of Staging Theory

The evidence I will cite for staging theory is based on this logic. If depictions are a form of make-believe play, and the theater is an evolved, institutionalized form of make-believe play, then depictions should reflect both what is common and what varies in staging scenes in the theater. The strategy here is to show how principles from the theater account for many features of everyday depictions. These principles fall into six areas:

1. *Participants*, such as actors, prop-managers, and spectators;
2. *Actor versus prop depictions*;
3. *Props*, both physical and imaginary;
4. *Spatial frames of reference* within scenes and within the discourse;
5. *Temporal frames of reference*, also within scenes and within the discourse;
6. Actual and displaced *worlds* in discourse.

As a characteristic example of the theater, I will refer to a performance I saw in London of Samuel Beckett’s *Waiting for Godot* in which the characters Estragon and Vladimir were played by Ian McKellen and Patrick Stewart.

Actors and Props

Those who execute scenes in the theater are either *actors* or *prop-managers*. The actors enact characters *inside* the scenes, working on stage in plain sight. The prop-managers realize props from *outside* the scenes. Even when the prop-managers are visible or audible, they are treated as invisible and inaudible. Actors have identities; prop-managers do not.

Staged scenes also typically include *spectators*. When I saw *Godot*, I, too, was a participant. As Jorge Luis Borges (1960/1999, see Walton, 1978) put it: “the actor . . . on the stage plays at being another before a gathering of people who play at taking him for that other person (p. 300).” Therefore, while McKellen and Stewart were playing Estragon and Vladimir, we in the audience were playing our own roles.

What, then, are these roles? The stage is usually treated as having a fourth wall. The pretense is that the actors on stage are unaware of the spectators, yet the spectators can see and hear everything that happens on stage. In the physical analog, then, the spectators play the role of *covert observers* of the distal scene. At *Godot*, we in the audience played characters who covertly watched Estragon and Vladimir as they bantered in the countryside. The general mappings are these:

Cast member → actor ⇒ character

Cast member’s action → enactment ⇒ character’s action

Prop-manager → ∅

Recipient → spectator ⇒ covert observer

Actor and Prop Depictions

If depictions are staged scenes, they, too, should be performed by actors or prop-managers, and they are. The two basic formats are illustrated by *Booming* and *Tapering*, repeated in Table 8. In *Booming*, MB placed himself inside the proximal scene and enacted what Smith said. In *Tapering*, MB placed himself outside the proximal scene and realized a V-shaped prop inside it. The first I will call an *actor depiction*, and the second, a *prop depiction*:

- An *actor depiction* is the selective enactment, by a producer within the staged scene, of a character in a distal scene.
- A *prop depiction* is the selective realization, by a producer outside the staged scene, of one or more props for a distal scene.

Actor and prop depictions are named for the main element the producer controls—an actor, or props. If the producer is an actor, the depiction is an actor depiction; if not, it is a prop depiction. Actor depictions can be further classified by the characters they enact (by whose persona, or mask, they put on): the producer him- or herself, an addressee, or a third person. These are 1P, 2P, and 3P actor depictions. Every character in a distal scene is associated with an *I*, *here*, and *now*, which anchor the character and his or her actions to that scene (see Bühler’s, 1934 notion of *origo*). In *Booming*, the person MB was quoting (“I”) was Smith, so this was a 3P actor depiction. In *Audition*, when Morgan said, “And I was kind of like ‘Whatever’ you know,” the person she was quoting was Morgan herself, so this was a 1P actor depiction (for further divisions, see Engberg-Pedersen, 2015). The characters depicted may be robots, bears, dogs, or other creatures. In make-believe play, one child enacted a baby fish with “‘Hello, I’m a baby fish’ (said in a high squeaky voice while making the fish swim toward the partner)” (Garvey, 1990, p. 88).

Iconic Gestures

Actor and prop depictions have been distinguished in iconic gestures for over a century (Kendon, 2004, Chapter 6). They have been called, among other things, *kinetographic* versus *iconographic gestures* (Efron, 1941), *kinetic* versus *pictorial illustrators* (Ekman and Friesen (1969), and gestures with a *character* versus an *observer* viewpoint (McNeill, 1992; for more on viewpoints, see later). The difference between the two types is clear. In retellings of *Snow White* recorded by McNeill (1992), one narrator enacted Snow White sweeping the floor by pretending to hold a full-sized broom and making sweeping motions. This was a 3P actor depiction. Another narrator depicted a person walking by wiggling two fingers pointed downward across an imaginary surface. This was a prop depiction. Both gestures depicted people, but the first was by an actor inside the scene, and the second, by a prop-manager outside the scene. As McNeill put it, the first “incorporates the

speaker’s body into the gesture space,” and the second “excludes the speaker’s body from the gesture space” (p. 119). This is a criterial difference between actor and prop depictions.

Quotations

Everyday quotations also come as both actor and prop depictions. The prototypical quotation is reported speech, a 3P actor depiction of what someone said. In *Booming*, MB enacted Smith, a third party, booming out, “Well, I can see you Wednesday at three o’clock.” Other quotations are 1P or 2P actor depictions. In one narrative, a speaker depicted herself screaming and gesturing dramatically, “I went out of my mind and I just screamed I said ‘Take that out! That’s not for me!’” (Polanyi, 1989, p. 92). In *Go-to-bed* (see Table 7), Sally told Alex “Say ‘Go to sleep now’” to which he responded “Go sleep now.” Her quotation was a 2P actor depiction, and his line was a 3P actor depiction. Other quotations depict actions that are *not* linguistic. In *Bartok*, MB produced two 3P enactments, not of a person speaking, but of a person playing the piano. His lecture was filled with such quotations, including 1P actor depictions such as “I might play [*demonstration on the piano*]” and 2P actor depictions (as in *Bink-bink* cited later). Some of these were embedded depictions; others were independent depictions. Still other quotations are *prop* depictions—depictions of objects, states, or events by a producer outside the scene. The examples in Table 9 are from callers talking to the hosts of the American radio program *Car Talk*. Both callers depicted car noises, the first embedded as a noun in “the car has a kind of a X,” and the second embedded as an adverb in “it goes X.” And in *Dee-duh-dum*, MB depicted three notes from a Mozart sonata embedded as the direct object in “he writes X.” Examples like these are further evidence against the widely held reproduction theories of quotation (e.g., Cappelen & Lepore, 1997, 2003, 2012; Davidson, 1979; Geurts & Maier, 2003; Goodman, 1978; Pagin & Westerståhl, 2010; Partee, 1973; Potts, 2007; Saka, 1998). According to these theories, quotations are the mention or reproduction of linguistic expressions. The materials quoted in *Bartok*, *Dee-duh-dum*, *Worry-noise*, and *Purring*, of course, are not linguistic expressions, or even forms of communication (see Clark & Gerrig, 1990; De Brabanter, 2005; Holt, 2009; Recanati, 2001, 2008; Tannen, 1989; Wade & Clark, 1993).

Hybrid Depictions

In the theater, a single person can perform two or more roles at once. In *Punch and Judy* shows, the puppeteer is both a prop-manager, manipulating the puppets, and an actor, enacting the puppets’ speech. Ventriloquists are also both prop-managers and actors as they manipulate and speak for their dummies. These are hybrid depictions, and in everyday discourse, they can be actor-actor, actor-prop, or prop-prop depictions.

Table 8
Paradigmatic Actor and Prop Depictions

Actor depiction (<i>Booming</i>)	[MB was speaking about a former academic colleague named Smith.] And you would go by his office and hear, “(in a booming voice with dramatic gestures) Well, I can see you Wednesday at three o’clock.”
Prop depiction (<i>Tapering</i>)	It means simply (tracing a sideways V in the air with R thumb and forefinger) <u>what we would call the tapering.</u>

Table 9
Prop Depictions of Car Noises

Label	Example
<i>Worry-noise</i>	My son and I noticed that the car has a-uh, a worry noise, kind of a “(low growling noise)” like when you used to put a baseball card in a-uh, the spokes of your bike.
<i>Purring</i>	And then, about one and a half or two seconds into my foot being on the accelerator, it goes “[gargling sound],” has a real deep lion-purring sound to it.

Hybrid depictions are ubiquitous in make-believe play. In *Go-to-bed* (see Table 7), Sally managed the doll and, at the same time, whined “Why?” in the doll’s voice. Like a puppeteer, she was both prop-manager and actor, creating an actor-prop depiction. Another girl pushed a toy car toward her brother and said “Beep. Brmm” (Dunn & Dale, 1984, p. 154) This was a prop-prop depiction; one prop was the car and its movements, and the other was its sounds.

Hybrid, or dual, depictions are also well documented in adults (see Dudis, 2004; McNeill, 1992; Parrill, 2009; Perniss, 2007; Perniss & Özyürek, 2008; Stec, 2012). Three examples are shown in Table 10:

1. In *Bink-bink* MB enacted a person striking piano keys (a 2P actor depiction) and, simultaneously, depicted the sounds of the keys being struck, “Bink bink bink bink” (a prop depiction).
2. In *Conductor*, MB sang “dee-duh dum” (a prop depiction) while gesturing like an orchestra conductor (a 3P actor depiction). (MB’s in-breath was a prop depiction of the silence between two notes.)
3. In *Lift-your-chin*, Seigny enacted *herself* lifting her chin (a 1P actor depiction) and, simultaneously, enacted a *cinematographer* placing the tips of his fingers under her chin and saying, “Lift your chin” (a 3P actor depiction in two parts).

How, then, do people perform and interpret actor and prop depictions? Let us start with actor depictions.

Actor Depictions

Actors are usually described as playing characters, but that is not right: As with any analog, they enact only a selection of the characters’ actions. In *Godot*, McKellen did not play Estragon as

a whole. All he enacted were Estragon’s vocal, facial, and bodily actions. In a radio play, all he would have enacted would have been Estragon’s vocal actions. Everyday depictions are just as selective. In *Booming*, MB enacted Smith’s posture, head movements, arm, hand, and facial gestures, sentence content, and voice amplitude, and yet moments earlier, in *Raised voice*, all he enacted was Smith’s hand gestures and voice amplitude.

So producers select a subset of a character’s actions to enact, but which ones? Many selections are formed around a section of the body—the voice, face, head, hands and arms, torso, or entire body.

Vocal Depictions

In depicting a person’s vocal action, producers have options in which of its aspects to enact. Although the quotations in Table 11 look simple enough, they are not what they seem.

1. In *Tiffany’s*, the narrator had memorized verbatim what the clerk had said, and yet she enacted the force of the clerk’s question and not his wording. Why? To fit the style of her ongoing narrative (Wade & Clark, 1993; see also Holt, 2000, 2009).
2. In *Rolled-r*, the speaker did not enact any of Paisley’s *actual* utterances. He invented an utterance with many r’s to depict the way Paisley rolled his r’s (Clark & Gerrig, 1990, p. 799).
3. In *Zschau*, what Ed Zschau enacted was not a particular utterance, but an instance of the *type* of question he had often been asked (Clark & Gerrig, 1990, p. 773).
4. And in *Haydn*, MB enacted what Haydn said, not in German, but in an English version his audience could understand. Here the aspect “language of producer” was not depictive, but supportive.

Table 10
Examples of Hybrid Depictions

Label	Example
<i>Bink-bink</i>	[MB was speaking to a woman in the audience.] All you have to do is say, “Look, there are four notes here, but it’s not really very natural to go ‘(thrusting R index finger four times in rhythm onto an imaginary piano keyboard) bink bink bink bink.’”
<i>Conductor</i>	[MB was speaking about a phrase from a Mozart sonata.] But “(singing and directing with hands) <i>dee-duh dum</i> ” has this little “(in-breath)” in it, which is so nice.
<i>Lift-your-chin</i>	[Actress Chloe Seigny was describing how a cinematographer told her to hold her chin up during a scene he was shooting.] Before every take, the (directing her hand, palm down, sagittally toward her throat, patting her chin from underneath, and lifting her chin with each pat) cinematographer would be like, “Lift your chin.” So when I watched the movie, I’m . . . (from the <i>Conan O’Brien Show</i>)

Table 11
Direct Quotations of Vocal Actions

Label	Example
<i>Tiffany's</i>	[Speaker was narrating a scene from <i>Breakfast at Tiffany's</i> and knew the script <i>verbatim</i> . In that scene, the clerk said, "May I ask how limited?"] And the clerk says, "Well, how much can you spend?"
<i>Rolled-r</i>	Old Jim Paisley goes, "oh dear oh dear it must be the beer," he always says that.
<i>Zschau</i>	Many people have come up to me and said, "Ed, why don't you run for the Senate?"
<i>Haydn</i>	[MB was speaking to DN.] Do you really think that Haydn went back from London to Vienna and said, "Now listen everybody. I wrote all these pieces in London, (<i>raising pitch of voice</i>) and I don't want any of you to play them on the pianos that you have here"?

Speakers can also select the dialect, lisping, drunkenness, and other aspects of a person's voice (Clark & Gerrig, 1990). In narratives, speakers often go further and dramatize a character's attitude, emotion, or tone of voice (see later).

Facial Depictions

Most facial gestures are actor depictions. In *Disgust*, MB had just finished a bad rendition of Beethoven when he turned to his audience and, putting on a look of disgust, said, "It couldn't be that. I wouldn't think so." He used the 1P actor depiction as an adjunct to his assertions. Note that he depicted his disgust, not during the bad rendition of Beethoven when he might have been disgusted, but during his assertions about it. Facial gestures of emotions are generally performed in the *absence* of those emotions (Bavelas & Chovil, 2000; Chovil, 1991a, 1991b). They are there to depict emotions, not express them.

Everyday conversations are filled with facial gestures. In a sample of 1,000 such gestures (Chovil, 1991a), 50% of them depicted how people felt. Two examples are shown in Table 12. In *Desserts* (Chovil, 1991a, p. 180), the speaker performed a 1P actor depiction of the face as an adjunct to her assertion about desserts. And in *Looked-down* (Chovil, 1991a, p. 182), the speaker "re-enacted the [guy's] look by tilting his head down, lowering his eyes, and forming a straight mouth" as a 3P actor depiction, which he embedded in his utterance as an adjective. Other facial gestures depicted "slight disgust," "disbelief and surprise," "amazement," and "pain."

The facial gesture in *Disgust* belongs to a class of 1P actor depictions that make public what the producer is thinking. Not all of these depict affect. One type is the *thinking face*, as in Table 12 (from Goodwin & Goodwin, 1986). Allan was speaking to Barbara when he gazed to one side and put on a conventional expression of someone thinking hard. He initiated the gesture at "uhm" and ended it at "tch!" to tell Barbara that he was searching for a word. Another such gesture is the so-called *funny face* to depict uncertainty (Krahmer & Swerts, 2005; Swerts & Krahmer, 2005). Both

of these gestures are marked by features such as depressed lip corners, stretched lips, widened eyes, moving eyebrows, and averted gaze (see also Chovil, 1991a).

Facial gestures can also be 2P actor depictions. In an experiment by Bavelas, Black, Lemery, and Mullett (1986), individual students watched a research assistant carry a large TV set into the lab room and, in a carefully contrived accident, drop it on his finger. The assistant, feigning pain, then did one of two things: He looked up at the student, or he hunched down without looking at the student. As captured on video, most students began forming a grimace a fraction of a second after the onset of the accident. When the assistant looked up, the nascent grimaces turned into full grimaces, but when he did *not* look up, they quickly faded to nothing.

The full blown grimaces were 2P actor depictions. If the students themselves had experienced pain, *all* of the nascent grimaces should have become full blown, but they did not. The students used the full blown grimaces to enact the emotion they attributed to the assistant ("you") as a way of communicating sympathy—"I feel your pain" (Bavelas et al., 1986; Bavelas, Black, Chovil, Lemery, & Mullett, 1988; Bavelas, Black, Lemery, & Mullett, 1990). A few students combined their grimaces with "Ouch!" in composite 2P actor depictions.

Broader Depictions

Vocal, facial, and manual enactments are often part of broader depictions, as illustrated in Table 13. In *Schubert*, DN did more than enact the tenor singing a few words from Schubert's "Winterreise." He combined the singing with gestures and posturing toward an imaginary concert audience. MB did much the same in *Booming*. And in *Come-out*, Lauren enacted not only Lo's conspiratorial speech style, but her conspiratorial face and three "come-here" gestures. Dramatizations like these are common in narratives (Bavelas, Gerwing, & Healing, 2014; Holt, 2000, 2009; Stec, Huiskes, & Redeker, 2015; Tannen, 1989; Wade & Clark, 1993).

Table 12
Facial Gestures as Actor Depictions

Label	Example
<i>Desserts</i>	(Speaker "raised her eyebrows, then raised one side of her upper lip, and squinted her eyes, which was seen as illustrating her dislike of the dessert") I hate, I hate desserts with alcohol in them.
<i>Looked-down</i>	[Speaker was describing "how a person had given him a particular kind of 'look.'"] . . . and the guy just sort of looked "[facial gesture]" you know, sorta looked down on me
<i>Thinking- face</i>	He put (gazing to one side with a thinking face) <u>uhm</u> , (.7) <u>tch!</u> Put crabmeat on th'bo:dum [bottom].

Table 13
Actor Depictions With Gestures

Label	Example
<i>Schubert</i>	[DN was telling MB about a tenor in a recital.] He found he could do his “(breaking eye-contact, turning 60 degrees left, gesturing dramatically toward an invisible audience, and singing) <u>Fremd bin ich ein-</u> ” in a much more conversational way . . .
<i>Come-out</i>	[Lauren, a California high school student, was telling an interviewer about her friend Lo.] I could see her through the window and she was trying to get my attention, and she was like “(breathlessly, with L hand and arm pointing upward, making three “come-here” motions toward her R shoulder) <u>Come out, come out.</u> ” (from <i>Laguna Beach</i>)

In staging theory, what producers are trying to do is stage scenes and not simply collections of aspects. If so, the aspects they combine should form coherent scenes, and they seem to. In *Schubert*, DN surely did not enact the tenor’s voice, gestures, and posture as three separate actions and then synchronize them. He enacted the bundle of actions as a whole.

Joint Depictions

In the theater, actors sometimes violate the fourth wall and enlist the audience as coactors. In American melodramas, actors invite the audience to cheer the hero and boo the villain, and in British pantomimes, some actors interact with the audience, and others do not (Taylor, 2007). In immersive theater, the audience is often treated as coactors (see White, 2012), and in Disneyland, the actors who play Mickey Mouse and Snow White speak freely with visitors.

The same happens in everyday depictions. In *Manager* (see Table 14), Ellen incorporated Kate into the proximal scene of her depiction and treated Kate as if she were the person Lorna was speaking to. That is, Ellen, in her role as Lorna, enlisted Kate to play the role of Lorna’s addressee. For similar examples, imagine Allan telling Sam “She was so mad that she shook me *like this* (shaking Sam by the shoulders)” or “The ball grazed my shoulder *like this* (brushing Sam’s shoulder with one hand).” The first is a 3P actor depiction, and the second, a prop depiction.

Joint depictions are a common part of many demonstrations. In a self-defense class analyzed by Stukenbrock (2012), the instructor demonstrated body throws by enlisting students and confederates as coactors. Furthermore, joint depictions are the core of much make-believe play. In *Go-to-bed* (see Table 7), Sally and Alex performed joint enactments throughout. Children are able to do that from about two-and-a-half (Garvey, 1990; Miller & Garvey, 1984).

Table 14
Example of a Joint Depiction

Label	Example
<i>Manager</i>	[Ellen was speaking to Kate.] And my friend Lorna, if something’s wrong like that, the first thing she’d say, “(pointing index finger at Kate, and looking Kate in the eye) <u>Give me the manager. Let me talk to your manager.</u> ”

Note. Example from video-recording by Kashmiri Stec (personal communication).

Physical Props

In the theater, most staged scenes have *props*. These are elements in proximal scenes (other than the actors and enactments) that map into real elements in distal scenes. Props can be physical or imaginary. In *Godot*, the physical props included a road, a tree, greasepaint, scenery, a moon rising, evening light. The imaginary props included a person McKellen created by looking off stage at someone in the distance. Physical props may be: both visual and auditory, as in *Godot*; visual only, as in silent mime shows; auditory only, as in radio plays; or even visual, auditory, tactual, and proprioceptive, as in immersive theater. And in silent mime shows, most props are imaginary.

Props can be further classified as *center*, *side*, or *back* props. Center props are those that are realized by prop-managers alone, such as the offstage noises and rising moon in *Godot*. Side props are the ancillary objects or events actors use in the course of their enactments, such as McKellen’s boots and bowler hat. *Back props* are those that do not change during a scene, such as the mound, tree, road, and scenery in *Godot*. Furthermore, physical props can range from genuine to figural, as illustrated in Table 15. If depictions are staged scenes, they, too, should include both physical and imaginary center, side, and back props, and they do. I begin with physical props.

Status of Props

First, a warning. It is often stipulated that for an action to be an iconic gesture, or a quotation, it must be performed *without* physical props. In a theory of depictions, this stipulation makes no sense. When I say “George holds his pen like this,” I could mold my hand *as if* I were grasping a pen, or I could mold my hand around a physical pen. The two actions are surely different, and yet both depict George holding a pen. The second is no less a depiction—or a gesture à la Kendon (2004)—simply because the prop is physical.

Furthermore, physical props are not tools. People can use pencils, sticks, or chalk for drawing pictures, and computers for creating pictures or sounds. These are tools, not props, because they do not map into elements in the distal scene. The focus here is on depictions that do not require tools.

Center Props

Center props are those that are realized by producers from outside the scene, and the depiction is centered on the prop itself. The props people use depend in part on availability. In a study by Engle (2000), tutors were asked to instruct students on how a door

Table 15
Five Types of Physical Props (Illustrated for a Prop Pistol and Prop Pistol Shot)

Prop type	Definition	Prop pistol	Prop pistol shot
1. Genuine	Real exemplar	Actual pistol	Actual pistol shot
2. Replica	Fake or toy example	Fake or toy pistol	Noise of cap gun
3. Somatic	Body part or bodily action	Fist with extended index finger	Hand clap, “bang”
4. Ad hoc	Arbitrary object, event, or state	Pen extended from the hand	Ruler rapped against a table
5. Figural	Exhibited depiction used as back prop	Sketch of a pistol	Recording of pistol shot

lock worked. Two of their depictions are shown in Table 16. When the tutors had an actual lock and key in hand, they tended to use them as genuine props, as in *Cotter-pins*. When they did not, they often created somatic props, as in *Grooves*, in which the tutor used his knuckles as a prop for a key’s grooves. In *Protractor*, a math instructor used a pen and protractor as ad hoc props for a man and a woman (Radford, 2009, p. 117).

In face-to-face dialog, somatic props are often the only option available. Some of these are novel, like the knuckles in *Grooves*, but others are more or less conventional, such as these:

- a. A hand-phone made with an extended thumb and little finger;
- b. Hand scissors fashioned from two fingers opening and closing;
- c. A hand pistol made from a vertical fist and extended index finger;
- d. A hammer or gavel made from a fist plus forearm; and
- e. A person walking formed from two fingers pointing down and wiggling on a surface.

Other somatic props are auditory, such as hissing to depict a snake and a hand clap to depict a door slam. Still others are tactile, such as drawing a finger across a friend’s arm to depict a spider.

Make-believe play relies heavily on props that are replicas (Bretherton, 1984; Garvey, 1990; Harris, 2000). Many of these are toys specifically manufactured to be used as center props—dolls, doll houses, train sets, toy cars, and toy animals. And many of these are replicas not only in shape, but in sound (e.g., dolls that cry, toy cows that moo) or in touch (e.g., soft animals, rubbery worms).

Children also use center props that are ad hoc. In one example (Harris, 2000), a 2-year-old put a teddy bear into a bathtub (a shoebox), rubbed its back with soap (a wooden block), and then,

because “He’s all wet,” wrapped it in a towel (a piece of paper). As a rule, children find it easier to pretend with replica than with ad hoc props (Fein, 1975; Mann, 1984).

Side Props

People enacting characters often use physical props as ancillary support—as side props. In *Godot*, McKellen used a genuine boot in enacting Estragon trying to take off his boot, and in Shakespeare, actors use fake daggers, fake swords, and fake poison. In everyday depictions, side props range from genuine to ad hoc.

A great many side props are somatic, as illustrated in Table 17. In *Hand-phone* and *Falcons* (repeated from Table 1), the producers J and WK used a conventional hand-phone and hand-pistol. In *Arm-viol*a, a music tutor created a novel prop, using his forearm as the neck of a viola and, at the same time, producing spoken syllables for its sounds in a hybrid depiction.

In make-believe play, children’s side props are often replica props. In *Go-to-bed* (see Table 7), Sally enacted Mother putting Baby to bed by using a doll and a toy bed as side props. Many toys are designed to be side props—Halloween costumes, water pistols, pedal driven cars, and hobby horses.

Genuine props are especially common in full-scale demonstrations. In *Bartok*, MB demonstrated on an actual piano. In other examples, music tutors performed demonstrations with actual violins, oboes, and marimbas; tennis coaches with actual rackets and balls; chefs with actual bowls, spoons, flour, eggs, water, and frying pans; and carpenters with actual wood, saws, hammers, screws, and screwdrivers.

Back Props

Back props are found in both actor and prop depictions, and they, too, range from genuine to ad hoc. In demonstrating how a lock works (Engle, 2000), one tutor used her finger to depict a key that she inserted into an actual lock. Her finger was a side prop, and the actual lock was a back prop.

Table 16
Prop Depictions With Physical Props

Label	Example
<i>Cotter pins</i>	You can probably hear it. (Slowly moves an actual key into the keyhole; three clear separate clicks can be heard). (Finishes sliding the key in with a fourth click) <u>You can</u> hear each pin go up?
<i>Grooves</i>	And these (lifting up L fist with largest knuckles on top) <u>grooves are</u> specially designed. . . .
<i>Protractor</i>	Okay, look (4.36 s) Okay, hey, I got it. She . . . That is him, this is her. (The speaker uses a pen and a protractor to simulate his and her motions on the desk) <u>He walks more quickly, so it does this</u> (the distance between the pen and the protractor increases), and there is a bigger distance.

Table 17
Actor Depictions With Physical Props

Label	Example
<i>Hand-phone</i>	[J placed a hand-phone to her ear whenever she mentioned calling someone on the telephone.]
<i>Falcons</i>	[Film director WK was speaking to reporter TF] “We were going to shoot on top of the MetLife Building . . . but falcons had nested there, so they shut it down. In L.A., they would have—” He leveled a finger at some imaginary nestlings and made a gun-cocking sound. (<i>New Yorker</i>)
<i>Arm-violin</i>	[Violist DM was telling a student about fingering on a violin.] We have to get this cleaner. [<i>While sitting, DM raises his R forearm and places LH fingers near his R wrist, simulating the neck of a violin.</i>] (<i>singing and fingering</i>) <u>Bum duh duh duh . . .</u> (for 5.77 s, then looking up at student and in slow deliberate phrases) <u>Buh-duh duh dum, . . .</u> (for 3.84 s).

Many back props are figural props such as maps, photographs, and diagrams. In one study (Clark, unpublished), one student was instructed to tell another how to get from one place on campus to another. When the two of them shared a map, the first described the route while tracing it on the map with her finger, as here:

So what you’re gonna do is you’re gonna (*tracing to a stop with R index finger*) *come up* (1.03 sec) (*tracing to a stop*) *on, uh, Fremont.* [Uh-huh.] (*tracing to a stop*) *Electioneer.* [Uh-huh] (*tracing to a stop*) *Up Campus Drive. . . .*

The student treated the map as a back prop and depicted the route on top of it. In other examples, people traced mechanical sequences on diagrams (Hegarty, Mayer, Kriz, & Keehner, 2005; Tversky & Suwa, 2009), and changes to architectural plans on blueprints (Heath & Healey, 2012; Murphy, 2004, 2005).

Imaginary Props

Imaginary props are invisible, inaudible, impalpable props that producers bring into existence in the very performance of their depictions. In *Godot*, McKellen used a physical hat in enacting Estragon doing routines with his hat. In a silent mime version, he would have used an imaginary hat. Everyday depictions are no different. In *Bartok*, the piano MB used was visible and audible to the audience, but in *Bink-bink*, there was no piano until he created one—an imaginary one—by acting as if it were visible and audible to the audience.

Imaginary props are found in all types of depictions, but I will focus on iconic gestures, where they have been studied in remarkable detail (e.g., Cienki & Müller, 2008; Enfield, 2009; Kendon, 2004; McNeill, 1992; Streeck, 2008; Streeck, Groethues, & Villanueva, 2009, among many others).

Center Props

When people are outside the proximal scene, they create imaginary props according to an *as-if-evidence principle*: “From out-

side the proximal scene, behave *as if* there is evidence for the physical props inside the scene.” Most of this evidence takes one of four forms: shape, location, movement, and sound.

Shape. Producers can *trace* an imaginary prop’s shape or exterior surface, as in *Tapering*. They can *delimit* part of a prop’s exterior surface, as in the gesture for “I caught a fish this long.” Or they can pretend to handle parts of its exterior surface.

Location. Producers can establish an imaginary prop’s location. In an example from Kendon (2004, p. 148), “M acts as if to sketch a large rectangular object lying on the table in front of him, understood as a depiction of the Christmas cake he refers to verbally.” M locates the cake by tracing its shape at a location in the larger scene.

Motion. Producers can imply an imaginary prop by depicting its motion. In *Sparks* in Table 18, the man interviewed by a TV reporter accompanied his description of what happened with vivid gestures. He used the second and third gestures to depict the motion of objects emitting sparks and exploding.

Sound. Producers can also imply an imaginary prop by depicting its sound. In *Dee-duh-dum*, MB created a piano and its sounds by singing “dee-duh dum.” In *Worrying-sound* and *Purring*, the callers implied the presence of a car engine by depicting its sounds. And in *Fire* in Table 18, the man implied the presence of the fallen electrical pole by depicting its audible aftereffects.

Side Props

When people are inside a proximal scene, they create imaginary props, instead, according to an *enacting-as-if principle*: “Perform your enactments as if the *physical props* were present.” The props created by this principle look distinctly different from those created by the *as-if-evidence principle*.

The true experts in enacting-as-if are silent mimes. To study their techniques, I examined video-recordings of sketches by the mimes Marcel Marceau (MM) and Red Skelton (RS). Segments of two sketches are described in Table 19. As actors in these sketches, MM and RS never once relied on the *as-if-evidence principle*.

Table 18
Imaginary Props About a Car Struck by a Falling Electrical Pole

Label	Example
<i>Sparks</i>	There was (<i>thumb pointing off to one side</i>) <u>arcin’</u> , (<i>two hands going up</i>) <u>sparkin’</u> , (<i>moving the two hands upward and outward in front of him</i>) <u>blowing up</u> .
<i>Fire</i>	And then (<i>pronouncing each word with an explosive phonation and with wild head and hand gestures</i>) “ <u>arc arc bam</u> ” fire was coming from everywhere.

Table 19
Imaginary Props Created by Two Professional Mimes

Label	Description of sketch
<i>Smoker</i>	Marcel Marceau (MM) mimed a man picking up a cigarette, striking a match, lighting the cigarette, tossing the match, and puffing on the cigarette.
<i>Parade</i>	Red Skelton (RS) mimed a person watching a parade go by. He implied the presence of a marching band, a troop of soldiers, and a horse by the direction of his gaze and many other reactions.

Their use of enacting-as-if fell mainly into three classes: acting on, reacting to, and communicating with.

Acting on. In *Smoker*, MM created an imaginary cigarette by picking it up, lighting it, and puffing on it. In another sketch, RS created imaginary swim goggles, waders, suspenders, and swim fins by putting them on.

Reacting to. In other sketches, both MM and RS moved their heads and eyes as they followed imaginary birds, airplanes, lions, bulls, horses, and passers-by. In *Parade*, RS created a marching band in a parade by watching it pass by while marching in place.

Communicating with. In other sketches, MM and RS “spoke” to, “yelled” at, gestured to, gestured for, saluted, and held hands out for imaginary people.

The same techniques should occur in everyday discourse, and they do. In *Bink-bink*, MB created an imaginary piano by acting on it—thrusting his right index finger onto its imaginary keys. In *Falcons*, WG depicted the location of the imaginary falcons by leveling his finger in their direction. And in *Booming*, MB depicted an imaginary student by addressing her as if she were there.

To sum up, imaginary props are created in both prop and actor depictions. Center props are brought into existence by prop-managers displaying *as-if* evidence of their presence. Side props are brought into existence by actors performing enactments *as if* the imaginary props were physically present.

Spatial Frames of Reference

In the theater, people keep track of several frames of reference (FRs) at once. They have to. At any moment in a play, there are three scenes, a base, a proximal, and a distal scene, and each has its own FR—its own spatial coordinates, or *place-frame*, and its own temporal coordinates, or *time-frame*. People represent the elements in each scene with respect to these frames. In this section, I focus on place-frames.

Typically, it is easy to map the place-frames from one scene to the next. If McKellen is facing Stewart in the base scene, then the actor playing Estragon is facing the actor playing Vladimir in the proximal scene, and Estragon is facing Vladimir in the distal scene. Where, however, is the audience in all this?

The base scene is itself part of a larger *discourse scene*. In *Godot*, the discourse scene included not only the stage (for the actors and props), but the backstage (for the crew) and the auditorium (for the audience):

Discourse scene \supset base scene \rightarrow proximal scene \Rightarrow distal scene

Base scenes have very different configurations in the discourse scene as one goes from standard theaters to theaters-in-the-round,

Punch-and-Judy shows, street theater, immersive theater, radio plays, and other venues.

Spatialization

In everyday circumstances, people have to configure each depiction within the discourse scene, and they do that by a process I will call *spatialization*. To spatialize a depiction is to establish the dimensionality, scale, location, orientation, and granularity of the base scene within the discourse scene (see Table 20).

The idea behind spatialization is easy to illustrate. At the opening of *Godot*, a lone cast member (McKellen) was sitting on a prop mound with a prop tree behind him. This was the base scene. That scene, however, was oriented in such a way that McKellen was downstage facing the audience and the tree was upstage. If the scene had been rotated 180 degrees, McKellen would have been upstage facing away from the audience with the tree downstage. The base scene would have been the same, but its spatialization would have been different.

People spatialize depictions largely to suit recipients. In *Godot*, the base scene (McKellen sitting in front of a prop tree) was designed to reflect the distal scene (Estragon sitting in front of a tree). Its orientation (with McKellen downstage and the tree upstage) was chosen to be optimal for the audience to see and hear. Everyday depictions are spatialized for the same reasons.

Orientation

In spatializing a depiction, people have a choice of orientation. Consider DN’s actions in *Schubert* in Table 21. In Line 3, DN enacted Enrico singing “Fremd bin ich ein-” to an imaginary audience. In doing so, he *could* have continued to gaze at MB, but if he had, MB might have thought he was singing to *him*. To prevent this, DN broke his gaze with MB and reoriented his body, head, and gaze 60 degrees to his left (Line 2), and only then did he gesture and sing the four words (Line 3). Even before he had finished singing, however, he had returned his body, head, and gaze to MB (Line 3). These are *orientation shifts*. DN used them

Table 20
Five Parameters in the Spatialization of Depictions

Parameter	Example values of parameter
1. Dimensionality	0D, 1D, 2D, 3D
2. Scale	Miniature, life-sized, enlarged
3. Location	In front of producer, surrounding producer
4. Orientation	Rotate left 60 degrees, horizontal-to-vertical
5. Granularity	Full detail, much detail, skeletal

Table 21
Detail of DN Performing Two Depictions in Schubert

1. [DN is facing and gazing at MB.]
2. He (*breaking eye contact and reorienting body, head, and gaze 60 degrees left*) found he could do his
3. (*thrusting left arm toward an imaginary audience*) Fremd bin ich (*reorienting body, head, and gaze back to MB*) ein-
4. in a much more natural way
5. (*reorienting body, head, and gaze 60 degrees left*) because he didn't have to
6. (*thrusting both arms toward imaginary audience*) pitch himself
7. and project himself
8. (*reorienting body, head, and gaze back to MB*) over a lot of noise.

to make clear the proximal scene did *not* include MB (see Sidnell, 2006).

Producers use the same orientation shifts in follow-up depictions. In Line 6 of Table 21, DN went on to depict Enrico “pitching himself” over a lot of noise. To mark that scene as having the same place-frame as the first, DN used the same orientation shifts, turning 60 degrees left before the enactment (Line 5) and turning back after it (Line 8).

Producers have a choice in which segments of their body to reorient, as in these examples:

- a. Head, eye gaze, body, and gestures (*Booming, Schubert, Bink-bink, Smoker, and Parade*)
- b. Head, eye gaze, and gestures (*Lift-your-chin*)
- c. Head and eye gaze (*Thinking-face, Funny-face*)
- d. Gestures alone (*Raised-voice, Come-out*)
- e. Eye gaze alone (*Audition, MTV* in Table 22)

In sign languages, quotations are marked by reorienting the entire body in what are called role shifts (Emmorey, 2002; Engberg-Pedersen, 1993).

Orientation shifts go with the characters enacted. In *MTV* (in Table 22), Lauren depicted alternating camera-like shots of Lo, herself, and Lo by means of slight orientation shifts of her eye gaze left, right, and then left again (along with shifts in her voice). We see and hear the sequence not as three unrelated actor depictions, but as a single conversation (see also Goodwin, 1990). Orientation shifts in eye gaze are especially common in oral narratives (Sidnell, 2006; Stec, Huiskes, & Redeker, 2015; Thompson & Suzuki, 2014).

Placement

People spatializing a depiction also have a choice of placement. People usually perform hand gestures in an imaginary 3D bubble in front of the body called a *gesture space* (McNeill, 1992). In McNeill's studies, 95% of the iconic gestures were performed there. People sometimes create imaginary stages within this space, about chest high, on which they perform proximal scenes, much as signers do in sign languages (Emmorey, 2002; Liddell, 1995, 2003). For people who are face to face, this is the space where the gestures are most visible (Gullberg & Holmqvist, 1999).

By the same logic, producers should place gestures in other locations if that would be better for recipients. In *Tapering*, MB was speaking to about 50 people, so he placed his gesture above his shoulders and left of his head where it was highly visible to everyone. And in a study by Özyürek (2002; see also Campisi & Özyürek, 2013), people who narrated stories to people on their left or their right accommodated by shifting their gesture spaces to their left or their right.

Producers also use placement in marking follow-up depictions. In an example analyzed by Enfield (2009, p. 119), a Lao fisherman described a fish-trap called a *sòdn5* (translated from Lao):

1. As for the *sòdn5*, they make it (*using both hands to form a funnel facing upward*) fluted at the mouth
2. and they (*holding LH in place while using RH to depict smallness of the opening*) make it small
3. (*holding LH in place but now using RH to depict a fish swimming into the opening*) going in.

In line 1, the fisherman depicted the mouth of the trap as a funnel facing upward (a somatic prop). In line 2, he kept his left hand in place to maintain half of the funnel, and depicted the small opening at the bottom of the funnel and then, in line 3, a fish swimming into that opening. He marked the continuity of these depictions by holding his left hand in place and depicting the opening and the fish with respect to it. One of Enfield's depictions went through 11 steps (for similar examples, see Kendon, 2004, e.g., pp. 161–163).

And, finally, people use placement in creating private depictions—depictions for their own use. In an unreported experiment by Teenie Matlock and myself, people were given a map and asked to locate routes and landmarks for partners who had their own maps. The instructors tended to gesture over the maps in a private gesture space before performing similar gestures in a public space for their partners. They intended only the public gestures for communication.

Scale and Granularity

To spatialize a depiction is also to select its scale. In the theater, one reflection of scaling is in stage voices and stage gestures. When McKellen produced “Nothing to be done,” he directed his speech at Stewart a few steps away (in a base place-frame), but in a voice loud enough to be heard in the balcony (a discourse place-frame). And when he gestured to Stewart, it was with move-

Table 22
Depiction of Dialogue About MTV

[Lauren was recounting a conversation with her friend Lo about producers from MTV coming to their school.]

1. I I don't know. I had to go to the bathroom and went out there,
2. and she's all, “(*shifting gaze slightly left, in excited voice*) You'll never guess who's in the office.”
3. and I was like, “(*shifting gaze slightly down, in second excited voice*) Who?”
4. and she's like, “(*shifting gaze slightly left, in first excited voice*) MTV.”

Note. From *Laguna Beach* (MTV).

ments large enough to be seen in the balcony. The audience would be mistaken if they thought it was *Estragon* who was exaggerating his speech or gestures.

Producers follow the same logic in everyday depictions (Bavelas, Gerwing, Sutton, Prevost, 2008; Campisi & Özyürek, 2013; Debreslioska, Özyürek, Gullberg, & Perniss, 2013; Gerwing & Bavelas, 2004; Holler & Wilkin, 2009). In *Tapering*, MB turned a 3 cm long diminuendo sign into a 30 cm long gesture to make it visible to his audience. And in *Fish-trap*, the fisherman reduced the size of the fish-trap to fit his gesture space. People also use smaller gestures when they are on the telephone (Bavelas, Gerwing, Sutton, & Prevost, 2008).

Spatializing a depiction also includes granularity. In one study (Gerwing & Bavelas, 2004), people were asked to tell others how a toy whirligig worked. Their initial depictions were large, detailed, and precise because the information about the toy was new. Their later depictions were smaller, less detailed, and less precise because, by then, the whirligig was already known and taken as given (see also Holler & Wilkin, 2009).

Dimensionality

Depictions can be 0D, 1D, 2D, or 3D. MB's depiction in *Dee-duh-dum* was effectively 0D, because it was produced by his vocal apparatus at a single point. Someone saying "I caught a fish this long" would use a 1D gesture of length. MB's depiction of the diminuendo sign in *Tapering* was a 2D drawing. Most of the other depictions were 3D.

How people spatialize a depiction depends on its dimensionality. For a 0D depiction, producers must select a location, but not a size or orientation. For 1D, 2D, and 3D depictions, they must fix all three parameters. Producers sometimes have a choice of dimensionality. In *Come-out*, Lauren quoted her friend Lo using "come-here" gestures in a 3D depiction. If all she had enacted was Lo's speech, it could have been a 0D depiction.

Viewpoints

To interpret what people are doing generally requires taking account of their viewpoints. By viewpoint I mean the position of a person in a scene—the location and direction from which he or she sees, hears, feels, and manipulates things. In face-to-face conversation, speakers and addressees tend to display their roles in part by the way they position their bodies, shoulders, heads, eye gaze, and actions in relation to each other (Goodwin, 1981; Kendon, 1990; Schegloff, 1998). These are discourse viewpoints.

Depictions require additional viewpoints. In *Schubert*, there were only the speaker's and the addressee's viewpoints as long as DN and MB were discussing pianos. Though once DN began singing, "Fremd bin ich ein—" he introduced three more. With his 60 degree orientation shift, he created distinct viewpoints (a) for the actor playing Enrico, (b) for the imaginary audience he was facing, and (c) for MB in the role of spectator, viewing his enactment from the side. These were proximal viewpoints.

In distal scenes, every character has a distinct viewpoint. In *Schubert*, DN and MB imagined a recital hall in which Enrico was facing an actual audience with a covert observer looking on from the side. The viewpoints of Enrico, his audience, and the covert observer were distal viewpoints.

Are all of these viewpoints really needed? In general, yes. Discourse viewpoints are needed for expressions such as *I*, *you*, *over there*, and *the other side*. Proximal viewpoints are needed for interpreting the spatialization of a depiction; with the orientation shift, DN separated the actor's, audience's, and spectator's viewpoints from the speaker's and addressee's. Distal viewpoints are needed for identifying the *I*, *here*, and *now* of the characters depicted. When MB uttered "Well, I can see you Wednesday at three o'clock" in *Booming*, the referents of *I* and *you* were *not* MB and his audience, but Smith and the student in the distal scene.

Temporal Frames of Reference

Every scene also has temporal coordinates, or a time-frame. Early in *Godot*, McKellen's "Nothing to be done" was followed by Stewart's "I'm beginning to come round to that opinion." We in the audience represented the timing of these two events with respect to the base scene. We also represented the timing of the base scene as a unit with respect to the time-frame of the discourse scene—the play as performed in London that evening.

Internal Timing

The events in a base scene typically map moment by moment into the distal scene. In *Dee-duh dum*, the tempo and rhythm of MB's three syllables mapped into the tempo and rhythm of Mozart's three notes. And yet base timings can be modified when there is reason to do so. Here are several techniques:

1. *Holds*: A producer can pause, or hold, a depiction at particular junctures. In demonstrating a tennis serve, a coach might add a hold at the beginning, the top, and the end of the swing.
2. *Slow-motion*: A producer can perform an entire depiction at half speed.
3. *Condensation*: A producer can depict an entire scene in a condensed form.
4. *Truncation*: A producer can truncate a depiction in the middle.

Holds and slow-motion are common in instructional demonstrations, as in this example:

[Pianist NT is asking a student about the phrasing of three bars from Brahms' first piano concerto.] Are you going (playing piano slowly and singing softly) "(D# E E) (rest) (C# C# D D) (A A B B C# C#)"? You could also do (playing and singing softly) "(D# E E C#) (rest) (C# D D A) (rest) (A B B C# C#)."

NT played the three bars in slow motion, first in one phrasing and then in another. She added the hold (a "rest") to mark the difference between them. Condensation is illustrated by the hand gestures called "markers" that ballet dancers use in discussing and rehearsing dance routines (Kirsh, 2010, 2011). Each marker is a compressed depiction of a specific ballet movement. And truncation is illustrated by "Fremd bin ich ein—" in *Schubert*.

Phases of Discourse

For most depictions, the discourse scene divides into three phases: preparation, execution, and follow-up. In *Go-to-bed* (see Table 7), Sally and Alex prepared their scene by setting up props (e.g., the doll and doll's bed) and issuing stage directions (e.g., Sally's "Say, 'Go to bed'"). They distinguished these from their enactments, such as Alex's "Go to bed" and Sally's "(whining) Why?" And in a follow-up phase, they coordinated in closing down their play and leaving.

In everyday discourse, the three phases are usually local to the utterance produced. In *Schubert* (see Table 21), the first clause of DN's utterance divided as follows:

1. *Preparation phase: (DN and MB are facing each other two meters apart. DN is speaking.) He (turning his body, head, and gaze 60 degrees left) found he could do his*
2. *Execution phase: (gesturing and singing to an imaginary audience) Fremd bin ich ein-*
3. *Follow-up phase: (at ein-, returning his body, head, and gaze to BM) in a much more natural way . . .*

DN produced the entire clause as part of DN and MB's discussion of pianos. Within that clause, however, he timed the depiction locally to fit the *N* slot in "He found he could do his *N* in a much more natural way."

Discourse Timing

For local timing, a depiction can be: (a) *slotted* within or between utterances; (b) *overlapped* with an utterance part; or (c) *fused* with an utterance part. Depictions, of course, cannot overlap descriptions produced with the same body part, or be fused with ones produced with a different body part. Otherwise, the three techniques are used for different purposes.

Embedding. The function of an embedded depiction is specified by the slot in which it is embedded. In *Schubert*, the depiction filled the slot for the nominal *N* in "his *N*." Other depictions cited in this article occupied slots for noun phrases, verbs, adjectives, and adverbs. The same logic applies to independent depictions. The student in *Manifest* placed her depiction in the slot for an answer to a question.

One way to embed depictions is with quotatives such as "he said," "they go," "she's like," or "Mozart wrote." In *MTV*, Lauren, speaking about her friend Lo, said: "and *she's all* 'You'll never guess who's in the office.'" Her quotative ("she's all") helped establish not one, but two functional mappings:

Lauren as base producer → actor playing *Lo*

Lauren executing "You'll never guess who's in the office" → enactment of *Lo* saying "You'll never guess who's in the office"

Indexing. The function of an indexed depiction is specified by the phrase that indexes it. In *Mozart*, MB's demonstration on the piano was indexed by "this," the subject noun phrase of "NP is what I want to write." In his lecture, MB placed his indexed depictions in various locations: (a) in a slot just before the current utterance; (b) in overlap with the indexing phrase, as in *Mozart*; (c) fused with the

indexing phrase, as in *Raised voice*; and (d) in a slot just after the current utterance. That is, MB placed indexed depictions near, but not necessarily cotemporal with, the phrases that indexed them.

Adjoining. The function of an adjunct depiction is specified by the word or phrase it is timed to go with. Table 23 shows two examples. In *Window*, as analyzed by Kendon (2004, p. 117), a woman was about to say "window displays" when she started to produce a gesture to go with it. As Kendon argued, however, the woman had not planned the gesture by the end of "wonderful," so she delayed "window displays" by 0.6 s so that she could perform the "stroke" of her gesture in overlap with it. In other examples, speakers delayed the gesture to overlap the yet-to-be-formulated affiliate (Kendon, 2004, pp. 134ff; see also Chu & Hagoort, 2014). A gesture should overlap its affiliate, and speakers can achieve that by delaying either the gesture or the affiliate.

In *Running*, a woman depicted the running speed of the gunmen she was describing by accelerating her speech within the words *take off* and *running*. That is, she fused her depictions with these two words to mark them as the affiliates (for other examples of fusion, see Perlman, Clark, & Johansson Falck, 2015; Shintel, Nusbaum, & Okrent, 2006).

Producers of a depiction, in brief, pay close attention not only to its spatialization, but to its timing. They choose the timing within a depiction for strategic reasons and use the timing of the depiction as a unit to specify its function in the discourse.

Displaced Worlds

In discourse, people engage in the *actual world*, but during each depiction, they also engage in a *displaced world*—a world displaced in location, time, or reality. In *Schubert*, the actual world contained the discourse scene (DN and MB discussing pianos), and within the depiction, the displaced world contained the distal scene (Enrico singing "Fremd bin ich ein-"). Both worlds were larger than these scenes alone. Enrico did not stop singing after just four words, and he surely had a life both before and after his singing. Displaced worlds, like actual worlds, are assumed to be continuous and complete.

During a depiction, then, people jointly engage in two simultaneous layers of activity (Clark, 1996):

Layer 1: Engaging in the actual world (with its discourse scene)

Layer 2: Imagining the displaced world (with its distal scene)

Table 23
Examples of Timing in Two Adjunct Depictions

Label	Example
<i>Window</i>	And they used to have wonderful (0.6 s silence, then outlining a window with two hands) <u>window displays</u> in there, didn't they.
<i>Running</i>	[Woman was describing some gunmen she witnessed to a TV reporter.] I come out here to the steps [pause] [speeded pronunciation, hand moving quickly to her right] <u>take off</u> down that way . . . and then I see some people [speeded pronunciation, hand moving quickly to her left] <u>running</u> that way."

The participants in a discourse are always engaged in Layer 1. It is only during a depiction that they engage in Layer 2 as well. How they engage in Layer 2 depends on the displaced world.

Real and Fictional Worlds

In literature and film, works are either nonfiction or fiction. Histories contrast with novels, news stories with short stories, and documentaries with movie dramas. The two differ in their displaced worlds. Fictional worlds include elements that are not real. Nonfictional worlds reputedly do not.

Nonfiction also contrasts with fiction in the theater. In Shakespeare, the histories contrast with the comedies and tragedies, and in solo performances, actors doing Mark Twain or Abraham Lincoln contrast with satirists and mimes. If so, nonfiction should also contrast with fiction in everyday discourse, and it does. In *Schubert*, the world of Enrico was real. In *Go-to-bed*, the world of Mother, Father, and Baby was fictional.

The difference is crucial. With nonfictional depictions, the recipient's goal is to fit the distal scene to the real world. The direction of fit is depiction-to-world. With fictional depictions, in contrast, the goal is to create a world, different from the real world, that fits the depiction. The direction of fit is world-to-depiction. Those of us at *Godot* created the world of Estragon and Vladimir as the play progressed. That world did not exist for us until it was created by the play. Still, evidence suggests that people assume that a displaced world is real unless it is marked otherwise.

Joint Access

The basic function of a nonfictional depiction is to establish *joint access* to a segment of the displaced world. In *Schubert*, DN drew MB's attention to Enrico singing just four words ("Fremd bin ich ein")—a brief segment of Enrico's world. Other depictions access segments of past or future actions, events, and objects. Examples range from the quotation of Smith in *Booming*, the musical depiction in *Dee-duh-dum*, and the diminuendo sign in *Tapering* to the car noises in *Worry-noise* and *Purring* and the notches in the key in *Grooves*. Each of the worlds accessed is displaced in place or time, but not in reality.

The item depicted may be either a *particular* individual or event, or an *example* of a type of individual or event. The quotation in *Tiffany's* represented what a particular speaker said on a particular occasion, but the quotation in *Zchau* was an example of the type of question Zchau was asked. The diminuendo sign in *Tapering* was also an example of a type, and so were the two measures depicted in *Mozart*. In full-scale demonstrations, tennis instructors are more

likely to depict a type of serve than, say, Serena William's last serve at Wimbledon in 2015.

Joint Pretense

Fictional depictions, in contrast, are designed to create joint pretenses. Children are explicit about this in make-believe play. In *Go-to-bed*, Sally and Alex would have said that they were just pretending (Garvey, 1990; Harris, 2000). As one two-and-a-half-year-old said, "I'm not realing. I'm pretending." Joke tellers are usually just as explicit, as in this joke Sam told to Reynard (Svartvik & Quirk, 1980):

- 1. Let me tell you a story, - - -
- 2. a girl went into a chemist's shop and asked for . contraceptive tablets - - -
- 3. so he said "well I've got . all kinds and . all prices what do you want"
- 4. she said "well what have you got"
- 5. [*proceeds with long joke*]

In Line 1, Sam announced he was going to tell Reynard a "story," and from Line 2 through the punch line, he enacted someone telling someone else in a fictional world about a girl who went into a chemist's shop. The depicted worlds in both make-believe play and jokes are displaced in place, time, *and* reality.

In another class of depictions, however, the depicted worlds are displaced in reality, but *not* in time or place. One example is the tease recorded in Table 24. Conan O'Brien had just welcomed Chloe Sevigny as a guest on his late night TV show when, in Line 3, he *appeared* to assert that Will Ferrell, who was sitting next to Sevigny, "almost wore that outfit" but "decided that would not be cool." He did not announce the pretense beforehand, but left Sevigny and his audience to discover it on their own. He enacted himself speaking to Sevigny at that moment, but in a fictional world in which Farrell "almost wore that outfit." Ferrell recognized the pretense and added "Inches away," and Sevigny showed her recognition by laughing (see Drew, 1987).

Certain pretenses, then, are designed to be discovered, and at least some are interpreted first as real (in Layer 1) and then reinterpreted as fictional (in Layer 2). Teasing belongs to a class of staged communicative acts that also includes verbal irony, hyperbole, understatement, rhetorical questions, banter, taunting, and ostensible speech acts (Clark, 1996, Chapter 12; Clark & Gerrig,

Table 24
A Joint Tease by O'Brien and Ferrell of Sevigny

[Conan O'Brien has just welcomed Chloe Sevigny, who is sitting next to Will Ferrell, to his late-night TV program.]	
1. O'Brien	You look (<i>LH gesturing toward Sevigny</i>) <u>beautiful</u> .
2. Sevigny	<i>Thank you.</i>
3. O'Brien	Will almost (<i>two hands gesturing at dress</i>) <u>wore that outfit</u> when he came here tonight.
4. Ferrell	(<i>leaning forward with thumb and finger inches apart</i>) <u>Inches away</u> .
5. Sevigny	[<i>laughs</i>]
6. O'Brien	Then he decided that would not be cool.

1984; Gibbs, 2000; Isaacs & Clark, 1990; see also Clift, 1999). As Grice (1978) wrote, “To be ironical is, among other things, to pretend (as the etymology suggests), and while one wants the pretense to be recognized as such, to announce it as a pretense would spoil the effect” (p. 125).

Recursion

Layering is also needed to account for recursive depictions (see Clark, 1996). In Act 5 of *A Midsummer Night’s Dream*, Quince and five other characters put on a play called *Pyramus and Thisby*. The audience at *Midsummer Night’s Dream* would readily recognize: (a) that they themselves belonged to the actual world, (b) that Quince and his friends belonged to the *Dream*-world, and (c) that Pyramus, Thisby, and *their* friends belonged to the *Pyramus-and-Thisby* world. They would know they were seeing a fictional depiction of a fictional depiction.

If there is recursion in the theater, there could also be recursion in everyday depictions, and there is. Recall that in *Bink-bink*, MB told a woman:

All you have to do is say, “Look, there are four notes here, but it’s not really very natural to go ‘(thrusting *R* index finger four times in rhythm onto an imaginary piano keyboard) bink bink bink bink.””

Starting with “look,” MB enacted the woman speaking, but starting with “bink,” he enacted the woman enacting a pianist playing four notes—a 2P actor depiction of a 2P actor depiction. And Sam’s joke became recursive when the fictional narrator quoted the chemist and the girl. Recipients had no trouble interpreting these recursions.

Conclusions

Depicting, I have argued, is a basic method of communication. It is on a par with describing and indicating, but it works by different principles. To describe something is to *tell* others about its properties—to represent it categorically. To indicate something (e.g., by pointing) is to *locate* it for others in space and time. To depict something, however, is to *show* others what it looks or sounds or feels like.

Depictions contrast most directly with descriptions. To highlight their differences, let me return to *Falcons*, TF’s report of WG telling him about shooting some falcons:

Version 1: “In LA, they would have *shot those falcons*.”

Version 2: “In LA, they would have—” He leveled a finger at some imaginary nestlings and made a gun-cocking sound.

Version 1 is what WG might have said. Version 2 is what he did say, with its embedded depiction. The question is this: How does the depiction in Version 2 differ from the description “shot those falcons” in Version 1?

Scenes

Pure descriptions are structured strings of arbitrary symbols, whereas depictions are physical scenes. In Version 1, WG produced a stream of sounds that was construed as the phrase *shot those falcons*. He used these words in turn to denote a composition

of the categories “kill or wound with a bullet,” “distant in space,” and “diurnal bird of prey” for elements in the scene he was describing. In Version 2, in contrast, WG staged a scene by “leveling a finger at some imaginary nestlings and making a gun-cocking sound.” He intended the scene as a physical analog of the scene he was trying to depict. In depictions, space maps into space, time into time, actors into characters, and props into objects. In descriptions like “shot those falcons” there are no spaces, actors, or props to map.

Descriptions are formed by linearization, and depictions by spatialization. In Version 1 of *Falcons*, once WG settled on what he wanted to describe, he selected, ordered, and timed the words *shot*, *those*, and *falcons* according to the conventions of English. This is a process of linearization (Levelt, 1989). In Version 2, once WG settled on which elements to depict, he spatialized them in the discourse by selecting a placement, orientation, scale, granularity, and dimensionality for them. Spatializing a depiction and linearizing arbitrary symbols are fundamentally different processes.

Layers of Action

Descriptions are created in one layer of action, and depictions in two. As a result, descriptions have two active participant roles, and depictions have four.

In the primary layer of a discourse, every utterance has a speaker and at least one addressee. The speaker, equivalent to Goffman’s (1981) *principal* and Grice’s (1991) *utterer*, is the participant who means something by producing the current utterance. In both versions of *Falcons*, WG was the speaker, TG was the addressee, and WG’s goal was to jointly establish with TG what people in LA would have done to the falcons.

In a depiction, however, people engage in a second layer of activity as well, and there they take on additional roles—and viewpoints. In the depiction in Version 2, WG was also temporarily an actor (playing someone shooting falcons), and MB a spectator (playing someone covertly watching the shooting). In this layer, WG’s goal was for the two of them to jointly imagine a displaced world in which someone was shooting falcons. In prop depictions, the added roles are prop-manager and spectator (as in *Tapering*), and in joint depictions, they are actor and coactor (as in *Manager*). The roles of speaker and addressee originate in the discourse; the roles of actor, prop-manager, spectator, and coactor originate in the staging.

Imagined Worlds

Ordinarily people try to imagine a displaced world whether it is described or depicted. In Version 2 of *Falcons*, TF would have used WG’s enactment to imagine someone shooting falcons. In Version 1, he might have used “shot those falcons” to imagine a similar scene (see Clark & Van Der Wege, 2015; Gerrig, 1993; Zwaan, 2009, 2014). These are really two types of imagining.

Imagining a depicted scene is a *percept*-bound process. When WG “leveled a finger at some imaginary nestlings and made a gun-cocking sound,” the scene TF imagined was bound by the depiction as he perceived it. It would have included: a rifle visually analogous to WG’s leveled finger; a gun-cocking sound auditorily analogous to WG’s tongue click; and a shooter aiming across at the falcons.

Imagining a scene from a description is, instead, a *concept-bound* process. For “shot those falcons,” the scene would have had to fit a composition of the categories “kill or wound with a bullet,” “distant in space,” and “diurnal bird of prey.” TF could not know whether the shot came from a pistol or rifle, or from above or below, so he would have filled in these details by guesswork. If the verb had been “done something,” it would all have been guesswork. Imagining distal scenes is an inherent part of interpreting depictions. It is *not* an inherent part of interpreting descriptions. Many descriptions, indeed, have no scenes to imagine.

Composite Utterances

Everyday discourse is a mix of descriptions, depictions, and composites of the two. It is one thing to account for descriptions and depictions on their own and quite another to account for composites.

One issue is function. Embedded depictions are used *as if* they were nouns, verbs, noun phrases, or other linguistic units. In Version 2 of *Falcons*, the depiction of someone shooting falcons functioned as the verb *V* in “In LA, they would have *V*.” And yet an enactment of someone shooting falcons is manifestly *not* a verb. How do people infer word-like categories for depictions that allow them to function as nouns, verbs, adjectives, noun phrases, and adverbs?

Another issue is reference. In *Mozart*, MB told his audience “This is what I want to write” while playing two measures from Mozart. He used “this” to refer, not to features of his performance (the proximal scene), but to features of the notes he depicted (the distal scene). Later, when MB said “This is a passage from *this* sonata” and then played two measures from a Mozart sonata, he referred, not to his performance, but to the sonata represented in his performance. How do recipients identify these referents?

Still another issue is modification. Adjunct depictions are nonrestrictive modifiers of the words or phrases they are affiliated with. In *Tapering*, it was as if MB had said “what we would call the tapering, *which looks like this*” as he depicted the diminuendo sign. And in *Disgusting*, it was as if he had said “It couldn’t be that. I wouldn’t think so, *which I consider like this*” as he depicted himself in disgust. How do recipients interpret *nonlinguistic* depictions as modifiers of *linguistic* expressions?

Standard theories of language processing account for descriptions, but to be complete, they must also account for depictions. They must deal with scenes as well as symbol strings, layered as well as nonlayered actions, proximal and distal viewpoints as well as discourse viewpoints, and percept-bound as well as concept-bound imagination. And they must say how people combine describing and depicting in composites of the two.

References

- Atkin, A. (2010). Peirce’s theory of signs. *Stanford encyclopedia of philosophy*. Stanford, CA: Stanford University.
- Bakhtin, M. M. (1981). *The dialogic imagination: Four essays*. Austin, TX: University of Texas Press.
- Bateson, G. (1955). A theory of play and fantasy; a report on theoretical aspects of the project of study of the role of the paradoxes of abstraction in communication. *Psychiatric Research Reports*, 2, 39–51.
- Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R., & Mullett, J. (1988). Form and function in motor mimicry topographic evidence that the primary function is communicative. *Human Communication Research*, 14, 275–299. <http://dx.doi.org/10.1111/j.1468-2958.1988.tb00158.x>
- Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). “I show how you feel”: Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, 50, 322–329. <http://dx.doi.org/10.1037/0022-3514.50.2.322>
- Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1990). Motor mimicry as primitive empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development*. Cambridge: Cambridge University Press.
- Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, 19, 163–194. <http://dx.doi.org/10.1177/0261927X00019002001>
- Bavelas, J. B., Gerwing, J., & Healing, S. (2014). Effect of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. *Discourse Processes*, 51, 619–655. <http://dx.doi.org/10.1080/0163853X.2014.883730>
- Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495–520. <http://dx.doi.org/10.1016/j.jml.2007.02.004>
- Bilson, M. (2005). *Knowing the score*. [DVD]. Ithaca, NY: Cornell University.
- Bloom, P. (2010). *How pleasure works: The new science of why we like what we like*. New York, NY: Norton.
- Borges, J. L. (1999). *Everything and nothing*. New York, NY: New Directions.
- Bretherton, I. (1989). Pretense: The form and function of make-believe play. *Developmental Review*, 9, 383–401. [http://dx.doi.org/10.1016/0273-2297\(89\)90036-1](http://dx.doi.org/10.1016/0273-2297(89)90036-1)
- Bretherton, I. (Ed.). (1984). *Symbolic play: The development of social understanding*. New York, NY: Academic Press.
- Bruner, J. S., Busiek, R. D., & Minturn, A. L. (1952). Assimilation in the immediate reproduction of visually perceived figures. *Journal of Experimental Psychology*, 44, 151–155. <http://dx.doi.org/10.1037/h0062693>
- Bühler, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: G. Fischer.
- Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, 47, 14–27. <http://dx.doi.org/10.1016/j.pragma.2012.12.007>
- Cappelen, H., & Lepore, E. (1997). Varieties of quotation. *Mind*, 106, 429–450. <http://dx.doi.org/10.1093/mind/106.423.429>
- Cappelen, H., & Lepore, E. (2003). Varieties of quotation revisited. *Belgian Journal of Linguistics*, 17, 51–75.
- Cappelen, H., & Lepore, E. (2012). Quotation. *Stanford encyclopedia of philosophy*. Stanford, CA: Stanford University.
- Carmichael, L., Hogan, H., & Walter, A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology*, 15, 73–86. <http://dx.doi.org/10.1037/h0072671>
- Chovil, N. (1991a). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, 25, 163–194. <http://dx.doi.org/10.1080/08351819109389361>
- Chovil, N. (1991b). Social determinants of facial displays. *Journal of Nonverbal Behavior*, 15, 141–154. <http://dx.doi.org/10.1007/BF01672216>
- Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for interaction in action. *Journal of Experimental Psychology: General*, 143, 1726–1741. <http://dx.doi.org/10.1037/a0036281>
- Cienki, A. J., & Müller, C. (Eds.). (2008). *Metaphor and gesture*. Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/gs.3>

- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511620539>
- Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7, 507–525. <http://dx.doi.org/10.1177/1461445605054404>
- Clark, H. H. (2012). Unspoken questions, unspoken answers. In J.-P. d. Ruiter (Ed.), *Questions: Formal, functional, and interactional perspectives* (pp. 81–100). Cambridge: New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139045414.006>
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84, 73–111. [http://dx.doi.org/10.1016/S0010-0277\(02\)00017-3](http://dx.doi.org/10.1016/S0010-0277(02)00017-3)
- Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, 113, 121–126. <http://dx.doi.org/10.1037/0096-3445.113.1.121>
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66, 764–805. <http://dx.doi.org/10.2307/414729>
- Clark, H. H., & Van Der Wege, M. M. (2015). Imagination in narratives. *The Handbook of Discourse Analysis*, 2, 406–421.
- Clift, R. (1999). Irony in conversation. *Language in Society*, 28, 523–553. <http://dx.doi.org/10.1017/S0047404599004029>
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–153. <http://dx.doi.org/10.1016/j.tics.2009.01.005>
- Davidson, D. (1979). Quotation. *Theory and Decision*, 11, 27–40. <http://dx.doi.org/10.1007/BF00126690>
- De Brabanter, P. (2005). *Hybrid quotations*. Amsterdam: John Benjamins.
- Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. M. (2013). Gestural viewpoint signals referent accessibility. *Discourse Processes*, 50, 431–456. <http://dx.doi.org/10.1080/0163853X.2013.824286>
- De Ruiter, J., & Wilkins, D. (1998). *The synchronization of gesture and speech in Dutch and Arrernte (an Australian aboriginal language): A cross-cultural comparison*. Paper presented at the Oralité et Gestualité: Communication multimodale, interaction, Paris.
- Drew, P. (1987). Po-faced receipts of teases. *Linguistics*, 25, 219–253. <http://dx.doi.org/10.1515/ling.1987.25.1.219>
- Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics*, 15, 223–238. <http://dx.doi.org/10.1515/cogl.2004.009>
- Duke, R. A., & Simmons, A. L. (2006). The nature of expertise: Narrative descriptions of 19 common elements observed in the lessons of three renowned artist-teachers. *Bulletin of the Council for Research in Music Education*, 170, 7–19.
- Dunn, J., & Dale, N. (1984). I a daddy: 2-year-olds' collaboration in joint pretend with sibling and with mother. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 131–158). New York, NY: Academic Press.
- Efron, D. (1941). *Gesture and environment*. New York, NY: King's Crown Press.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48, 384–392. <http://dx.doi.org/10.1037/0003-066X.48.4.384>
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behaviour. *Semiotica*, 1, 49–98. <http://dx.doi.org/10.1515/semi.1969.1.1.49>
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124–129. <http://dx.doi.org/10.1037/h0030377>
- El'konin, D. (1969). Some results of the study of the psychological development of preschool-age children. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology* (pp. 34–61). New York, NY: Basic Books.
- Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Erlbaum.
- Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511576737>
- Engberg-Pedersen, E. (1993). *Space in Danish sign language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: SIGNUM-Press.
- Engberg-Pedersen, E. (2015). Perspective in signed discourse: The privileged status of the signer's locus and gaze. *Open Linguistics*, 1, 411–431.
- Engle, R. A. (2000). *Toward a theory of multimodal communication combining speech, gestures, diagrams, and demonstrations in instructional explanations* (Ph. D. dissertation). Stanford University, Stanford, CA.
- Farver, J. A. M. (1992). Communicating shared meaning in social pretend play. *Early Childhood Research Quarterly*, 7, 501–516. [http://dx.doi.org/10.1016/0885-2006\(92\)90047-3](http://dx.doi.org/10.1016/0885-2006(92)90047-3)
- Fein, G. G. (1975). A transformational analysis of pretending. *Developmental Psychology*, 11, 291–296. <http://dx.doi.org/10.1037/h0076568>
- Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, 52, 1095–1118. <http://dx.doi.org/10.2307/1129497>
- Fillmore, C. J. (1997). *Lectures on deixis*. Stanford, CA: CSLI Publications.
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, 20, 226–241. <http://dx.doi.org/10.1016/j.cogbrainres.2004.02.012>
- Garvey, C. (1990). *Play*. Cambridge, MA: Harvard University Press.
- Garvey, C., & Kramer, T. L. (1989). The language of social pretend play. *Developmental Review*, 9, 364–382. [http://dx.doi.org/10.1016/0273-2297\(89\)90035-X](http://dx.doi.org/10.1016/0273-2297(89)90035-X)
- Gelman, S. A., & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition*, 76, 91–103. [http://dx.doi.org/10.1016/S0010-0277\(00\)00071-8](http://dx.doi.org/10.1016/S0010-0277(00)00071-8)
- Gendler, T. (2011). Imagination. *Stanford encyclopedia of philosophy*. Stanford, CA: Stanford University.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170. http://dx.doi.org/10.1207/s15516709cog0702_3
- Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science*, 10, 139–146. <http://dx.doi.org/10.1111/j.1467-7687.2007.00576.x>
- Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. New Haven, CT: Yale University Press.
- Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157–195. <http://dx.doi.org/10.1075/gest.4.2.04ger>
- Geurts, B., & Maier, E. (2003). Quotation in context. *Belgian Journal of Linguistics*, 17, 109–128.
- Gibbs, R. (2000). Metarepresentations in staged communicative acts. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 389–410). Oxford: Oxford University Press.
- Giffin, H. (1984). The coordination of meaning in the creation of a shared make-believe reality. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 73–100). New York, NY: Academic Press.
- Gleitman, L. R., & Liberman, M. (1995). *An invitation to cognitive sciences, Vol. 1: Language*. Cambridge, MA: MIT Press.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.
- Goffman, E. (1981). *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.
- Goldrick, M., Ferreira, V., & Miozzo, M. (2014). *The Oxford handbook of language production*. New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/oxfordhb/9780199735471.001.0001>
- Goldstein, J. H. (1994). *Toys, play, and child development*. Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511527616>
- Gombrich, E. H. (1960). *Art and illusion: A study in the psychology of pictorial representation*. Princeton, NJ: Princeton University Press.

- Goodman, N. (1974). On some questions concerning quotation. *The Monist*, 58, 294–306. <http://dx.doi.org/10.5840/monist197458223>
- Goodman, N. (1978). *Ways of worldmaking*. Indianapolis, IN: Hackett Publishing.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York, NY: Academic Press.
- Goodwin, M. H. (1990). *He-said-she-said: Talk as social organization among black children*. Bloomington, IN: Indiana University Press.
- Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62, 51–76.
- Gregory, R. L. (1968). Perceptual illusions and brain models. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 171, 279–296. <http://dx.doi.org/10.1098/rspb.1968.0071>
- Gregory, R. L. (1970). *The intelligent eye*. New York, NY: McGraw-Hill.
- Gregory, R. L. (2005). The Medawar Lecture 2001 Knowledge for vision: Vision for knowledge. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 360, 1231–1251.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66, 377–388.
- Grice, H. P. (1978). Some further notes on logic and conversation. In P. Cole (Ed.), *Syntax and semantics: Vol. 9. Pragmatics* (pp. 113–127). New York, NY: Academic Press.
- Grice, H. P. (1991). *Studies in the way of words*. Harvard, CT: Harvard University Press.
- Gullberg, M., & Holmqvist, K. (1999). Keeping an eye on gestures: Visual perception of gestures in face-to-face communication. *Pragmatics & Cognition*, 7, 35–63. <http://dx.doi.org/10.1075/pc.7.1.04gul>
- Haight, W. L., & Miller, P. J. (1993). *Pretending at home: Early development in a sociocultural context*. Albany, NY: SUNY Press.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9, 697–704. <http://dx.doi.org/10.1093/cercor/9.7.697>
- Harley, T. A. (2013). *The psychology of language: From data to theory*. New York, NY: Psychology Press.
- Harris, P. L. (2000). *The work of the imagination*. Oxford: Blackwell Publishing.
- Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, 58, 1–107.
- Heath, C. P., & Healey, P. G. (2012). Making space for interaction: Architects design dialogues. In E. Efthimiou, G. Kouroupetoglou, & S.-E. Fotinea (Eds.), *Gesture and sign language in human-computer interaction and embodied communication* (pp. 250–261). Berlin: Springer. http://dx.doi.org/10.1007/978-3-642-34182-3_23
- Hegarty, M., Mayer, S., Kriz, S., & Keehner, M. (2005). The role of gestures in mental animation. *Spatial Cognition and Computation*, 5, 333–356. http://dx.doi.org/10.1207/s15427633scc0504_3
- Holler, J., & Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24, 267–289. <http://dx.doi.org/10.1080/01690960802095545>
- Holt, E. (2000). Reporting and reacting: Concurrent responses to reported speech. *Research on Language and Social Interaction*, 33, 425–454. http://dx.doi.org/10.1207/S15327973RLSI3304_04
- Holt, E. (2009). Reported speech. In S. D'hondt, J.-O. Östman, & J. Verschueren (Eds.), *The pragmatics of interaction* (pp. 190–205). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/hoph.4.11hol>
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295–355. http://dx.doi.org/10.1207/s15516709cog1303_1
- Isaacs, E. A., & Clark, H. H. (1990). Ostensible invitations. *Language in Society*, 19, 493–509. <http://dx.doi.org/10.1017/S0047404500014780>
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and non-verbal communication* (pp. 207–227). The Hague: Mouton.
- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, 27, 175–200. http://dx.doi.org/10.1207/s15327973rlsi2703_2
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kirsh, D. (2010). *Thinking with the body*. Paper presented at the 32nd Annual Conference of the Cognitive Science Society, Austin TX.
- Kirsh, D. (2011). How marking in dance constitutes thinking with the body. *Versus: Quaderni di Studi Semiotici*, 113–115, 179–210.
- Klippel, E., & Gurney, J. (2002). Some observations on deixis to properties. In K. Van Deemter & R. Kibble (Eds.), *Information sharing: Reference and presupposition in language generation and interpretation* (pp. 355–390). Stanford, CA: CSLI Publications.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, 2, 635–642. <http://dx.doi.org/10.1038/35090055>
- Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and Speech*, 48, 29–53. <http://dx.doi.org/10.1177/00238309050480010201>
- Leslie, A. M. (1987). Pretense and representation: The origins of “theory of mind.” *Psychological Review*, 94, 412–426. <http://dx.doi.org/10.1037/0033-295X.94.4.412>
- Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT press.
- Liddell, S. K. (1995). Real, surrogate, and token space: Grammatical consequences in ASL. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 19–41). Mahwah, NJ: Erlbaum.
- Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511615054>
- Lillard, A. S., Pinkham, A. M., & Smith, E. (2011). Pretend play and cognitive development. In U. Gowwami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 285–311). New York, NY: Wiley-Blackwell.
- Lodge, D. (1990). *After Bakhtin: Essays on fiction and criticism*. London: Routledge.
- Lowe, M. (1975). Trends in the development of representational play in infants from one to three years—An observational study. *Journal of Child Psychology and Psychiatry*, 16, 33–47. <http://dx.doi.org/10.1111/j.1469-7610.1975.tb01870.x>
- Mann, B. (1984). Effects of realistic and unrealistic props on symbolic play. In T. Yawkey & A. Pellegrini (Eds.), *Child's play: Developmental and applied* (pp. 359–376). Mahwah, NJ: Erlbaum.
- Maynard, P. (1994). Seeing double. *The Journal of Aesthetics and Art Criticism*, 52, 155–167. <http://dx.doi.org/10.2307/431162>
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2008). *Gesture and thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (Ed.), (2000). *Language and gesture*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511620850>
- Melton, J. (Ed.), (1996). *Encyclopedia of occultism and parapsychology* (4th ed., Vol. 2). Detroit, MI: Gale.
- Miller, P., & Garvey, C. (1984). Mother-baby role play: Its origins in social support. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 101–130). Orlando, FL: Academic Press.
- Mitchell, R. W. (2002). *Pretending and imagination in animals and children*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511542282>
- Murphy, K. M. (2004). Imagination as joint activity: The case of architectural interaction. *Mind, Culture, and Activity*, 11, 267–278. http://dx.doi.org/10.1207/s15327884mca1104_3

- Murphy, K. M. (2005). Collaborative imagining: The interactive use of gestures, talk, and graphic representation in architectural practice. *Semiotica*, 156, 113–145.
- Nelson, K., & Seidman, S. (1984). Playing with scripts. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 45–71). New York, NY: Academic Press.
- Özyürek, A. (2002). Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688–704. <http://dx.doi.org/10.1006/jmla.2001.2826>
- Pagin, P., & Westerståhl, D. (2010). Pure quotation and general compositionality. *Linguistics and Philosophy*, 33, 381–415. <http://dx.doi.org/10.1007/s10988-011-9083-8>
- Parrill, F. (2009). Dual viewpoint gestures. *Gesture*, 9, 271–289. <http://dx.doi.org/10.1075/gest.9.3.01par>
- Partee, B. H. (1973). The syntax and semantics of quotation. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 410–418). New York, NY: Holt Rinehart and Winston.
- Peirce, C. S. (1932). The icon, index, and symbol. In C. Hartshorne & P. Weiss (Eds.), *Collected papers of Charles Sanders Peirce* (Vol. 2, pp. 156–173). Cambridge, MA: Harvard University Press.
- Peirce, C. S. (1974). *Collected papers of Charles Sanders Peirce* (Vol. 3). Harvard, CT: Harvard University Press.
- Perlman, M., Clark, N., & Johansson Falck, M. (2015). Iconic prosody in story reading. *Cognitive Science*, 6, 1348–1368.
- Perniss, P. M. (2007). *Space and iconicity in German sign language (DGS)*. (Ph. D.). Nijmegen, Netherlands: Radboud University.
- Perniss, P. M., & Özyürek, A. (2008). Representations of action, motion, and location in sign space: A comparison of German (DGS) and Turkish (TİD) Sign Language narratives. *Signs of the time: Selected papers from Theoretical Issues in Sign Language Research*, 8, 353–378.
- Piaget, J. (1962). *Play, dreams and imitation in childhood*. London: Routledge & Kegan Paul.
- Polanyi, L. (1989). *Telling the American story*. Cambridge, MA: MIT Press.
- Potts, C. (2007). The dimensions of quotation. In C. Barker & P. Jacobson (Eds.), *Direct compositionality* (pp. 405–431). Oxford: Oxford University Press.
- Preissler, M. A., & Bloom, P. (2008). Two-year-olds use artist intention to understand drawings. *Cognition*, 106, 512–518. <http://dx.doi.org/10.1016/j.cognition.2007.02.002>
- Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics*, 70, 111–126. <http://dx.doi.org/10.1007/s10649-008-9127-3>
- Recanati, F. (2001). Open quotation. *Mind*, 110, 637–687. <http://dx.doi.org/10.1093/mind/110.439.637>
- Recanati, F. (2008). Open quotation revisited. *Philosophical Perspectives*, 22, 443–471. <http://dx.doi.org/10.1111/j.1520-8583.2008.00153.x>
- Saka, P. (1998). Quotation and the use-mention distinction. *Mind*, 107, 113–135. <http://dx.doi.org/10.1093/mind/107.425.113>
- Sawyer, R. K. (1997). *Pretend play as improvisation: Conversation in the preschool classroom*. Mahwah, NJ: L. Erlbaum Associates.
- Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266–296). Cambridge: Cambridge University Press.
- Schegloff, E. A. (1998). Body torque. *Social Research*, 65, 535–596.
- Shakhashiri, B. Z. (1989). *Chemical demonstrations: A handbook for teachers of chemistry*. Madison, WI: University of Wisconsin Press.
- Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, 91, 417–447. <http://dx.doi.org/10.1037/0033-295X.91.4.417>
- Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, 55, 167–177. <http://dx.doi.org/10.1016/j.jml.2006.03.002>
- Sidnell, J. (2006). Coordinating gesture, talk, and gaze in reenactments. *Research on Language and Social Interaction*, 39, 377–409. http://dx.doi.org/10.1207/s15327973rlsi3904_2
- Sokoloff, D. R., & Thornton, R. D. (2004). *Interactive lecture demonstrations: Active Learning in Introductory Physics*. New York, NY: Wiley.
- Spivey, M., McRae, K., & Joannis, M. (Eds.). (2012). *The Cambridge handbook of psycholinguistics*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139029377>
- Stec, K. (2012). Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture*, 12, 327–360. <http://dx.doi.org/10.1075/gest.12.3.03ste>
- Stec, K., Huiskes, M., & Redeker, G. (2015). Multimodal analysis of quotation in oral narratives. *Open Linguistics*, 1, 531–554. <http://dx.doi.org/10.1515/opli-2015-0018>
- Streeck, J. (2008). Depicting by gesture. *Gesture*, 8, 285–301. <http://dx.doi.org/10.1075/gest.8.3.02str>
- Streeck, J., Groethues, J., & Villanueva, J. (2009). *Gesturecraft: The manufacture of meaning*. Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/gs.2>
- Strömquist, S. (1984). *Make-believe through words: A linguistic study of children's play with a doll's house* (Ph. D. thesis). University of Göteborg, Göteborg Sweden.
- Stukenbrock, A. (2012). Imagined spaces as a resource in interaction. *Bulletin Suisse de Linguistique Appliquée*, 96, 141–161.
- Svartvik, J., & Quirk, R. (Eds.). (1980). *A corpus of English conversation*. Lund: Gleerup.
- Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, 53, 81–94. <http://dx.doi.org/10.1016/j.jml.2005.02.003>
- Tannen, D. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (Ed.), *Direct and indirect speech* (pp. 311–332). Berlin: Mouton. <http://dx.doi.org/10.1515/9783110871968.311>
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Taylor, M. (1999). *Imaginary companions and the children who create them*. Oxford: Oxford University Press.
- Taylor, M. (2007). *British pantomime performance*. Bristol, UK: Intellect.
- Taylor, M., & Mannering, A. M. (2006). Of Hobbes and Harvey: The imaginary companions created by children and adults. In A. Göncü & S. Gaskins (Eds.), *Play and development: Evolutionary, socio-cultural, and functional perspectives* (pp. 227–245). New York, NY: Erlbaum.
- Thompson, S. A., & Suzuki, R. (2014). Reenactments in conversation: Gaze and reciprocity. *Discourse Studies*, 16, 816–846. <http://dx.doi.org/10.1177/1461445614546259>
- Tversky, B., & Suwa, M. (2009). Thinking with sketches. In A. B. Markman & K. L. Wood (Eds.), *Tools for innovation* (pp. 75–84). Oxford: Oxford University Press.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language*, 32, 805–819. <http://dx.doi.org/10.1006/jmla.1993.1040>
- Walton, K. L. (1978). Fearing fictions. *The Journal of Philosophy*, 75, 5–27. <http://dx.doi.org/10.2307/2025831>
- Walton, K. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.

- White, G. (2012). On immersive theatre. *Theatre Research International*, 37, 221–235. <http://dx.doi.org/10.1017/S0307883312000880>
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47, 9–12. <http://dx.doi.org/10.1016/j.neuron.2005.06.013>
- Zwaan, R. A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology*, 39, 1142–1150. <http://dx.doi.org/10.1002/ejsp.661>
- Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, 18, 229–234. <http://dx.doi.org/10.1016/j.tics.2014.02.008>

Received May 26, 2015

Revision received December 22, 2015

Accepted December 23, 2015 ■