

Barbara Dancygier* and Lieven Vandelanotte*

Internet memes as multimodal constructions

<https://doi.org/10.1515/cog-2017-0074>

Received September 19, 2016; revised June 7, 2017; accepted June 9, 2017

Abstract: This paper considers a range of so-called image macro Internet memes and describes them as emerging multimodal constructions relying as much on image as on text, and apportioning roles to images much like constructional slots, for instance to fill in a subject role in a subjectless clause, or even to provide the main clause content to a textually given *when*-clause. In addition to existing or partially altered linguistic constructions, many examples also rely on specific top text/bottom text division of labor, and crucially depend on frame metonymy, with limited formal means quickly cueing richly detailed frames (for instance by using iconic images). The popularity of memes, forming series and cycles of iterations and remixes, and their role in establishing and maintaining discourse communities seems to be driven by a need to express and reconstrue viewpoints, often starting from ideas, affects or stereotypes assumed to be intersubjectively shared with viewers, whose responses they solicit. This paper argues that a proper description of Internet memes of the type considered requires a construction grammar approach, complemented by an understanding of viewpoint dynamics in terms of a Discourse Viewpoint Space regulating the network of spaces and viewpoints.

Keywords: constructions, frame metonymy, Internet memes, intersubjectivity, multimodality, viewpoint

1 Introduction

Internet discourse relies more and more pervasively on artifacts combining images and linguistic expressions to express and comment on viewpoints.

*Corresponding authors: Barbara Dancygier, Department of English, University of British Columbia, Vancouver, Canada, E-mail: barbara.dancygier@ubc.ca

Lieven Vandelanotte, English Unit, University of Namur, 5000 Namur, Belgium; Functional and Cognitive Linguistics: Grammar and Typology Research Group, KU Leuven, 3000 Leuven, Belgium, E-mail: lieven.vandelanotte@unamur.be

In this paper, we argue that Internet memes in particular offer important arguments for expanding the scope of cognitive linguistics and construction grammar to the study of their constructional and multimodal properties.

The *meme* concept finds its origin in the work of Richard Dawkins and his followers. Dawkins (1976) is primarily interested in replicability of genetic material, but he notices that culturally transmitted concepts also have a way of perpetuating themselves – hence the term “meme” as a complement to “gene”. The concept of memes covered various phenomena (melodies, catch-phrases, the habit of celebrating birthdays, etc.), and in fact language as a trait of human behavior has occasionally been seen as a meme (without the underlying claims of innateness, Merge, or universal grammar). The fact that Internet memes spread very quickly in online communication was the reason why these image/text combinations have been named memes. However, recent work by scholars including Shifman (2012; 2013; 2014) and Wiggins and Bowers (2015) has shown that biological analogies only partly apply to Internet memes, which tend to concern relatively short-lived fads and which involve considerable human agency (in the form of user-generated variations and remixes) compared to the kinds of units of culture Dawkins had in mind, which are transmitted largely unaltered and relatively passively, over long periods of time.

Among the broad variety of phenomena that come under the contemporary notion of Internet memes, as surveyed by Shifman in her chapter on meme genres (2014: Ch. 7), we focus particularly on (non-moving) images with superimposed text, sometimes referred to as *image macros*. These require comparatively little expert editing knowledge to create, and dedicated “meme generator” sites exist which make creating new iterations very easy indeed. The replicability and refashioning of Internet memes clearly drives their popularity as a means of communicating different, often humorous, viewpoints online.

In a single paper no comprehensive construction grammar theory and typology of Internet memes can be attempted. Instead, we analyze selected examples to demonstrate that construction grammar is the right framework from which to approach this emerging usage, and that the framework itself can, by the same token, usefully be expanded by taking into account the specific multimodal properties involved.

One thing that is needed is a better understanding of the different roles images play in their interaction with text in the overall meaning of memes. In the first example we will analyze below, that of *said no one ever* memes, for instance, the image – while relevant and tightly integrated in the meme’s setting up of mental spaces – is less in focus. At the other end of the spectrum, in the *when*-memes analyzed in Section 5, there would simply be no meme

(and no meaning) without the image, the mere text of the meme being incomplete or even ungrammatical on its own. In intermediate cases, discussed in Section 4, the meanings of memes with the same text can be reversed completely by changing the image from one stock character to another; in addition the images fill in varying essential roles not expressed in the linguistic form.

These interactions between linguistic and visual form are what make Internet memes *multimodal* in a specific sense. Within cognitive linguistics, multimodality has mostly been taken to involve the study, alongside the linguistic channel, of co-speech gesture, eye gaze, facial expressions, posture and possibly other kinesic/visual channels (see, e.g., Green 2014). However, we believe that “the other multimodality” (in Dancygier’s 2017 phrase), focusing on image and text (e.g., Kress and van Leeuwen 2001), urgently needs more detailed analysis in cognitive linguistics circles too (cf. Forceville and Urios-Aparisi 2009), given the ubiquity of contemporary forms of communication relying on novel combinations of both. In a similar vein, in proposing, as we do below, that Internet memes not only build on existing constructions but have unique, genre-specific constructional properties of their own, we are proposing “another” multimodal construction type compared to the pairings of physically embodied and verbally expressed components of the kind reviewed and exemplified in Brône et al. (2017), as in the multimodal “package” (McNeill 2006) of verbal fillers, gaze aversion and turn-holding they describe. As some of our examples will show, once established as multimodal constructions of this “other” type, memes can additionally give rise to new linguistic (monomodal) constructions appearing in more standard usage contexts such as journalism and advertising.

We thus want to make the case that Internet memes are emerging pairings of form – specifically multimodal configurations of forms – with identifiable meanings, thereby qualifying as *constructions* in various understandings of construction grammar (e.g., Langacker 1987; Fillmore 1988; Fillmore et al. 1988; Goldberg 1995; Croft 2001). One aspect which our analysis will highlight is that constructional meaning can be signaled even when some of the formal features of the full construction are missing. This is a particular constructional version of frame metonymy, by which characteristic parts of a frame are sufficient to call up whole frames. In like manner, as Dancygier and Sweetser (2005) have argued, selected, characteristic constructional features may be salient enough to prompt broader frames of constructional meaning; they refer to this as constructional compositionality, and elaborate the example of various coordinated constructions which share sufficient skeletal structure with predictive conditional constructions to share the latter’s meaning

(e.g., *Raise your voice again and I quit* or even the (partly) verbless *Three strikes and you're out*). Such reduced predictive constructions are also discussed below in the memetic context.

In addition to making the case for a constructional approach to Internet memes, our analysis stresses the importance of *viewpoint*, with memes relying on pre-existing attitudes and beliefs, and shifting or manipulating viewpoint to achieve a new viewpointed construal, often an ironic one. Two interrelated aspects stand out in particular. One, the viewpoints expressed take into account and respond to attitudes, beliefs, stereotypes, clichés and the like assumed to be shared by or at least known by the addressee, and are thus inherently intersubjectively established in the sense of Verhagen (2005). Second, memes tend to form chains of successive responses and refashionings, recycling initial combinations to refer to new developments, current events, fashions, fads, and the like. In other words, Internet memes cannot be studied in isolation, as individual artifacts. They participate in complex networks not only of previous examples of the same meme and new blended combinations of existing memes, but also of various non-memetic artifacts (photos, films, iconic figures, ads, etc.). These artifacts are well-known in a given discourse community and provide instant access to rich frames whose contribution to the emerging meaning is central. The final interpretation which emerges out of this network, we argue, can best be understood as being resolved at the level of a supervisory mental space which we have labeled the Discourse Viewpoint Space (Dancygier and Vandelanotte 2016).

We propose to put flesh on these bones of our argument to treat Internet memes as viewpoint-driven multimodal constructions as follows. In Section 2, we focus on the *said no one ever* meme for its constructional and viewpoint interest – relying on but reconfiguring existing viewpoint constructions. The degree of multimodal interaction, while present, is still modest in this example. The importance of the visual component increases considerably in the *One does not simply* meme to which we turn in Section 3, as it helps to evoke a detailed frame without which the meaning of the meme – to categorize some activity as a futile undertaking – would be very unclear. Section 4 turns to *Good Girl Gina, Scumbag Steve* and related memes, where the meme characters provide their own framing, but also fill in particular constructional roles left unexpressed linguistically, and where the reliance on familiar stereotypes about male and (especially) female behavior is exploited intersubjectively in subtle ways. The *when*-memes looked at in Section 5 feature a very prominent role for the images in complementing an incomplete linguistic construction in order to construe a simile multimodally. Section 6 concludes with an overview of the constructional, multimodal and viewpoint phenomena highlighted by the cases studied in this paper, and opens up further questions left to be explored.

2 *Said no one ever: Speech representation and viewpoint reversal*

In this section we consider the *said no one ever* meme, which relies on but modifies the direct speech construction. Constructions of speech and thought representation (STR) are varied, and display many grammatical parameters connected to a range of viewpoint phenomena (see, e.g., Sanders and Redeker 1996; Vandelanotte 2009). They all profile a higher-level telling/thinking space, aligned with a narrator or another kind of teller, and an embedded space representing the words or thoughts, aligned with the narrator herself or with another speaker/character. In the simplest case, Direct STR, the (assumed) words spoken or thought are represented from the deictic perspective of the actual speaker/thinker, as in (1):

- (1) “Your suggestion really made me change my mind”, said Joe.

The narrator is here reporting Joe’s statement, presumably verbatim, though this assumption is only a matter of convention; more often than not, people’s words are not actually remembered nor represented verbatim (see, e.g., Vandelanotte 2009: 118–123).

The direct speech construction is the foundation of the *said no one ever* meme, first discussed in Dancygier and Vandelanotte (2016: 28–34), which uses this constructional formula, but violates the assumption that there is a specific speaker whose words are being represented, as in (2) and Figure 1¹:

- (2) “Your Facebook status really made me change my political views.” Said no one ever.

The meme thus starts with a sentence, often given in quotes, to maintain the pretense of an actual utterance by an actual speaker. The image representing a character in the meme supports this initial impression of being presented with a genuine quoted clause. Indeed, some versions of the meme feature not drawn characters as in Figure 1, but photos of actual people; one amusing example shows a fearsome threesome of unimpressionable hipster girls accompanied by the text *Can I wear your crocs (Said no one ever)*. In the meme, the apparently

¹ The memes which we use as examples can all easily be found via Internet search engines. Meme databases such as knowyourmeme.com and memebase.cheezburger.com compile informal usage histories and selected examples of different memes, which readers might find helpful.

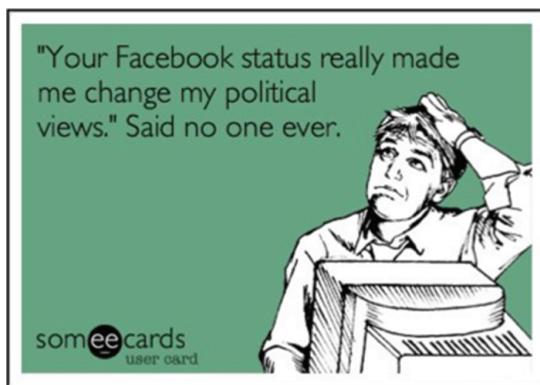


Figure 1: The *Said no one ever* meme.

quoted statement is often naïve or improbable, but not in itself absurd, so it could have been genuine. The added *said no one ever* clause is, thus, what finally and decisively has the effect of a complete viewpoint reversal – the clause in quotes is now not a quotation, but a token of an unlikely opinion. Even though the first and second person pronouns used suggest, by default, a deictic center and spoken interaction, the added clause fails to embed the discourse in any conversational context. The construction which denies the possibility of any speaker using the faux-reported clause in earnest changes the viewpoint of the reporting clause from ‘containing reported discourse’ to ‘containing impossible discourse’, and as such joins the ranks of constructions identified crosslinguistically as “non-quotative” uses of direct speech (Pascual 2014: Ch. 4). However, in order to appreciate the viewpoint shift, the hearer/viewer needs to construct a higher-level discourse space which allows him to reconcile the two viewpoints – a report and the denial of report. The final interpretation, such that reasonable people are not expected to be swayed by opinions expressed on Facebook, can only be arrived at in the course of that resolution.

We propose to analyze cases such as (2) in terms of a multi-level viewpoint structure, represented in Figure 2. The box on the left represents the utterance as it could be initially understood, i.e., as spoken by the character depicted in the meme (MC) in what we label Discourse Space 1. The addressee here could be taken to be the meme maker in the meme-making space, if we analyze the meme maker’s process; if we analyze the meme viewer’s interpretive process, the addressee is the meme viewer (cf. Rohrer 2005). This space is then embedded (the big arrow represents the process) in Discourse Space 2 by means of the

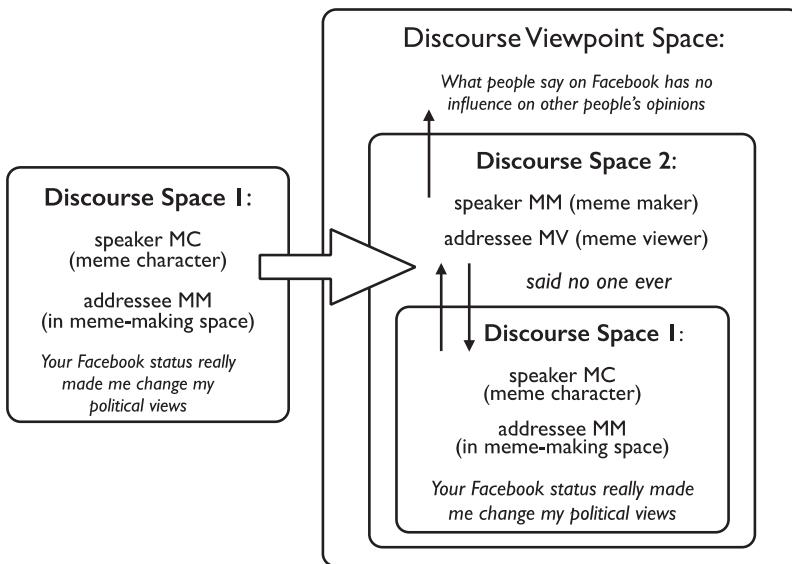


Figure 2: Viewpoint in a typical *Said no one ever* meme.

"faux" reporting clause (*said no one ever*), producing a clash between the initially assumed claim and the *said no one ever* faux-reporting clause, indicating that the apparent "quote" is in fact addressed by no one to no one. In order to meaningfully resolve this incongruous clash, a unifying higher space is set up, represented by the outer box, which we have labeled the Discourse Viewpoint Space (Dancygier and Vandelanotte 2016). As we argued there (2016), viewpoints in discourses of different sizes are best thought of as organized in networks rather than as hierarchical lists, and these networks can take on many different forms but are all governed or supervised by a top-level, unifying viewpoint space where the different lower-level viewpoints are reconciled and understood in final interpretation. The small black arrows indicate how the faux-direct statement is affected by being embedded in the *said no one ever* clause, and then re-interpreted by seeking coherence at the higher (Discourse Viewpoint Space) level.

The solution of a Discourse Viewpoint Space we use is similar to the zoom-out effect proposed by Tobin and Israel (2012) to explain the emergence of irony. In the case of irony, the resolution of incongruent viewpoints also requires an independent viewpoint level where a lower-level viewpoint is re-evaluated. However, the incongruity in the case of irony is not explicitly, constructionally present in the way it is in (2), due to the presence of *said no*

one ever. Indeed, when someone comments *What a beautiful day!* in the middle of a torrential downpour, it is contextual factors which require a zoom-out to a higher viewpoint in order for this incongruity to achieve relevance. Each in their different ways, the networks involved in our later examples too constructionally prompt for resolution in a Discourse Viewpoint Space, which is in turn related quite broadly to the concept of intersubjectivity – coordination of construals between the speaker and the hearer, in the course of discourse negotiation.

Figure 2 and our later figures do not represent the meaning emergence process formally or exhaustively, but signal the configuration of spaces and viewpoints enabling a final discourse viewpoint to emerge. In this and later figures, it is also important to point out that the various types of text or discourse spaces in memes involve different participants as speakers, addressees or other agents: the meme maker, the character represented, etc. This is a central aspect of meme use, since the profiling of different conceptualizers is aligned with profiling of different viewpoints, and viewpoint is the driving force behind exchanging, commenting on and responding to memes.

While the discourse viewpoint reached in (2) may seem quite generally shared – presumably some people would agree that Facebook posts do not fundamentally influence people's political views much – the construction does not in fact rely on pre-conceived assumptions. In another example, *I just don't find British accents attractive (said no one ever)*, the final interpretation assumes that everyone finds British accents attractive, since the thought that someone would not do so is being presented as too outlandish ever to be uttered by anyone. However, accents are very much a matter of individual tastes, so there is little ground for assuming that the discourse viewpoint arrived at here is a broadly shared belief. And yet, remarkably, the construction here re-constitutes a possible belief into an opposite one, held intersubjectively. By virtue of the *said no one ever* construction, not only has the meme maker made it clear that the thought that British accents might not be attractive is too ridiculous to contemplate, he or she has also presented this as a position shared (or now to be shared) by the meme viewers. The degree to which memetic views are indeed shared, or just forceful, cannot be evaluated in a reliable way. But it is clear that there is an assumption of a degree of intersubjectively shared views (where intersubjectivity is understood in terms of Verhagen's (2005) discussion of constructions of intersubjectivity).

Further variations of the *said no one ever* meme reveal other aspects of interest to our argument. In terms of intersubjectively negotiating beliefs and attitudes, the variant which replaces *no one* by a *no + noun* NP labelling a more specific subset of people works as stereotype-reinforcing, as in the example "*I*

can't wait for class to start" said no student ever (cf. Dancygier and Vandelanotte 2016: 32–34). Forms which condense the *said no one ever* phrase to just "no one ever" or even just "no one" illustrate the constructional compositionality discussed in the Introduction, i.e., the possibility for salient parts of constructions to be sufficient to call up the full constructional meaning.

As we have suggested above, the role of the image in the *said no one ever* meme supports the initial garden path onto which the viewer is briefly led, in creating the initial default suggestion that a speaker is genuinely being quoted. Interestingly, and perhaps partly because the multimodal nature of this meme is not as strong as in the other cases we will turn to, the meme appears to have given rise to a novel extra-memetic construction in more standard usage. In an editorial article published on May 25, 2016, the Canadian newspaper *The Globe and Mail* comments on the prime minister who has lost the recent election. He was generally not liked, and the editorial talks about the lack of nostalgia voters experience. The first sentence of the article is: "The announcement on Wednesday of Stephen Harper's imminent retirement from politics caused a massive outpouring of emotion, said no one ever". Here, the so-far multimodal form of expression is used as a linguistic construction, appropriate to journalistic prose. The viewpoint effect is very similar – the spurious discourse that nobody is likely to produce precedes the faux-reporting clause which in effect confirms the view that this particular politician will not be missed as a person. What is interesting to note is how the form originating in a multimodal context is coopted into journalistic prose. It is likely that we will start seeing these effects more often; we will in fact encounter a few similar examples in later sections. In the next section, we first turn to a class of examples in which the background image provides a frame which is essential to the overall meaning of the meme.

3 *One does not simply: Frames and categorization*

A typical image macro meme consists of three elements²: a background image, top text (TT), which is often formulaic and easily recognizable, and bottom text (BT), which often delivers the punch line of the meme. In many memes, the background image is constant, paired with TT. For example, there is a popular meme known as *One does not simply*. The background image is a screenshot

² The *said no one ever* meme discussed in Section 2 is not such a typical image macro meme, especially since there is no constant background image. While its text does contain two clearly distinguishable parts, it is often not presented in a clearly distinguished way (top vs. bottom).

from the movie *The Lord of the Rings*, showing the troubled face of Boromir (Sean Bean). At least for meme viewers who recognize the film still, as one assumes the original receivers of the first such memes would have, the image thus calls up a frame filled with a lot of quite specific knowledge. In particular, in the scene represented this way, characters are deciding on the best course of action, especially the idea of walking into the deadly land of Mordor. In the discussion, Boromir says: “One does not simply walk into Mordor”. The phrase *One does not simply* along with the image has come to metonymically evoke the scene, but also functions as a memetic ‘prefix’ to a description, in the bottom text (BT), of any proposed course of action which the speaker deems destined to fail. The image does not change, TT remains the same, while BT represents the specific undertaking that cannot succeed: from a simple observation about cleaning up glitter (*One does not simply clean up glitter*) to a more serious point about saving Africa by donating \$1 (Figure 3).



Figure 3: The *One does not simply* meme.

One does not simply memes are very popular, and the rich Mordor frame from *The Lord of the Rings* has come to function as a tool for categorization: new instances of ‘futile undertakings’ seem to be being added to the growing category almost daily. The formal stability of the textual component structure *One does not simply do X* is such that it has in fact been picked at by self-reflexive memes drawing attention to this fixity: in one such meme, the top text is changed to *One not does simply ...*, and the bottom text then comments ... *notice the words are switched*. It is important to stress that the *One does not simply* phrase in itself is not enough to determine the meaning of the overall construction; an ordinary image-less use of this clause could have other meanings, for

instance to point out that something is improper (*One does not simply barge into a meeting like that*). The meme construction is thus built multimodally, image and text together, to produce the understanding of categorizing undertakings as futile and bound to fail.

However, it is also built intersubjectively. In an example such as *One does not simply ... save Africa by donating \$1*, the meme naturally evokes various campaigns which urge people to donate, arguing that even a one dollar donation makes a difference. The viewer of the meme needs to be able to evoke such instances, and perhaps also think of his or her own trust in supporting worthy causes with micro-payments. To process the meme, one needs to rely on the Discourse Viewpoint Space to recall appropriate instances and one's own opinion. Figure 4 represents the configuration: the initial constructional framing including top text and the image macro, the embedding of bottom text, relying on frame knowledge of charities and fundraising, and the resulting understanding of small charity donations as futile.

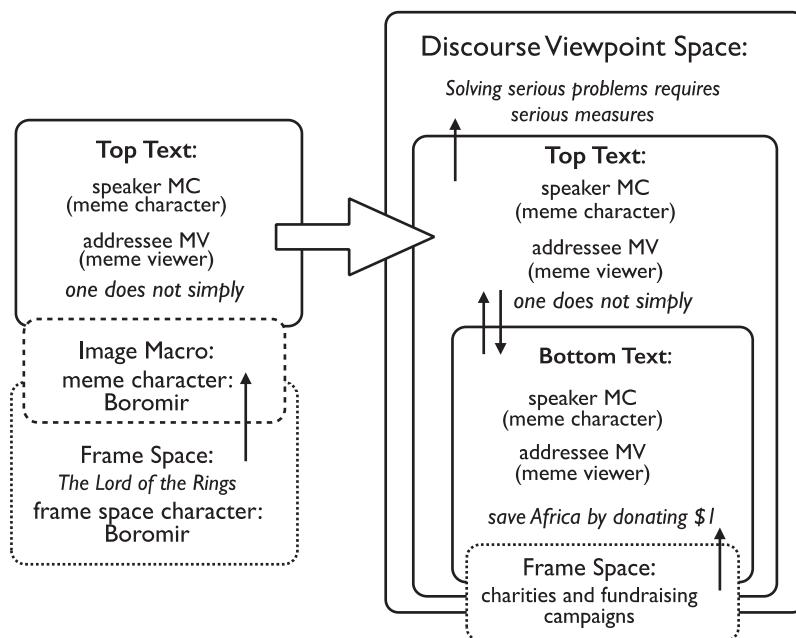


Figure 4: Viewpoint and frame configuration in a *One does not simply* meme.

Other examples build on much more pervasive viewpoints, especially when male and female stereotypes are evoked. In one example (Figure 5), the text



Figure 5: The *One does not simply* meme and stereotyping.

says: *One does not simply ... win an argument with a woman*. In another, non-memetic, context, a sentence such as *I can never win an argument with my sister* might simply mean that the speaker acknowledges his/her sister's power of argumentation, so it could be an expression of admiration. In the meme, which uses a male face and constructionally describes 'something bound to fail', what is being evoked is a sexist stereotype of women as people not listening to reason and driven by emotions. That stereotypical viewpoint is accessed intersubjectively, playing on known assumptions and prejudices, as a sort of 'wink' at the male viewers. Female viewers will most likely respond differently.

As we saw with *said no one ever*, what starts as a multimodal meme can live on in more standard usage without the accompanying meme conventions – background image, TT, BT and even font choice (cf. Brideau and Berrett 2014). Indeed, it is possible to find examples of *One does not simply* which arguably draw their meaning from the full multimodal construction, with its specific meaning of categorizing something as a futile undertaking. Thus, in a 2017 online article about "movies or shows that are only available on Netflix Canada", appropriately under the heading *The Lords of the Rings trilogy* we read "One does not simply fire up Netflix in the U.S. and expect to watch Peter Jackson's epic fantasy about two small men trying to destroy jewelry".³ Similarly, in a 2013 article about "9 stunning next-gen airport designs cleared for takeoff", the section on Rock Terminal at Wellington International Airport starts

³ Source: <http://www.fortmcmurraytoday.com/2017/04/17/shows-that-are-only-available-on-netflix-canada>.

with “One does not simply walk out of New Zealand”.⁴ While of course *One does not simply* could be used, purely textually, before the meme was first created, with a broader range of contextual meanings as suggested above, here given the contexts of use of these examples, it seems clear the meme is being evoked, without its full array of formal features (image, text arrangement, font, etc.).

The *One does not simply* meme uses TT for its signature phrase, but there are other memes where BT, not TT, is used that way. For example, the Kermit meme, featuring the image of Kermit the Frog drinking tea, has the signature BT line *but that's none of my business*. In one example, the TT says: *You offer relationship advice to people but you are single*. In a typical case such as this, TT is a critical or at least puzzled remark on people's behavior, and BT plays the role of a sort of ‘suit yourself’ attitude. Interestingly, this usually makes TT much longer than what you normally expect in a meme (as illustrated in Figure 6), making it less snappy and perhaps lessening the ‘constructional feel’ of the whole expression. Nevertheless, the overall pattern remains stable through a range of examples, with the image and BT providing a recognizable template ready for use in various contexts.

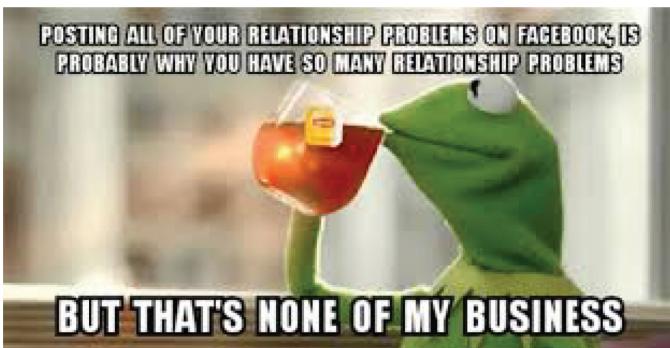


Figure 6: A *But that's none of my business* meme with very long top text.

4 *Good Girl Gina and friends: Conjunctionless predictive constructions and argument suppression*

While the memes discussed so far rely among other things on a set phrase (*said no one ever*, *One does not simply*, *but that's none of my business*), in this section

⁴ Source: <https://www.wired.com/2013/02/airport-architecture/>.

we turn to a family of memes which do not, and where the relationship between TT and BT is more complex, and resembles temporal or conditional clause relationships. In particular, we focus on a productive group of memes focusing on good, considerate vs. bad, inconsiderate behavior against a background of male and female stereotypes. The way in which the image functions in these memes is different again from the cases treated in previous sections: instead of an apparently random possible speaker (*said no one ever*) or a highly specific frame-evoking image (*One does not simply*), the image here represents a stock character who comes, so to speak, with built-in framing. Consider, for instance, Good Girl Gina (GGG), as in Figure 7. GGG is not a famous actress from a film or commercial, not a politician or some other well-known figure: at heart she is a noble unknown, but she represents a perfect girlfriend and her familiar picture resurfaces in each iteration of the GGG meme. Every such meme has the TT located at the upper edge of the photo, introducing a type of behavior which GGG might engage in, and the BT, located at the lower edge, indicating how GGG's conduct in the TT situation confirms her being a "good girl". For example, in Figure 7, when TT says *Gets mad at you*, BT then completes the construction: *tells you why*.



Figure 7: The *Good Girl Gina* meme.

The construction is the closest to a conjunctionless variant of a conditional or temporal construction, with a generic meaning, signaled by the habitual present tense (*If/When/Whenever/Every time GGG gets mad at you, she tells you why*), which is an example of a predictive construction, as described in Dancygier (1998) and Dancygier and Sweetser (2005). The characteristic features which the GGG meme shares with full predictive constructions (including the biclausal structure and the habitual present tense) appear sufficient to give access to

the full constructional meaning, in the frame-metonymic mechanism of what we discussed as constructional compositionality in the Introduction.

Perhaps the most interesting formal feature is the suppression of the subject argument. This happens predictably, in all GGG and related memes, and so is constructionally determined. David (2016) explains argument suppression examples through frame-analysis; we argue here that the constructional frame of a predictive sentence in combination with the image-macro yields the possibility of grammatical suppression of the subject. Specifically, the subject suppression here is an example of what Ruppenhofer and Michaelis (2010) have labeled “genre-based” argument omission, which they identified in genres such as instructional imperatives as used in food recipes, the language of labels, and diary style, and which they argue has obligatorily anaphoric interpretation: “the omitted arguments (...) are entities that are not only mutually identifiable to speaker and hearer but also a current joint focus of attention” (2010: 164). Thus, in food recipes the food item is in focus; in labelese the product being described is; and in diary style the diarist is. What is peculiar to the meme genre we are adding to the discussion here is the multimodal nature of the recoverability of the omitted agent/subject which is in focus. Indeed, the discourse which maintains the identity of the subject is the image, not a linguistic form; in effect, the meme says: *When GGG gets angry, she tells you why*, and the reference to GGG is provided by the familiar photo in the background of the meme.

As in other cases, the consistency of the pattern – the repeated image, the division of labor between TT and BT, the reduced, conjunctionless form of the construction, and the subject omission – is what gives the GGG meme a constructional status. In addition to Good Girl Gina, there are also series of memes describing Good Guy Greg, Scumbag Steve, and Scumbag Stacy, though for some reason the latter group is very small (perhaps girls aren’t easily thought of as “scumbags”, which might be a term of abuse used characteristically of male targets). Each of the four series works from its specific image, and uses generic predictive constructions, with suppressed subjects, as ways to define the categories of ‘perfect girl/boy’ and ‘awful girl/boy’. In each of the meme series, the names of the characters also function as the names of the memes, and at least proficient meme users and viewers would be aware of these names, whose form is also consistent in using alliteration. As Geeraerts and Zenner (2016) have pointed out, some memes require knowledge of the meme’s name to really be interpretable, so they propose to include a meme’s name as one of its constructional components. Examples include “Socially Awkward Penguin,” in which BT provides an awkward response to a situation, or “Anti-Joke Chicken,” in which BT provides an anticlimactic non-joke. The effect is arguably somewhat weaker

here, but knowing the name does provide a clearer sense of the (good vs. scumbag) frame evoked by the characters than merely relying on the looks of the characters – even if the looks do provide clues, with for instance Good Girl Gina looking like a dream girl and Good Guy Greg like a friendly chap who is enjoying himself.⁵

In one interesting case, the image has taken on a significance beyond filling in the suppressed subject, also making possible a striking metonymic pattern. The case in point is that of Scumbag Steve, who is always represented by a photo in which he is wearing a characteristically patterned baseball hat (see Figure 8). The object distinguishes him and so has become known in the meme world as the Scumbag Hat. Meme-making sites have the Scumbag Hat feature available independently of the meme-making formulae, so that a Scumbag Hat can be put on anything, in any image. For example, there is a whole series of memes featuring a Scumbag Brain, complaining about tricks our minds play on us, related to memory, language learning, face recognition, etc. The hat shown in the Scumbag Steve memes has thus come to



Figure 8: The *Scumbag Steve* meme.

⁵ An anonymous referee points out that the names (Gina, Greg, Steve, Stacy) are reminiscent of names used in children's literature and in TV public service announcements in the US in the 1950s, in which the message was essentially 'look at the good behavior of X; be like X' and 'look at the bad behavior of Y; don't be like Y'. As modern equivalents of these normative forms of communication, the memes both produce or reinforce stereotypes of their own, but also intersubjectively criticize this type of normative mentality.

metonymically evoke ‘untrustworthy character’, and does so on its own, in any context (though perhaps most typically it is put on the heads of politicians). The hat on its own seems thus to have become a meaningful form – a subconstruction in its own right, not unlike previously unanalyzed parts of words taking on morphemic status, as in the case of *-holic* formations deriving from *alcoholic*, such as *shopaholic* or *chocaholic*. Functionally, the Scumbag Hat as an independent operator seems to overlay an existing representation with a subjective meaning, modifying the whole meaning of the representation, not unlike grammatical operators. This is one particular case where thinking of memes as multimodal constructions is beginning to allow us to see deeper parallels than might have been expected.

As we signaled earlier in the discussion of the GGG meme, the TT and BT of these memes have interesting formal features. They use a conjunctionless generic predictive construction (with P as a temporal or conditional clause and Q as the main clause) such as *Uses shopping cart ... leaves it in empty parking space*, without subject NPs and also, as in newspaper headlines, without articles, thereby further condensing the linguistic form. Each such meme starts with a P clause describing a common and unremarkable event, and then uses BT Q clause to show an example of behavior which can unambiguously be judged as good or bad. The constructional and viewpoint configuration is represented in Figure 9. While the viewpoint structure is

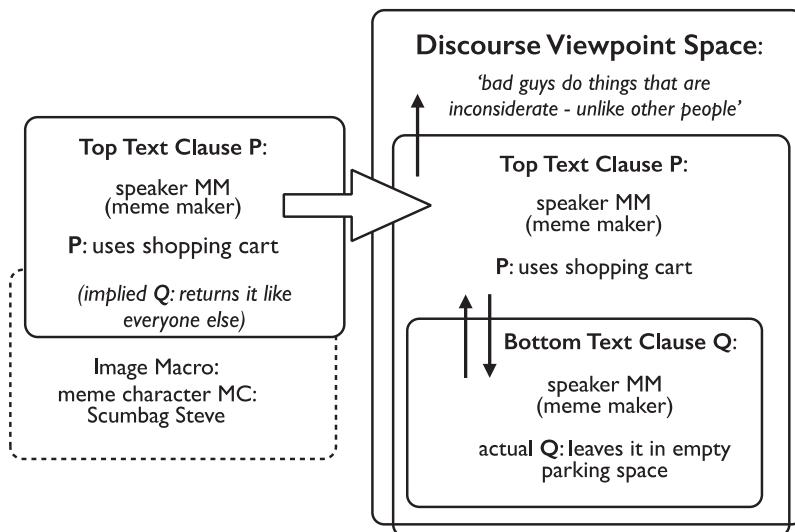


Figure 9: Viewpoint in a typical *Scumbag Steve* meme.

again best understood in terms of a network, as we argued for the examples in Sections 2 and 3 above, the constructional pattern represented is quite different. The diagram also adds the viewpoint role of the image, recognized as a representation of an inconsiderate, neglectful, ‘bad’ person, which is needed to help resolve, in the Discourse Viewpoint Space, the countering of expectations: rather than a predictable sort of Q (‘returns it like everyone else’), we get an instance of irresponsible and inconsiderate behavior in the actual Q (“leaves it in an empty parking space”).

Figure 10 presents a more complex variation on this predictive pattern. Here, BT is not straightforwardly the main clause of a predictive construction, but is an utterance, in quotation marks, which we resolve as representing the character’s response to the situation. For example, TT in Figure 10, a Scumbag Steve meme, reads *Breaks something expensive of yours* and is followed by the BT line “*Why would you spend so much on it anyway?*”, suggesting the defensive response on Scumbag Steve’s part when being confronted over the breakage.



Figure 10: *Scumbag Steve* and quotation.

In such cases the multimodal configuration is quite complex. On the one hand, as in the previous examples discussed in this section, the image provides referential material to fill in the suppressed subject in the first clause given in the TT line. On the other hand, the role of the person represented in the image changes in the BT line – Scumbag Steve is now the speaker of the question in quotation marks, at least fictively (in the sense

of Pascual 2002; 2014), with the quote providing a typical, dismissive response, refusing to take responsibility. Constructionally, the role of the BT line is thus quite complex. While it takes the form of direct speech, there is in fact no linguistically signaled deictic center, with a speaker of these words, the addressee, and the here and now. Clearly, no specific discourse situation is available on the basis of the meme as such, so in addition to imagining the narrative events described in the way the previous examples did, here the meme viewer also imagines a speech exchange taking place between Scumbag Steve and the owner of the expensive object, who we might label Scumbag Steve's friend (SSF), as a concrete token of the generic 'yours' (*something expensive of yours*). We imagine that SSF complains about the loss, but this discourse is not included in the meme's text. What we do get, in the BT line, is an imagined retort, which counts as Scumbag Steve's rude response to the complaint. On the basis of very minimal cueing – essentially, the use of quotation marks – to process the BT we fill in quite a detailed scene, including SSF's previous turn in the exchange which is not represented in the meme.

To complete the structure of the meme text in the way we did for the earlier examples, we would need to fill in not just conjunctions and the subject of the first clause, but a lead-in setting up the exchange, and a reporting clause to clarify the quote (*If/When Scumbag Steve breaks something expensive of yours, and you complain to him about it, then he'll respond by saying "Why would you spend that much on it anyway?"*). Constructional compositionality, frame metonymy more generally, and the multimodal interaction between image and text are all integral to fully access this generic, predictive understanding in spite of the strongly reduced verbal form, and to resolving its point at the level of the Discourse Viewpoint Space. The analysis is represented in Figure 11, which resembles Figure 9, except in the complex way in which the Q element is construed, with Scumbag Steve as speaker and SSF and, secondarily, meme viewer (MV), as addressee.

What complicates this example compared to that in Figure 8, then, is the speaker switch effected in the BT line, something which seems strictly speaking ungrammatical without proper linguistic marking ('then he'll respond'), but which respects the schema of the overall construction – TT sets up a situation, BT presents the consequence of that situation. As meme viewers, we understand Scumbag Steve's addressee in the virtual, fictive interaction prompted by the BT line to be an imagined, possible specific instance of the generic 'you' in the TT line (*something expensive of yours*). Fundamentally, of course, Scumbag Steve is also an imagined speaker: the meme invites us to think of a category of inconsiderate, rude brutes, of which we may know

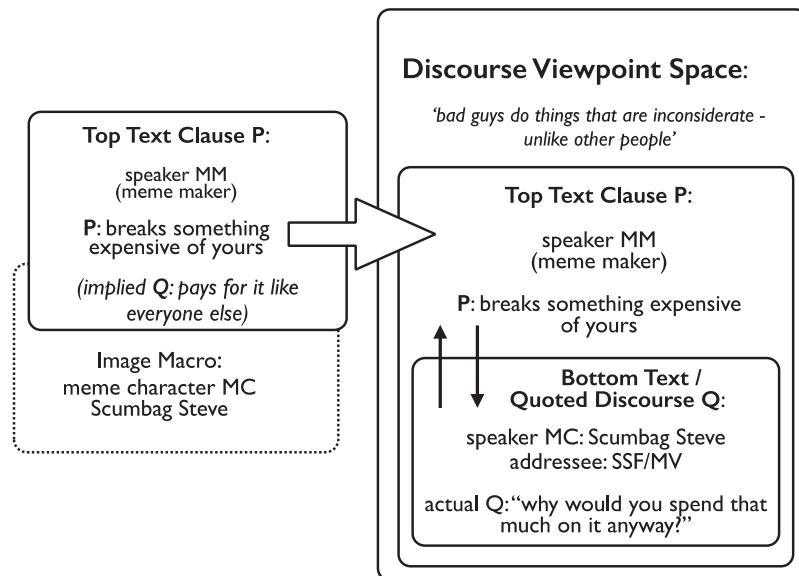


Figure 11: Viewpoint in a *Scumbag Steve* meme involving quotation.

specific examples in our real day-to-day existence, so we can naturally empathize with the role we labeled as that of SSF – the victim of Scumbag Steve’s appalling behavior.

The use of the quotation in the BT of this example is thus particularly effective. This ties in with the understanding of quotation generally as involving “demonstration” (Clark and Gerrig 1990) or “performed depiction” (Clark 2016), irrespective of the question whether any actual person actually uttered the quote. Here indeed, the BT clause demonstrates and typifies a rude person’s behavior. We could further refer to such usage as an instance of “fictive interaction” (Pascual 2002; 2014). This is speech which represents an attitude, not an actual exchange, so the general definition of fictive interaction applies: a fictive quote is used in order to typify characteristic behavior. However, it is questionable whether this example may be claimed not to involve a proper exchange but only typified behavior. In fact, we could argue, the meme’s BT does have a specific speaker, provided by the image, and it occurs in the context of an assumed conversation within the ‘object breaking’ narrative space set up by the construction. At a minimum, we can say that multimodal discourse complicates not only formal linguistic categories, such as the expression of the subject or the form of the construction, but also usage-related ones, such as fictive interaction.

The examples in this section also raise important questions related to intersubjective coordination and viewpoint in relation to gender stereotypes. One can ask, for example, how it makes sense that only an ideal girlfriend tells you why she is angry. Clearly, the GGG memes start with male assumptions about women's behavior that men find annoying, for instance the assumption that girls act offended without explaining why on the grounds that the boyfriend would guess the reason if he tried. The belief about women (held by men, but easily recognized by women) is what the meme works from. In effect, these memes are 'man-talk', describing a perfect girlfriend by contrasting them favorably with stereotypical undesirable behavior on the part of women (e.g., getting mad at you for no apparent reason). Similar stereotypes are also the basis of many of the Good Guy Greg, Scumbag Steve and Scumbag Stacy memes.

In each of these memes, as we suggested for the cases in previous sections too, the interpretation needs to involve a level wherein all the assumptions about behavior are reconciled and the meme maker's viewpoint becomes acknowledged. 'Being angry and not explaining why' is a behavior which the meme viewer will evaluate from their perspective, and possibly the reading of the meme will be different depending on whether the viewer is male or female. Reaching intersubjectively available viewpoints not expressed in the meme itself and reconciling them with the meme maker's standpoint and one's own requires a bird's eye view of the network from what we call the Discourse Viewpoint Space. The emergence of such a space could be talked about in terms of blending (Fauconnier and Turner 2002), but the primary mechanism we are relying on here is viewpoint compression (as discussed in Dancygier 2012). Since viewpoint emergence is our main concern here, we restrict our focus to that.

As a final example in this section, we need to consider the composite meme in Figure 12, which emphasizes the extent to which the image in these memes becomes fully representative of character. In Figure 12, the same TT and BT is used for both Good Guy Greg and Scumbag Steve: *Sees that you're drunk ... takes your keys*. Quite naturally, the two matched memes are interpreted differently in each case, based on the "Good Guy" or "Scumbag" framing provided by the image. Thus, the Good Guy scenario might involve Greg taking your keys to drive you home safely, while in the Scumbag scenario Steve takes them to get you into trouble, perhaps even steal your car. Neither scenario is suggested by the meme's text, but the image prompts the filling in of the good or bad intentions. While this effect is visually, not verbally, achieved, it can be likened to linguistic phenomena such as positive vs. negative semantic prosody (e.g., Sinclair 1991; Louw 1993), as when *provide work* is positive but *cause work* negative



Figure 12: *Good Guy Greg* meets *Scumbag Steve*.

(Stubbs 1995), or positive vs. negative polarity items (e.g., Israel 2011), as when *would you like some help* is oriented more towards a positive response than the more neutral *would you like any help*. Here again, the meaning of the meme is arrived at multimodally and accessed intersubjectively, responding to shared beliefs about types of behavior.

5 *When-memes: Missing clauses, simile, and metonymy*

One of the most popular types of meme uses the conjunction *when* as the means to set up a narrative space. There are various subtypes of the meme (see Lou 2017), but here we want to look at one subtype, because of its interesting constructional effect. In this type, the top text starts the *when*-clause, and the clause either fits into TT in its entirety or is split between TT and BT, but the expected main clause describing the expected outcome of the *when*-situation is not provided at all. In one such example, the text is *When you're about to leave work and the boss says "before you go ..."* (Figure 13); in another the TT is *When your parents*, and BT continues *make you talk to relatives* (Figure 14). Both describe situations that should prompt some reactions, but they are not verbally



Figure 13: The *when* meme with visual completion.



Figure 14: The *when* meme with ‘run-on’ TT/BT and visual completion.

described. Instead, the background images make it clear what the response might be like. In the first case, the image is the face of Leonardo Di Caprio, apparently at a red-carpet event, with his eyes closed and his lips in a grimace to represent disappointment, annoyance, tiredness, and possibly other emotions. The role of the background image is then to provide the completion of the construction, but not in a textual form. The usage reminds us of the way in which (among other things) facial expressions can be expressed in online communication as skeletal drawings referred to as emoji. The popularity of emoji is on the rise, as they provide a very economical way of making it clear how the writer of a tweet or a Facebook message feels about a situation. But emoji are typically independent expressions, added to the text to clarify one’s

attitude. Images such as that of DiCaprio in Figure 13 are essential components of multimodal constructions, in which they complete the textual component. The images in *when*-memes thus function as equivalents of main clauses of generic temporal constructions, but they only explain the emotional response, not other consequences. In Figure 13, DiCaprio's "Oh, no!" face completes the sentence, which now can be read as something like *When you're about to leave work and the boss says "before you go ...", you feel really disappointed and annoyed*, but not as *When you're about to leave work and the boss says "before you go ...", you just get back to work*.

Our second example (*When your parents make you talk to relatives*) uses a more complex image in which a young woman is facing the camera and holding a mobile phone to her ear; on the left side, we see a hand holding a gun to her head. The image targets the emotion felt by the person forced into unwanted conversations through a form basically resembling simile (*When your parents make you talk to relatives, you feel as though someone were holding a gun to your head*). Perhaps ironically, at least for viewers familiar with the scene from the film *Clueless* the image is taken from, in the actual scene the character Cher was being forced to hand over her phone and handbag to a man raiding her in a car park, so she was not being forced to talk at all.⁶

In his analysis of *when*-memes Lou (2017) argues that they are best viewed as instances of multimodal simile, not multimodal metaphor, since the source and target domains are not blended, but continue to be available as dissimilar domains, with only selected aspects being crossmapped (see Lou 2017: 108–109; 113–115). When we find ourselves in the situation described verbally in the *when*-clause text, we respond in ways similar to the response presented visually in the meme's image. One anonymous reviewer suggested that the meme types analyzed in previous sections also involve some form of simile, but we would argue they involve at most simulative thinking in resolving their interpretation, but are not constructionally cases of multimodal simile. Cases like Figures 13–14 inherently rely on a simulative frame in order to resolve the incomplete verbal part into a complete constructional meaning: 'when you are in situation X, you act or feel like Y', in which Y is the depicted emotion (or, in other types of *when*-memes, the depicted scene or event; see Lou 2017). This is not the case with the other memes discussed previously: for instance, in *One does not simply win an argument with a woman* we

⁶ We are grateful to one of the anonymous referees for identifying the source of the image for us.

know thanks to the frame provided by the stock phrase and the image that winning an argument with a woman is a futile undertaking, without having to construe this as being in any specific sense like walking into Mordor (indeed, many meme viewers will lack the detailed knowledge about *Lord of the Rings* for this). And more fundamentally, what is described in the text about the difficulty of arguing with a woman is not ‘like’ what we see depicted in the image, namely Boromir’s face, just as in the Good and Scumbag Girl/Guy memes, whatever the text describes is not ‘like’ the image of Gina, Steve, etc.

It is also worth noting the difference between *when*-memes and existing linguistic uses of ‘non-completed’ subordinate clauses. Without the image the *when*-clause clearly cannot be interpreted in the same way, as it is syntactically and informationally incomplete. This is different from what we see in some truncated conditional constructions (such as *If you would be so kind*), which do not provide the main clause because they frame other discourse elements as a request, and have arguably been conventionalized as so-called “insubordinated” constructions (in the sense of Evans 2007). In the case of the memes, what we see within the visual frame is all there is, even if it relies on the viewer’s ability to recognize the facial expression as signaling annoyance or a situation as threatening. Importantly, as in the other cases of memes we looked at above, these memes also have a categorizing role – having work time extended unexpectedly or being forced into unwanted social situations are added to categories of events that can annoy or threaten. Such memes are also ways of sharing experiences, as it is not likely that meme makers consider such situations as abstract; on the contrary, it is possible that events in one’s daily life prompt the need to share frustrations, and perhaps elicit understanding from other social media users. Shared sentiments about life’s small daily annoyances are part of the intersubjective power of memes, addressed to groups of users whose lifestyles and expectations may be similar.

The type of *when*-meme we have discussed requires recognition and alignment of viewpoints. Recognizing the frustration in DiCaprio’s face (Figure 13) or the terror in Cher’s face (Figure 14) is, in a way, sufficient to fill in the meaning of the image: for most meme viewers, it is not necessary to know the context in which DiCaprio made the face, or to know the details of the film scene in order to understand the meme. Rather, the emotional viewpoint expressed in the image is attributed to the meme maker, based on the meme viewer’s alignment with the annoyance of having to work longer or the stress involved in unwanted social situations. The memes tap into experiences the viewer either has or can easily imagine, and having the meme maker and the meme viewer share the emotional viewpoint is the point of the meme. A lot of

memetic creativity is about venting minor frustrations, which are understood intersubjectively, but evaluated from the discourse viewpoint level; there is no other way to connect the meme maker, meme viewer and DiCaprio into one communicative event.

The pervasiveness of *when*-memes has become noticeable in many people's social media feeds, but it is also starting to enter broader usage. One recent series of Flemish radio commercials for a chain of opticians called Pearle would be difficult to imagine without a context in which *when*-memes are very popular.⁷ (3) presents a translation (from the original Dutch) of the text of a single 30 second radio ad (the parts in italics are voiced by different actors than the ad's main speaker), consisting of a series of scenes introduced by "that moment when" (a precursor to the shorter "when" form we've focused on; see Lou 2017: 110–113):

- (3) That moment when you become the most annoying person in the cinema.
‘*Er, excuse me sir, is that guy naked or is that a beige pair of trousers? – Oh come on!*’. That moment when you can no longer read the menu. ‘*Same as usual, sir? – Er ... yes.*’ That moment when your car mechanic says ‘*That crack is not in your windscreen, you know, it’s in your glasses*’. And that moment when you hear on the radio “Now at Pearle opticians: 50% discount on all glass packets. Conditions at pearle.be” ... then obviously you go straight to Pearle. Duh!

The first two "moments", the cinema and restaurant scenes, in fact are not very simulative, since the fictive/exemplary interaction which follows the *when*-clause instantiates the same type of event rather than a comparable one from another domain. The third scene, at the garage, is intriguing: it only gives the *when*-clause, and leaves the rest to be imagined by the hearer. And the final punchline is a complete 'when/then' construction. While the first three moments chronicle frustrating and embarrassing moments, however, this consistent categorization is deliberately reversed in the final moment, which brings the resolution to all the embarrassing problems resulting from bad eye sight.

The usage here is thus quite different compared to the visual-textual cases we discussed, but arguably it is made possible by the context of many people sharing or liking *when*-memes, and constitutes a further illustration of what we saw for *said no one ever* and *One does not simply*, namely memes living on – in modified form – in more standard, offline usage.

⁷ We thank Eline Zenner for drawing our attention to this ad campaign.

6 Conclusion: Internet memes as viewpoint-driven multimodal constructions

As the cases we have surveyed have shown, memes are complex artifacts, in spite of their simple and replicable form. The *said no one ever* meme puts viewers on a garden path, aided by the image in the meme, only to result in a viewpoint reversal, revealing the apparent quote to be unattributable to anyone, and categorizing its content as untenable, ridiculous, and so on. The *One does not simply* meme originally relies on a rich cultural frame evoked by the face of a character in a film scene, now providing a template to continue to categorize ever more events as examples of futile undertakings. The Good Girl Gina and related memes allow verbally strongly reduced forms of generic predictive constructions (lacking conjunctions, determiners and overt subjects or embedded reporting clauses) to nevertheless call up the full predictive meaning thanks to the filling in of missing elements by the image and its built-in framing as good or scumbag character, categorizing behaviors as typical good girl/guy or scumbag girl/guy behavior. The type of *when*-memes we focused on, finally, complete an incomplete verbal construction visually to indicate what people behave like in certain types of stressful or embarrassing situations. All these memes intersubjectively rely on the assumption that viewers will recognize certain sets of beliefs and certain frames, and also assume that viewers are likely to share the discourse viewpoint ultimately expressed by the whole construction.

In this paper, we have not intended to provide a full account of all the issues, but rather signal the reasons why such multimodal artifacts deserve the attention of linguists. They pose questions on the nature of discourse and the use of linguistic resources in multimodal contexts. One of the central questions concerns the replicable nature of memes. The technical conditions under which memes are easy to spread but also to respond to creatively by reconfiguring and editing them need not concern us as linguists, but the linguistic reasons for their replicability need to be adequately explained. Our analyses suggest that the reasons are complex, involving constructional meaning, frames, constructional compositionality, and, perhaps primarily, the need for communicating about emotional responses to daily experiences.

We have seen that memes use existing constructions, such as direct speech or predictive conditional constructions, but in novel, multimodal ways, in the process creating their own constructional offspring, both multimodally (as with the Scumbag Hat acquiring its own specific function) and monomodally, in coining constructions used in journalism and advertising. We are not arguing that all types of Internet memes (including LOLcats, video spoofs, etc.)

are constructions pure and proper; this would need more research on each type. We do propose that the specific cases of image-macro memes be recognized as multimodal constructions, which often rely heavily on available frames. Just as construction grammar has long recognized clines of constructionality in dimensions of size (from morphemes to argument structure constructions, and beyond) and abstractness (from concrete constructs to abstract schematic constructions), we might begin to conceive of gradations in terms of modalities involved (from monomodal to multimodal). In addition, just like lexical concepts and grammatical constructions (Taylor 2003), multimodal constructions show prototype structure too, with core examples that are more characteristic and recognizable and more peripheral examples deviating from the prototypes in some respects, as in the case of *said no NOUN ever* as a variant of *said no one ever*, or of self-reflexive meta-memes compared to the memes they show reflexive awareness of.

As we have shown, constructional compositionality plays an important role in explaining how constructions may vary their form while not changing their core meaning, allowing for instance an expression such as *No pain, no gain* to be interpreted similarly to full conditionals such as *If you do not work hard, you will not achieve your goals*. Constructional compositionality is a correlate to Goldberg's (1995) argument regarding the meaning of lexical verbs in constructions, addressing the question why *laugh* can be used in a causative construction such as *They laughed him off the stage* for instance, or Croft's (2001) approach known as radical construction grammar, whereby individual grammatical forms obtain their meanings from constructions, rather than constructions obtaining meaning from the forms. What constructional compositionality adds is that, having developed their constructional meaning, grammatical forms can then express it without the support of the full construction. Memes naturally relax expectations about formal aspects as important as subject NPs or as minor as articles, as we have seen, and some of them treat complex expressions such as *One does not simply, said no one ever, or but that's none of my business* as constructional elements imposing formal restrictions on the realization of the expression as a whole, in that they determine the sequence of clauses (*One does not simply* initiates the text; *said no one ever* and *but that's none of my business* come at the end). They also determine the meaning of the constructional slots, so that for instance the BT of *One does not simply* counts as a description of a futile undertaking. Finally, the visual structure of a meme is as important as its linguistic structure, so that there is typically a clear visual frame that linguistic material fits into.

Another aspect of the replicability and malleability of memes we have stressed is heavy reliance on available frames. While the frames may become

backgrounded in time, they provide good starting points. This is the case with the *One does not simply* meme. A viewer may not recognize the *Lord of the Rings* background now, but initially it provided the framing to the phrase which made it easy for it to become a token of futile undertakings within a certain core discourse community familiar with the source frame. At the same time, the Good and Scumbag series of memes create frames too – what is expected of a nice person or a bad person. The fact that aspects of attire, such as the Scumbag Hat, may acquire framing and metonymic roles of their own shows quite clearly that the replicability depends on restrictions on form on the one hand, and frame metonymic patterns on the other.

Even more importantly, memes operate in complex viewpoint networks, which give support and stability to the constructions. They often carry elements of surprise, humor, or irony, so frame-shifting (as described in Coulson 2001) is an important aspect of how memes are processed. What we argue, though, is that frame-rich constructions, like the ones discussed here, need to be processed from the viewpoint level of the Discourse Viewpoint Space (Dancygier and Vandelanotte 2016). Consider, for instance, the meme in Figure 15 which portrays Barack Obama in a Scumbag Hat; the TT says *promises change*, the BT says *changes promise*. Not only does the viewer have to recognize Obama's face, she also has to know the meaning of the Scumbag Hat so as to appreciate that the politician is represented in a negative light rather than to start pondering the man's taste in headwear. It also helps if she can recall campaign promises and the way they are often felt to remain



Figure 15: A meme featuring the “Scumbag Hat” to add another viewpoint.

unfulfilled. This network of spaces and viewpoints is jointly signaled through the construction and the visual means, and reconciling the viewpoints and spaces profiled is central to the reading. Discourse Viewpoint Space is a necessary addition to the network.

The final aspect of memes we want to emphasize is the way they represent emotional and experiential meanings. They are very often about everyday situations or current political events, they are often jocular, and they provide opportunities for immediate commentary on current events and observed behaviors. However, rather than explain at length, they rely on intersubjectively accessible assumptions and salient visual and frame-metonymic representations. Counting on others to recognize facial expressions, evoking rich experiential scenes (such as having someone hold a gun to your head), quickly assigning behaviors to Good or Scumbag categories, using ready-made opinion-expressing formulae (*One does not simply, said no one ever, but that's none of my business*) – all these are ways to profile emotional and experiential viewpoints on situations. While the subject matter may often seem trivial, in a world of online communication these are important ways to quickly and effectively share opinions with groups of people one does not daily interact with in real life. Because of their emotional impact and opinion-forming power, these are forms of communication which are already occupying much of younger generations' communicative life, and their role may grow further. They allow not just sharing but co-creating of economically expressed, clever cultural artifacts, relying on ideas and emotions assumed to be shared or at least known by peers within a given discourse community, and they are shared precisely to elicit responses (likes, tags, comments) and further iterations, in a creative cycle both creating and sustaining the discourse community.

As illustrations of this, consider Figures 16 and 17. Figure 16 presents a meme sequence or 'chain meme', as we might call it, which forms a single image file and which demonstrates well how much knowledge of frames and assumed viewpoints can be left implicit among proficient users of memes. They will know that the first image features "Socially Awesome Penguin", who is popular and attractive, the second "Bad Luck Brian", whose every undertaking fails, and the third Scumbag Steve, whose behavior is deplorable and insensitive. Without the images and the knowledge and viewpoints associated with them, the text on its own would seem weird, choppy, and not very funny. Figure 17 is an example of a meme blend, in which the image derives from one popular meme, but a set phrase is taken from another. This adds further viewpoint complexity, inviting a kind of ironic distancing similar to the effect achieved in metamemes (memes about memes, as in "*Remember when I said 'said no one ever' out loud in conversation?" said no one ever*"). In Figure 17, for instance, there is humor



Figure 16: A chain meme.



Figure 17: Blending the *Kermit* meme (image) with *The most interesting man in the world* meme (text: “I don’t always X but when I do Y”).

both in the absurd contradiction within the text itself, and in its combination with the image clearly showing tea, suggesting, perhaps, that the target of the irony is ultimately the meme maker and meme consumer, and their silliness in producing and enjoying memes.

There are, of course, various other aspects of memes which merit further study, including the role of puns and wordplay and the area of multilingual meme combinations (Geeraerts and Zenner 2016). We have focused in this paper on the issues which seemed closest to some of the core interests of cognitive linguists: constructionality, multimodality, viewpoint and intersubjectivity. More generally, in spite of their apparent simplicity, memes present a challenge to the way we view communication. The interaction between language and image is at the core of their popularity, but, thanks to meme-making tools on line, they are extremely easy to make – it is often more efficient to use the existing multimodal formula than to type in an original comment, which is why many online exchanges are conducted through memetic forms. We need an approach which allows us to capture the ways such artifacts work, and how they tap into linguistic resources. We have pointed out a number of directions here, but, as one meme maker noticed, *One does not simply ... write a paper on memes*. Certainly, not just one.

Funding: The authors gratefully acknowledge the support provided by KU Leuven (Senior Fellowship, SF/16/004) which allowed Barbara Dancygier to carry out research at KU Leuven (September-December 2016) and which greatly facilitated our joint work on this project.

References

- Brideau, Kate & Charles Berrett. 2014. A brief introduction to impact: ‘The meme font’. *Journal of Visual Culture* 13(3). 307–313.
- Brône, Geert, Bert Oben, Annelies Jehoul, Jelena Vranjes & Kurt Feyaerts. 2017. Eye gaze and viewpoint in multimodal interaction management. *Cognitive Linguistics* 28(3).
- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347.
- Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805.
- Coulson, Seana. 2001. *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge: Cambridge University Press.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Dancygier, Barbara. 1998. *Conditionals and prediction: Time, knowledge and causation in conditional constructions* (Cambridge Studies in Linguistics 87). Cambridge: Cambridge University Press.
- Dancygier, Barbara. 2012. *The language of stories: A cognitive approach*. Cambridge: Cambridge University Press.
- Dancygier, Barbara. 2017. Viewpoint, discourse, and multimodal artifacts. Plenary lecture, AFLiCo 7 - Discourse, Cognition & Constructions: Implications & Applications, 7th International Conference of the French Association for Cognitive Linguistics, Liège University, Belgium, 31 May–3 June.
- Dancygier, Barbara & Eve Sweetser. 2005. *Mental spaces in grammar: Conditional constructions* (Cambridge Studies in Linguistics 108). Cambridge: Cambridge University Press.
- Dancygier, Barbara & Lieven Vandeborre. 2016. Discourse viewpoint as network. In Barbara Dancygier, Lu Wei-Lun & Arie Verhagen (eds.), *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities* (Cognitive Linguistics Research 55), 13–40. Berlin & Boston: De Gruyter Mouton.
- David, Oana. 2016. *Metaphor in the grammar of argument realization*. Berkeley: University of California, Berkeley dissertation.
- Dawkins, Richard. 1976. *The selfish gene*. Oxford: Oxford University Press.
- Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.
- Fauconnier, Gilles & Mark Turner. 2002. *The way we think: Conceptual blending and the mind’s hidden complexities*. New York: Basic Books.
- Fillmore, Charles J. 1988. The mechanisms of ‘construction grammar’. *Berkeley Linguistics Society* 14. 35–55.
- Fillmore, Charles J., Paul Kay & Mary Catherine O’Connor. 1988. Regularity and idiomativity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.
- Forceville, Charles J., Urios-Aparisi, Eduardo (ed.) 2009. *Multimodal metaphor* (Applications of Cognitive Linguistics 11). Berlin & New York: De Gruyter Mouton.
- Geeraerts, Dirk & Eline Zenner. 2016. One does not simply borrow a meme: Memetics from the perspective of cognitive contact linguistics. Plenary talk at the symposium “The dynamics of wordplay”, University of Trier, 29 September–1 October.

- Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.
- Green, Jennifer. 2014. *Drawn from the ground: Sound, sign and inscription in Central Australian sand stories*. Cambridge: Cambridge University Press.
- Israel, Michael. 2011. *The grammar of polarity: Pragmatics, sensitivity, and the logic of scales*. Cambridge: Cambridge University Press.
- Kress, Gunther & Van Leeuwen, Theo. 2001. *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar. Volume I: Theoretical prerequisites*. Stanford: Stanford University Press.
- Lou, Adrian. 2017. Multimodal simile: The “when” meme in social media discourse. *English Text Construction* 10(1). 106–131.
- Louw, Bill. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Mona Baker, Gill Francis & Elena Tognini-Bonelli (eds.), *Text and technology. In honour of John Sinclair*, 157–176. Amsterdam: John Benjamins.
- McNeill, David. 2006. Gesture, gaze, and ground. In Steve Renals & Samy Bengio (eds.), *Machine learning for multimodal interaction*, 1–14. Berlin & Heidelberg: Springer Verlag.
- Pascual, Esther. 2002. *Imaginary trialogues: Conceptual blending and fictive interaction in criminal courts*. Utrecht: LOT.
- Pascual, Esther. 2014. *Fictive interaction: The conversation frame in thought, language, and discourse* (Human Cognitive Processing 47). Amsterdam & Philadelphia: John Benjamins.
- Rohrer, Tim. 2005. Mimesis, artistic inspiration and the blends we live by. *Journal of Pragmatics* 37(10). 1686–1716.
- Ruppenhofer, Josef & Laura A. Michaelis. 2010. A constructional account of genre-based argument omissions. *Constructions and Frames* 2(2). 158–184.
- Sanders, José & Gisela Redeker. 1996. Perspective and the representation of speech and thought in narrative discourse. In Gilles Fauconnier & Eve Sweetser (eds.), *Spaces, worlds and grammar* (Cognitive Theory of Language and Culture), 290–317. Chicago: University of Chicago Press.
- Shifman, Limor. 2012. An anatomy of a YouTube meme. *New Media & Society* 14(2). 187–203.
- Shifman, Limor. 2013. Memes in a digital world: Reconciling with a conceptual troublemaker. *Journal of Computer-Mediated Communication* 18. 362–377.
- Shifman, Limor. 2014. *Memes in digital culture*. Cambridge, MA: The MIT Press.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stubbs, Michael. 1995. Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2(1). 23–55.
- Taylor, John R. 2003. *Linguistic categorization*, 3rd edn. Oxford: Oxford University Press.
- Tobin, Vera & Michael Israel. 2012. Irony as a viewpoint phenomenon. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 25–46. Cambridge: Cambridge University Press.
- Vandelanotte, Lieven. 2009. *Speech and thought representation in English: A cognitive-functional approach* (Topics in English Linguistics 65). Berlin & New York: Mouton de Gruyter.
- Verhagen, Arie. 2005. *Constructions of intersubjectivity*. Oxford: Oxford University Press.
- Wiggins, Bradley E. & G. Bret Bowers. 2015. Memes as genre: A structurational analysis of the memescape. *New Media & Society* 17(11). 1886–1906.