10

Viewpoint

Lieven Vandelanotte

In the humanities, viewpoint has long been studied in the context of narrative, with a focus on such questions as focalization (e.g. Genette 1980, Bal 2009) and speech and thought representation (e.g. Banfield 1982, Fludernik 1993). In formal linguistics, analyses of logophoric pronouns (e.g. Cantrall 1974, Culy 1997) and so-called empathetic syntax (Kuno 1987) provide early foci of interest, but arguably cognitive linguistics, with its principled consideration of construal choices and viewing arrangements, has been well-placed since its inception to elaborate the notion. However, viewpoint has recently started to take center stage, as witnessed by edited volumes such as Dancygier and Sweetser (2012), Dancygier, Lu, and Verhagen (2016), Dancygier and Vandelanotte (2017a), and Vandelanotte and Dancygier (in prep). This focusing of efforts is owed in part to recent developments in cognitive science and psychology, where the embodied simulation hypothesis was formulated and tested (see, e.g., Bergen 2012), fostering further viewpoint studies across a range of modalities and research paradigms, whether experimental or more theoretical. The pervasive importance of viewpoint offers a challenge and an opportunity to demonstrate the usefulness of tools such as image schemas, frames, mental spaces, and blends to coherently formulate and analyze viewpoint complexities in grammar, discourse, and multimodal artifacts.

This chapter proceeds as follows. In its first section, general discussion shows that viewpoint is not only ubiquitous, but also intersubjectively and multimodally construed. The second section looks at specific viewpoint phenomena at different levels of structure in different modes of expression, and suggests ways of integrating multiple local viewpoints into a global viewpoint network. Based on this exploration, the third section proposes to conceive of types of viewpoint along a deictic–cognitive divide.

10.1 General Properties of Viewpoint: Embodied, Multimodal, Intersubjective

The potential application of the notion of viewpoint can seem bewilderingly broad, as ultimately any aspect of construal reveals something about the views of the language user: choosing one word over another, one focus over another, one temporal or aspectual form over another, and the list goes on (cf. Langacker 2008a: 55–89). At the same time, the term 'viewpoint' etymologically gives pride of place to the visual modality, although not all aspects of viewpoint can be or need to be visually imagined. This reflects the way in which vision is prioritized among the senses (Bergen 2012: 49), and its links with conceptualization and understanding are deeply rooted (Sweetser 1990, Langacker 1995).

By way of initial definition, viewpoint will be understood here as "a discourse participant's alignment with an aspect of a frame or situation" (Dancygier and Vandelanotte 2016: 14). In other words, to speak of viewpoint requires a conceptualizer in a discourse event assuming a position (for instance, in terms of perception and spatiotemporal location, likelihood and knowledge, attitude and feeling, or solidarity/power dynamics) toward an element within a described situation or knowledge structure. Viewpoint is, in this view, a fairly comprehensive notion moving well beyond the spatial and, by extension, temporal, location toward or within a described scene; the advantage of such an inclusive understanding is that it points to a unified conception of low-level phenomena all the way up to viewpoint flows in discourses of different sizes, from Internet memes to whole novels. In analyzing viewpoint, we are trying to get at a discourse participant's conceptualization via its observable realizations in forms of language and in various forms of embodied behavior.

Viewpoint is ubiquitous and inescapable: in Sweetser's formulation, "we never have experience of the world except as a viewpoint-equipped, embodied self among other viewpointed, embodied selves" (2012: 1). This summing-up focuses attention on both the embodied nature of viewpoint and the fact that it is construed intersubjectively, in a negotiation with other participants in a given speech event. Both concepts are addressed more fully in other chapters of this handbook, but their specific relation to and relevance for viewpoint warrants discussion here.

An immediate sense in which viewpoint is embodied pertains to the body articulators' involvement in expressing viewpoints, as a string of studies on co-speech gesture have been making clear at least since the seminal work by McNeill (1992, 2005), who distinguished between observer viewpoint and character viewpoint in so-called 'iconic' gestures ("bear[ing] a close formal relationship to the semantic content of speech," McNeill 1992: 12). For example, a cartoon prompt involving a skunk hopping up and down in an experiment led by Parrill (2009, 2010) leads some

participants retelling the cartoon to use their finger to trace the pathway traversed by the skunk (observer viewpoint), whereas others 'embody' the skunk, imitating its motion and possibly other (e.g. facial) features (character viewpoint).

As many researchers have observed, ignoring co-speech gesture in analyzing spoken discourse leads to impoverished or misleading results, also in terms of viewpoint shifts (e.g. Narayan 2012). Conversely, comparing results may prove to be mutually reinforcing, as in Parrill's (2012) study showing that given versus new discourse status, whose relevance to linguistic referring expressions has long been established, similarly impacts on gesture choice: in the absence of shared knowledge, speakers produce significantly more gestures, and significantly more character viewpoint gestures. Another such correlation is that between transitivity and gestural viewpoint, with transitive and intransitive utterances showing a significant preference for character and observer viewpoint utterances respectively (Beattie and Shovelton 2002, Parrill 2010).

Unlike the linear processing of spoken or written linguistic structure, the visual presence of a whole body in spoken interaction affords a simultaneous presence of different viewpoints in what is known as body partitioning (Dudis 2004), whereby different sections of the body can express different viewpoints. For instance, Sweetser and Stec (2016) describe cases where gaze and facial expression represent one viewpoint, and the speaker's hands and body simultaneously represent another. While "single-viewpoint gestures are the most frequent" (Stec 2012: 333), multiple viewpoint gestures have been recorded and discussed for both co-speech gesture (e.g. Parrill 2009, Sweetser 2013) and sign language (e.g. Janzen 2004).

As the example of multiple viewpoint gestures shows, viewpoint embodiment involves more than mere manual gesturing, and detailed analyses of interactions need also to consider such parameters as body posture, gaze and facial expressions, head tilts and nods and, at the aural level, "vocal gestures such as pitch, intonation and accent" (Stec 2012: 332), all of which can help to effect shifts in viewpoint (see Stec 2012 for a review). Current work is starting to address the finer points of these other parameters, for instance in studying the role of head nods and tilts and of facial expressions in sarcasm (Tabacaru 2014) or in describing the ways in which gaze performs viewpoint functions such as enacting characters, visually 'checking' that interlocutors are engaged and 'on board,' or signalling that speakers are 'consulting' their memories (Sweetser and Stec 2016).

While sign languages such as ASL or BSL use broadly similar viewpoint strategies to co-speech gesture (Stec 2012, Quinto-Pozos and Parrill 2015), embodied viewpoint in them arguably is more structured, and "clearly more conventionalised" (Earis and Cormier 2013: 340). As an example, consider types of viewpoint shift in ASL. The default type shows some combination of body shifting, pointing and/or eye gaze directed toward

some static location to mark whose viewpoint is in focus when comparing characteristics of different referents (Janzen 2012b). In the special case of mentally rotated space (Janzen 2004, 2012b), on the other hand, the signer rotates the other viewpoint mentally to their own position. In one example discussed by Janzen, a signer, in explaining how she was signalled by the police to pull over, switches, across two consecutive constructions, from a clear leftward moving gesture (adopting the police officer's viewpoint) to a rightward move (adopting her own again).

In addition to 'outwardly perceivable' viewpoint embodiment, there is now a rich literature on embodied cognition focusing on phenomena in the brain, whose effects are demonstrated in experimental work. An accessible introduction covering this field is Bergen (2012); here I only highlight some of the research showing *that* and *how* "in understanding language, we use our perceptual and motor systems to run embodied simulations" (Bergen 2012: 195).

Various picture-sentence matching tests have shown that people respond faster when an entity's orientation or shape in the picture matches the implied orientation or shape in the sentence prompt in examples such as *The carpenter hammered the nail into the* ***floor*** versus *into the* ***wall*** or *The ranger saw the eagle in the* ***nest*** vs. *in the* ***sky*** (Stanfield and Zwaan 2001, Zwaan, Stanfield, and Yaxley 2002). This suggests that we visually simulate objects we read or hear about, and there is good evidence that this simulation is viewpointed. More importantly, the viewpoint adopted in simulations can be manipulated by language. For instance, in one experiment (Yaxley and Zwaan 2007), the sentence prompts were *Through the* ***clean*** *goggles, the skier could* ***easily*** *identify the moose* versus *Through the* ***fogged*** *goggles, the skier could* ***hardly*** *identify the moose*; remarkably, people were quicker to identify a high-resolution picture of a moose following the 'clean' sentence and a low resolution picture in the 'fogged' case. Understanding language really does seem to be "almost like being there," in Bergen's phrase, with simulations occurring in language comprehension adopting the viewpoint of an immersed experiencer (Bergen 2012: 66).

In addition to words, grammatical structures such as argument structure constructions and grammatical aspect influence different facets of simulations (Bergen 2012: ch. 5). Of particular interest is the influence of grammatical person: second-person sentences have been shown to prompt internal, participant viewpoint; third-person sentences external, observer viewpoint (Brunyé et al. 2009). For instance, in an experiment run by Bergen and his lab (2012: 110–13), reaction times were quicker when an 'away' visual prompt (showing two images of a baseball of which the second is smaller) followed a second-person sentence (***you*** rather than ***the pitcher*** *threw the baseball to the catcher*).

To describe viewpoint as fundamentally embodied implies describing it as fundamentally multimodal as well: combining spoken language with

co-speech gesture, for instance, naturally relies on both the vocal/auditory modality and the kinesic/visual one (Green 2014: 9). An impressive case study of multimodality in this strict sense is Green's (2014) analysis of Central Australian sand stories, in which spoken language and song are supported by co-speech gesture, sign language, and sand drawings, the latter using both graphic schemas and 'props' such as leaves representing story characters. In a looser sense, however, a form of multimodality can also be posited for communicative artifacts that use different semiotic resources, even if, for instance, they are all visuo-spatial, as in combining text and image in comics or advertising. There is a rich tradition of communication research centered on this broader conception of multimodality (e.g. Kress and Van Leeuwen 2001), and in cognitive linguistics the question has been focused on multimodal metaphor (Forceville and Urios-Aparisi 2009) and multimodal interaction (Feyaerts, Brône and Oben this volume Ch. 9). One recent study in this vein shows how a variety of Internet memes serve constantly to reconstrue evoked viewpoints, making effective use of the (semi-)constructional status which linguistic and visual elements alike acquire (Dancygier and Vandelanotte 2017b).

A final general property of viewpoint to be highlighted here is its intersubjective nature. At first glance, this is common sense: suppose you are teaching in a classroom, for instance, you take into account the fact that what you see in front of you on your computer screen students see projected behind you. We are constantly aware of other viewpoints. Less commonsensically, perhaps, we appear to be hard-wired to monitor and, in an attenuated sense, even simulate other humans' actions around us, in the sense that some of the brain structures used to perform a given action are activated when we "recognize, predict and understand" (Bergen 2012: 77) someone else carrying out that action. Research on monkeys' so-called mirror neurons (e.g. Rizzolatti, Fogassi, and Gallese 2001) made scientists alive to this reality; more indirect brain imaging evidence suggests humans too operate with mirror systems (e.g. Aziz-Zadeh et al. 2006).

Converging trends in a string of related fields of inquiry have demonstrated the pervasive significance of intersubjectivity. In both phylogeny and ontogeny, prelinguistic intersubjective skills such as imitation of actions, joint attention and pointing 'ground' language, to use Zlatev's (2008) term, and more complex, linguistic intersubjective skills continue to develop as language evolves. For instance, Tomasello (1999: 173) lists disagreements, repairs/explanations and meta-discourse as types of discourse in which children need to be able to model another's viewpoint; at the more basic level of selecting lexical items (e.g. *cat* versus *animal*), children appear to take into account which viewpoint is pragmatically being cued for (Clark 1997). Children's acquisition of advanced 'theory of mind' skills (such as false-belief understanding or pretence) further illustrates that being able to share and understand the experiences of others underlies much of what we do.

In linguistics, intersubjectivity has inspired work in various fields. In diachrony, for instance, much work has focused on changes from less or non-subjective, to subjective and possibly intersubjective, taking into account the addressee's viewpoint (e.g. Traugott 2010); in linguistic typology, a wide range of multiple viewpoint constructions are being analyzed (e.g. Evans 1995); and in construction grammar, constructions such as complementation, connectives, and sentential negation have been shown to have particular intersubjective orientations, in the sense of serving to coordinate mental states between conceptualizers (Verhagen 2005). It is to viewpoint constructions and how best to analyze them that the next section turns.

10.2 Tools for Analyzing Viewpoint Networks

This section focuses on new contributions that Cognitive Linguistics has made in the study of viewpoint in terms of range of phenomena studied and of analytical tools used. Cognitive linguists have contributed to existing research programs on narratorship and focalization (e.g. Bal 2009, Hühn, Schmid, and Schönert 2009b), speech and thought representation (e.g. Banfield 1982, Fludernik 1993), and deixis and evidentiality (e.g. Hanks 1992, Aikhenvald 2004, Evans 1995). At the same time, they have also revealed the viewpoint potential of hitherto neglected constructions; they have brought into focus the manifold interactions between viewpoint and image schemas, frames and mental spaces, on which metaphorical and metonymic mappings are built; and they have ventured beyond sentence-level interpretations to analyze the management of viewpoints within smaller and larger discourses.

Speech and thought representation was interpreted in terms of mental space embedding in Sanders and Redeker (1996), with the embedded space being accessed directly, as a new Base Space, in direct speech, and only indirectly, with varying degrees of 'narrator' influence, in (free) indirect speech/thought; they also distinguish forms of implicit viewpoint (cf. Sanders 2010). The specific kind of blend involved in the free indirect mode was characterized by Vandelanotte (2004, 2009, 2012) as locating viewpoint with the represented speaker (or character), in contrast to a separate type, distancing indirect speech/thought, in which the current versus represented speaker discourse blend locates viewpoint with the current speaker and is decompressed to construe a higher (attitudinally 'distanced') space from which the blend is ultimately viewed. An example is (1), in which the I-narrator Max is questioned by Carlo Grace about how he and his parents spend their vacation in a chalet; the *we* corresponds to an original addressed *you*, drawn into the deictic viewpoint of the current speaker, who incorporates Carlo's discourse without shifting the deictic centre.

(1) Now and then from the other end of the table Carlo Grace, chewing vigorously, would bend a lively gaze on me. What was life like at the chalet, he wanted to know. What did we cook on? A Primus stove, I told him. "Ha!" he cried. "Primus *inter pares*!" (Banville 2005: 208; italics original)

This work on discourse representation phenomena broadly assumes a notion of "discoursally motivated" (Nikiforidou 2012: 193) construction. The reliance particularly of blended forms such as free and distancing indirect speech/thought on the narrative context clearly sets them apart from more readily identifiable constructions such as morphemes or argument structure constructions. Recent suggestions concerning the question of what counts as a construction promise a better handle on this problem by focusing on smaller-scale component constructions which may frame-metonymically evoke the larger constructional frame. This is what Dancygier and Sweetser (2005) refer to as constructional compositionality (see also Dancygier 2009, 2011), and it is in this spirit that Nikiforidou (2010, 2012, 2015) analyzes the *was* + *now* pattern as one construction which, in marking a viewpoint shift away from the current speaker, may prompt the higher-level narrative construction of free indirect discourse (as in Banfield's example *How my heart was beating now!*, 1982: 99). Along similar lines other typical 'features' of free indirect discourse, such as logophoric or 'viewpoint' reflexives (e.g. Zribi-Hertz 1989, Van Hoek 1997) or viewpoint-shifting connectives (Sotirova 2004), could be reinterpreted as such subconstructions.

Further scrutiny by Dancygier (2012a: 190–93) has shown that *now* in past tense contexts acquires a resumptive function, marking a return to the narrative events currently in focus, as in example (2), in which the narrator uses *now* to shift away from what the library stacks have always represented to him in the (pre-story) past, back to the specific moment being narrated:

(2) Realizing I was stuck awhile, I began to see the place differently. The stacks had always been a purely functional means to an end. But now, I *lived* there. A long night ahead, and the third-biggest collection in the country to pass it on. (Richard Powers, *The Gold Bug Variations*; quoted in Dancygier, 2017b)

Now is only one among various examples where the viewpoint potential of 'small' constructions is reconsidered or even discovered in recent work, also including, for example, genitives, determiners, and negative markers (Dancygier 2009, 2011, 2012b). Even a widely studied language such as English still holds constructions with unexpected viewpoint function in certain contexts, though they are not necessarily grammaticalized, 'dedicated' viewpoint constructions like *was* + *now* or free indirect speech (cf. Dancygier 2016a: 283).

Some uses of genitives premodifying proper names – which naturally call up rich frames of knowledge – were shown by Dancygier (2009, 2011) to function as markers of viewpoint. In talking about "one Hillary who has not yet reached the summit of her political Everest," for instance (quoted in Dancygier 2009: 171), the frame of Mt. Everest and Edmund Hillary's climbing of it is selected to reflect the viewpoint of Hillary Clinton as seeing the presidency as the ultimate achievement. 'Her' thus means something like 'as she sees it,' which is not standardly the case in speaking of, for instance 'Hillary's success.' A similarly viewpointed use of genitives is found in the pattern *one person's trash is another person's treasure*, in which one and the same referent is evaluated negatively by some (*trash*) and positively by others (they view it as a *treasure*). Note again the difference with ordinary uses as in *I stumbled over my neighbor's trash*, in which *my neighbor* is not the 'viewer' of the trash. A striking example of the pattern is (3), commenting on the question of whether you should respect gay people if you find homosexuality 'immoral.' The value-laden expressions here are *sacred cow* and *double cheeseburger*, with the latter (unlike *treasure*) receiving its evaluative charge by virtue of the construction, and receiving ample metaphorical elaboration in the next sentence.

(3) Your sacred cow is somebody else's double cheeseburger. If you will simply look the other way when they are eating, not make rude noises, and don't try to legislate against their lunch, that'd be plenty respect right there. (www.slate.com/blogs/outward/2014/01/06/should_you_respect_gay_people_if_you_find_homosexuality_immoral.html)

As a final example of recent extensions of the repertoire of viewpointed uses of constructions, consider the case of intersubjective negation described by Dancygier (2012b) and exemplified in (4), in which *I don't hope* refutes the hopeful stance which the interviewer attributes to the actor being interviewed, thereby invalidating the question itself:

(4) – What do you hope the mainstream audiences will learn about this artistic/bohemian culture. . .
– I don't hope they learn anything . . . I don't know who goes to see films and why. (Internet example quoted in Dancygier 2012b: 85)

The role of both established and more discoursally based constructions is one important strand in current viewpoint research, and suggests that starting from low-level, small-scale constructions, which may or may not form part of larger constructions, is a fruitful approach. Another strand, represented in Dancygier and Sweetser (2014), highlights the role of viewpoint in figurative language. Even relatively mundane examples of metonymic and metaphoric usage turn out to subtly imply specific viewpoints: a metonymic example such as *Table 5 needs more water*, for instance, relies on the Restaurant frame from the viewpoint of the waiting staff, not the customer (Dancygier and Sweetser 2014: 21); or the oft-analyzed metaphor

That surgeon is a butcher only really works when the viewpoint adopted is that of the patient in the Surgery frame (Dancygier and Sweetser 2014: 216–17).

Recent work by Dancygier (2016b) has drawn attention to an even more fundamental level at which viewpoint comes in, and on which subsequent framing, metaphor, and blending operates, and that is the level of image-schematic structure as defined in the work of, for example, Johnson (1987), Lakoff and Johnson (1999), and Mandler and Pagán Cánovas (2014). Using the example of a barrier, involved in infants' primary scenes such as being enclosed in a play pen, Dancygier (2016b) shows a range of increasingly elaborate construals in which barriers are crossed, removed, or reconstrued in textual and visual artifacts involving walls and borders, whether real or imaginary. The essential viewpointed experience of a barrier – being on either side of it, causing the experiencer to be restricted in terms of motion and vision, or having an aerial view of both sides – provides the scaffolding for these creative elaborations. To cite one visual example Dancygier analyzes, Banksy's West Bank wall art creates different viewpointed experiences of 'unwalling' the wall, for instance by smashing it open with a wrecking ball, peeling it off like wallpaper or flying over it by hanging onto a bunch of balloons.

In addition to image schemas and frames, also mental spaces and blending have been used productively to analyze viewpoint structures, and this across different modes and modalities. In his analysis of a visual poem by bpNichol, Borkent (2010) for instance unpacks a series of blends (including the 'pond' frame and fictive motion, among others) involved in 'viewing' a poem consisting of three handwritten lines containing the graphemes 'fr,' 'pond,' and 'glop,' with a curvy line connecting the 'r' and 'g' via the 'o' of 'pond.' As Borkent shows, the poem construes, with minimal formal means, the viewpoint of the agent (the frog), lying on the muddy edge of a pond, jumping in, producing a 'glop' sound and leaving behind it only o-shaped ripples on the water. In his study of gesture in political rhetoric, Guilbeault (2017) demonstrates how politicians often paraphrase their opponents' viewpoints verbally, while criticizing and contextualizing these viewpoints in the more implicit gestural modality, thereby reducing cognitive load. In analyses of co-speech gesture and sign language more generally, the actual physical environment (including the body) of a language user serves as an input (the 'Real Space,' Liddell 1998) to viewpoint blends of varying complexity.

Returning to written narratives, Dancygier (2005) first proposed the mechanism of viewpoint compression which allows lower-level viewpoints to be integrated with higher-level (narrative) viewpoint. In (5), for instance, a travel writer, Jonathan Raban, recounts his experience of seeing a TV report covering his own departure; the lower viewpoint is that of

the writer-as-TV viewer, enriching the higher-level viewpoint of the narrator of *Old glory*:

(5) The TV news went local. An Englishman had left Minneapolis that day in a small motor boat . . . In the picture on the screen his face had a cheesy pallor . . . He looked to me like a clowning greenhorn. (Jonathan Raban, *Old glory*; quoted in Dancygier 2005: 109)

This is one case of what Dancygier (2012a: 100–02) calls decompression for viewpoint: for the duration of the description of the TV report, the narrative viewpoint is compressed with the 'TV viewer' persona as a decompressed subidentity of Raban, and the traveller persona moves temporarily to the background. A more complex case is (6), in which the 'insider' viewpoint, compressed with the narrative viewpoint, is contrasted with the 'outsider,' tourist's viewpoint on Cairo, but both remain accessible:

(6) My fellow-diners and I had come at Cairo from different angles, and we'd arrived at different places. They'd flown from Gatwick to the land of the Pharaohs, while I had made a homecoming of sorts from Sana'a. (Jonathan Raban, *Arabia*; quoted in Dancygier 2005: 110)

These and related phenomena were integrated into a cognitive model of how readers construct stories from narrative text in Dancygier (2012a), which distinguishes a Main Narrative Space (MN) containing various embedded Narrative Spaces (NS), with all of this structure ultimately supervised by the Story-Viewpoint Space (SV) which comprises the features of narratorship; in this model, readers construct stories by blending elements from the different spaces available. The approach has been applied to fragmented contemporary novels (Dancygier 2007, Vandelanotte 2015) as well as news narratives (Van Krieken, Sanders, and Hoeken 2016). As can be expected, the construals can be very complex across multiple embedded spaces, but by way of illustration, let us reconsider two examples previously discussed.

Figure 10.1 represents the decompression of Raban in example (5) above into 'traveller' and 'TV viewer,' and the subsequent compression of the 'TV viewer' with the narrative viewpoint. Figure 10.2 revisits example (1), in which the character Carlo Grace's questions 'What is life like at the chalet?' and 'What do you cook on?' get to be represented as 'What was life like at the chalet? What did we cook on?' from the deictic viewpoint of the I-narrator Max. As Figure 10.2 shows, the pronoun *we*, referring to the I-narrator Max and his parents, functions at the level of the embedded Discourse Space, where it has the role of addressee, while simultaneously representing the I-narrator in the top-level Story-Viewpoint Space (for similar analyses, see Dancygier 2012a: 187–88, 2017b).

The notion of Story-Viewpoint Space was generalized in Dancygier and Vandelanotte (2016) into that of Discourse Viewpoint Space, regulating and supervising a networked configuration of viewpoints. This approach

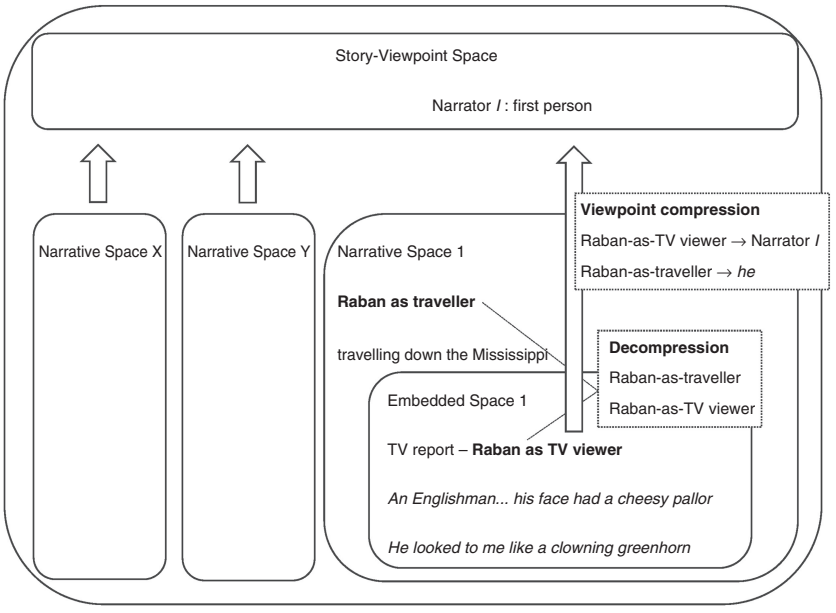

Figure 10.1 Decompression for viewpoint

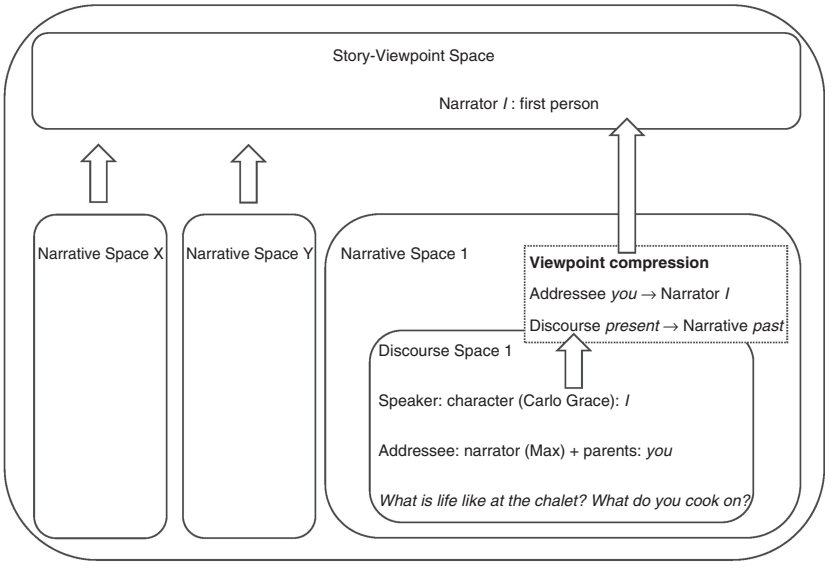

Figure 10.2 Viewpoint compression from Discourse Space to Story-Viewpoint Space

can help explain why and how, as a function of its place in the network, one and the same linguistic form (e.g. a first-person pronoun, an indefinite determiner) can have a different meaning. Across example sources such as film discourse, Internet memes, and video art, the analysis drives home the point that these often condensed discourse forms can show

considerable viewpoint complexity, which is clearly not the sole preserve of long forms such as novels. Consider, for instance, the popular *said no one ever* meme: while naturally understood to ridicule unlikely or untenable viewpoints (as in *Can I wear your Crocs? said no one ever*), it requires a zoom-out from its lowest level Discourse Space, in which the person depicted in the meme ostensibly asks someone (by default, the meme's reader) a straightforward question, to a distanced, higher-level space prompted by the 'said no one ever' part, usually presented visually as bottom text, from which the lower-level viewpoint is viewed and understood as 'non-quotative' direct speech, actually attributable to no one. The resulting Discourse Viewpoint arrived at is that Crocs are too ugly to be desirable.

The notion of zoom-out to achieve a view of a viewpoint was first proposed for the analysis of irony proposed by Tobin and Israel (2012). Irony on their account involves a layered configuration of mental spaces, where a perceived incongruity prompts a shift in attention from an inner to an outer layer (zooming out) resulting in a dynamic blended construal of an event from two incompatible viewpoints, one of which is rejected and 'looked down on.' Interestingly, sometimes the zoom-out does not have an unambiguous and stable nature, leaving the overall Discourse Viewpoint deliberately unclear. One fascinating example they discuss is the short story 'Borges and I,' written by José Luis Borges, where a kind of infinite regress of possibilities makes it impossible to pin down whether the overarching Discourse Viewpoint is that of Borges as public figure, or Borges as 'real' person.

Work in cognitive linguistic humor studies more generally incorporates an important viewpoint component (e.g. Ritchie 2006). To take one example, Brône's (2008) analysis of interactional humor in the British sitcom *Blackadder* relies on complex embedded viewpoints. In one scene involving misunderstanding on the part of dim-witted Lieutenant George, the soldiers are in no-man's land and George is looking at a map, guessing that the 'area marked with mushrooms' which he sees on the map means 'that we're in a field of mushrooms.' The following exchange between Captain Blackadder and Lieutenant George ensues:

(7) – Lieutenant, that is a military map, it is unlikely to list interesting flora and fungi. Look at the key and you'll discover that those mushrooms aren't for picking.
– Good Lord, you're quite right sir, it says "mine." So, these mushrooms must belong to the man who made the map.
– Either that, or we're in the middle of a mine-field. . .
– So, he owns the field as well? (Curtis et al. 1998: 362)

Figure 10.3 represents just one small aspect of the misunderstanding, that pertaining to the meaning of the mushroom symbol (for a more detailed analysis, see Brône 2008: 2036–42). As Figure 10.3 suggests, the viewer derives pleasure from seeing the whole picture from the Discourse Viewpoint Space, particularly the counterpart relations established across

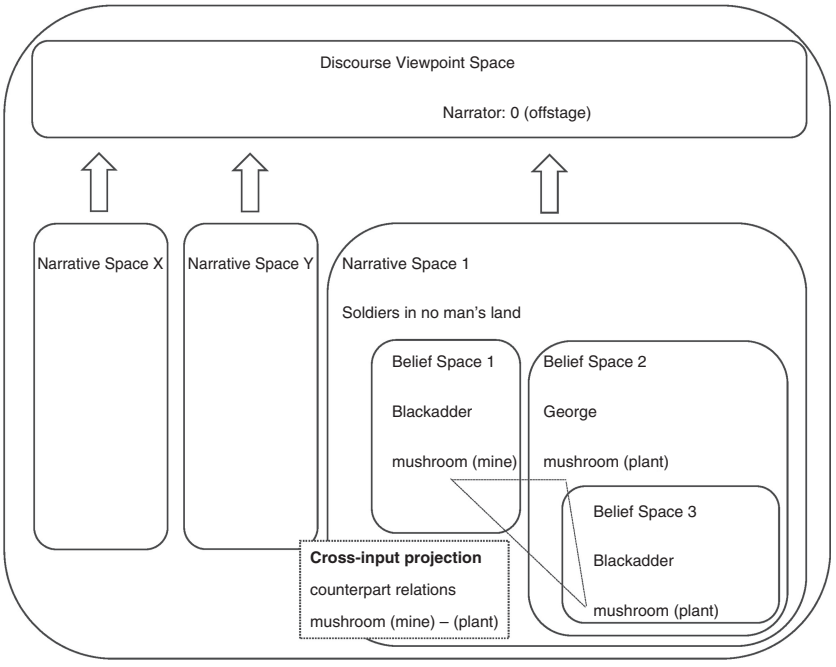

Figure 10.3 Embedded viewpoints with cross-input projection

the different Belief Space inputs: Blackadder's actual belief, George's mistaken belief, and, embedded within the latter, the belief which George wrongly attributes to Blackadder (given his lack of uptake following Blackadder's sarcastic follow-ups, he clearly believes Blackadder takes the mushrooms to represent mushrooms, not mines).

Irony and interactional humor are just some of the genres in which viewpoint mechanisms have been analyzed in mental space terms. Dancygier and Sweetser (2014: ch. 8), for instance, analyze discourses on illness or political discourse in which frames and blends are driven by viewpoint. The general idea of seeing discourse viewpoint as a matter of networked configurations of lower-level viewpoints, prompted by constructions, frames and blends, aims to capture the complexity of the phenomena by reconciling local multiplicity with the global coherence achieved at the level of the Discourse Viewpoint Space. In the closing section which follows, terminological and classificatory questions adhering to viewpoint are addressed.

10.3 Viewpoint Types and Relations to Other Concepts

The terminological diversity in viewpoint research can feel overwhelming and unhelpful. The following passage illustrates the problem:

> Every point of view belongs to a character, hence every focalization happens according to a character's perspective; similarly, every character's voice presupposes the existence of a focalizing perspective in which that character is vocalized. (Mey 1999: 149–50)

The potential for confusion is understandable, given the number and nature of linguistic and embodied construal choices which reveal a conceptualizer's viewpoint on smaller or larger aspects of a frame or situation. Certainly the narratological distinction between narrative voice ('who speaks?') and focalization ('who sees?' and, more broadly, 'who perceives?') articulated by Genette (1980, 1988) captures an important intuition, though as Niederhoff (2009a, 2009b) shows, it has spawned reifications (for instance by Bal 2009) capturing different aspects than those intended, and the issues remain hotly contested. For Genette, focalization replaces notions such as perspective or point of view, and pertains to "selections of narrative information" (Niederhoff 2009a: 122), distinguishing between unrestricted access to the storyworld (zero focalization), restriction to the experience of a character (internal focalization), or a further restriction to observable information only without access to characters' minds (external focalization). In cognitive narratology, Herman (2009) has attempted to overcome disagreements by proposing a theory of narratorial mediation in which narrating and perspective-taking (focalization) are intricately interconnected, given that "all storytelling acts are grounded in the perceptual-conceptual abilities of embodied human minds" (2009: 128).

Even if full narratological consensus on these issues is likely to remain elusive, a basic distinction along the lines suggested by Genette should be attempted. Given the choice to use viewpoint as the overarching concept, the distinction could be expressed in terms of *deictic viewpoint*, referring to the deictic center (conceptualizer, time, and place) from which some aspect of a situation is viewed, as opposed to various kinds of *cognitive viewpoint*, referring to whose cognitive states (thoughts, knowledge states, experiences, beliefs, attitudes, etc.) are being viewed, recognizing both the basic observation that various conceptualizers' cognitive states can be accessed from a constant deictic viewpoint and the possibility for the deictic viewpoint itself to shift (cf. Galbraith 1995), as for instance in direct speech reports. In example (8) below, for instance, the deictic viewpoint remains with the I-narrator (who is trying to block a woman from the view of her hawk, to avoid upsetting it), but the cognitive viewpoint shifts from the I-narrator to the woman's viewpoint of her ('a weirdo in a tattered jacket'), and subsequently blends with that of the hawk, a process completed in the final two clauses which compress the hawk's perception with the narrative viewpoint.

(8) I hold her close to my chest and turn in a slow circle to block the woman from view. The woman doesn't see the hawk. What she sees is a weirdo in a tattered jacket and baggy corduroy trousers revolving on the spot for no

> good reason . . . I've been with the hawk so long, just her and me, that I'm seeing my city through her eyes. She watches a woman throwing a ball to her dog on the grass, and I watch too, as baffled by what she's doing as the hawk is. I stare at traffic lights before I remember what they are. Bicycles are spinning mysteries of glittering metal. The buses going past are walls with wheels. (Macdonald 2014: 100–02)

We know of course that buses are not walls with wheels, but having assimilated the hawk's visual perception, the narrator conceptualizes them as precisely that (compare the notion of 'fictive vision' discussed in Dancygier 2012a: 102–06).

While arguably a useful distinction to make, the deictic–cognitive divide is not impermeable, and some forms straddle the boundary. Social deixis is a case in point: choice of address or reference form for people (for instance *Dirk* or *Professor Delabastita*) includes a deictic component (third person, not the speaker) but also a cognitive, social one, pertaining to the degree of familiarity to the speaker but also, as Sweetser (2013: 242) points out, intersubjectively, to the hearer (as I might use 'Professor Delabastita' in the presence of an undergraduate student, but 'Dirk' in private correspondence with my close colleague). Similarly, as suggested above, deictic elements such as *now* may be recruited to reflect an 'immersed,' present cognitive viewpoint in contexts where the overall deictic viewpoint may be past.

Things are thus rarely simple, and cognitive viewpoint in particular is by its very nature multifaceted, covering, for instance, phenomena that have been explicated in terms of experiential viewpoint (as in the viewpoint genitives discussed in section 10.2), grammatical aspect (which modulates conceptualization; Parrill, Bergen, and Lichtenstein 2013), and epistemic and emotional viewpoint. These latter notions have often been analyzed under the rubric of stance-taking (e.g. Du Bois 2007), and in Cognitive Linguistics have featured prominently in analyses of conditional constructions (e.g. Fillmore 1990, Dancygier and Sweetser 2005).

Broadly defined, along a general deictic–cognitive divide, as a conceptualizer's alignment with an aspect of a frame or situation, viewpoint is reflected in many linguistic and gestural constructions expressive of underlying conceptual viewpoint (Parrill 2012). This chapter has not sought to draw up an inventory of such constructions, since any such list would likely be incomplete, and unlikely to yield the kinds of generalizations sought here. However, descriptive work on such forms across different languages remains needed, while their applications across different modalities, domains, and genres are increasingly being tested and described using methodologies from across the cognitive humanities and sciences.