Compound Figures: A Multi-Channel View of Communication and Psychological Plausibility

Mihaela Popa-Wyatt[1]

Abstract

Philosophical views of language have traditionally been focused on notions of truth. This is a reconstructive view in that we try to extract from an utterance in context what the sentence and speaker meaning are. This focus on meaning extraction from word sequences alone, however, is challenged by utterances which combine different types of figures. This paper argues that what appears to be a special case of ironic utterances—ironic metaphorical compounds—sheds light on the requirements for psychological plausibility of a theory of communication and thus presents a different view of communication and language to that dominant in philosophy of language. In the view presented here, the hearer does not extract the speaker's communicative intention from the sequence of words in the utterance, but from other channels (gesture, intonation, facial expression), so as to constrain the inferential space for the sentence and speaker meaning. Specifically, we examine an example of ironic metaphor discussed by Stern (2000). He argues that ironic content is logically dependent on metaphorical content, but makes no claims about how psychologically plausible this is in terms of the processing order. We argue that a straightforward translation of logical order into temporal order makes little sense. The primary sticking point is that without a prior understanding of the speaker's communicative intentions, it is computationally more challenging to process the sub-component meanings. An alternative solution based on communicative channels leads us to a more psychologically plausible account of the structure of communicative acts and intentions. This provides support for the psychological realism of a richer theory of communicative intent.

1 Communication: Language and Signalling

The human faculty for language is one of our key tools to engage with the world around us: it enables us to shape our worldview through communication and coordination with others. We communicate and interact with each other by exchanging information, planning, debating, expressing opinions and attitudes, and creating new ways of thinking about and experiencing the world. We are thus able to influence each other's thinking and behaviour to the extent that we can form successful projects of coordinating on signalling acts.

There are three features of signalling that makes human communication special. The first feature is the *use of natural language*. Human languages have a syntax and a semantics, with huge combinatorial power, which enable us to communicate a potentially infinite number of different messages using a finite system of symbols/words. Indeed, the landmark of human communication is productivity and creativity. This enables us to produce not only new utterances that we never made before, but also to put to new uses linguistic devices already used with specific purposes, to different purposes, thereby ensuring the stabilization of those devices (cf. Millikan 2004).

The second feature of human communication is the exploitation of *higher order intentional systems*. This involves "recursive" mechanisms employed to generate higher-order intentional states (such as the 2nd-order belief that Jones intends to sit down). Davidson (1982) holds that thought is impossible without language. Indeed, we are

✉ Mihaela Popa-Wyatt
 popa.michaela@gmail.com

1 Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstraße 18, 10117 Berlin, Germany

able to communicate not only because we have a language, but because we can show and recognize their intentions to communicate something through a wide variety of ostensive stimuli which are manifestly intended to attract the audience's attention, and to focus it on the message being communicated. To this end, we use ostension and inference as particular ways for doing things to each other. But most importantly they engage in what Tomasello et al. (2005) call "shared intentionality," i.e. the motivation and ability to participate in collaborative activities with shared goals and intentions. This requires not just sophisticated mindreading abilities, but also a willingness to share psychological states with others and the cognitive capacity to conceive of a mutually shared goal. Indeed, it is the feature of ostension, which comes with a communicative intention to communicate something to someone, that is the hallmark of human (intentional) communication.

The third feature of human communication is the presence of *communicative acts*, by means of which humans not only communicate information but affect each other in a huge variety of ways. We send signals to each other not only to carry information about our worldview, but also to talk about things that are not present (past or imaginary) in the situation of communication. Moreover, linguistic communication involves more than transmitting a signal and decoding the message; *inference* to information from the context (not only from the signal) is required on the part of the audience in order to grasp the message conveyed. As Origgi and Sperber (2000, p. 142) write:

> "[Language] allows individuals to benefit from the perceptions and inferences of others and increases their knowledge well beyond that which they could acquire on their own. It allows elaborate forms of co-ordinated planning and action. It can be used for manipulation, deceit, display of wit, seduction, maintenance of social relationships, all of which have fitness consequences."

To understand how these three features work together, it's important to think of human communication as a purposeful or goal-directed activity occurring between at least two participants. Participants with independent goals may attain their own purposes by coordinating their behaviours around a common/joint purpose, that of succeeding in communicating. To that extent, both the production of a stimulus/ utterance and the understanding of it or a corresponding action need to be *coordinated*. The mechanisms underlying the communicators' independent tasks are designed to facilitate their coordination in the sense that it serves their own purposes by bringing their separate behaviours into a coordination that benefits both.

In this paper, we approach some core features of human communication and signalling from the perspective of the philosophy of language. Philosophers of language have taken a view of language as primarily being concerned with the transmission and processing of word sequences that convey meaning. The meaning of a sentence is the result of compositional semantics applied to the sequence of words. This view found formal expression in elegant theories of syntax and semantics that also have claims to psychological plausibility (Steedman 2000). The meaning of an utterance is commonly referred to as the *speaker meaning* and arises from the sentence meaning coupled with pragmatic processes that incorporate contextual information to derive the speaker's intention.

However, communicative acts do not always bear their meaning on their sleeves. Many aspects of what we are mean to get across are not explicitly stated in an utterance but need to be interpreted by means of pragmatic and common world knowledge. This is especially the case when understanding figurative language. Figurative (or non-literal) utterances encompass additional information associated with the contextual situation in which they occur, and convey more subtle and often different meanings that go beyond literal sentence meanings. In order to comprehend non-literal utterances, contextual information plays an important role in determining a speakers' intended meaning. The question of when and how contextual information is processed and integrated in understanding figurative sentences is thus key to a successful theory of communication.

In this paper we suggest a modification to this view of language, driven by observations about certain kinds of utterances: compound figures. These suggest that, if we wish a psychologically plausible account of language that we should take into account the fact that human communication relies on the use of *multiple channels*: intonation, gesture, facial expression as well as word sequence. Only in this way, can the view of language proposed by philosophers be reconciled to psychologically plausible accounts.

2 From Communicative Intentions to Multiple Meanings

Communication is a joint activity where speakers and hearers seek to coordinate as efficiently as possible on a range of meanings. Key to this is our ability to act with, and attribute to others, communicative intent. This idea has been the cornerstone in Grice's (1989) seminal work. He argues that what distinguishes communicative acts from other forms of intentional actions is their "*overtness*" or what relevance theorists call "mutual manifestness"—i.e. the fact that an utterance or an action is made manifest publicly. To capture the overtness in our linguistic behaviour, Grice (p. 151) defines the distinctive mental states underpinning our communicative intentions as below—where (*i1*) gives the content of

what is conveyed with a given communicative act, whereas (*i2*)–(*i3*) are responsible for the publicity requirement:

> "Speaker U meant something by uttering x" is true iff, for some audience A, U uttered x intending
>
> (*i1*) A to produce a particular response *r* [usually forming a belief *p* or performing
>
> an action *q*]
>
> (*i2*) A to think (recognize) that U intends (*i1*)
>
> (*i3*) A to fulfill (*i1*) on the basis of his fulfillment of (*i2*)."[1]

The Gricean underlying cognitive structure to communicative intent has been taken to be too cognitively demanding for it to be a theoretically adequate account of the prerequisites of a successful communication.[2] (We shall return to this in due course.)

Despite such worries, Grice's ideas have been foundational to theories of meaning and communicative behaviour more generally. While different types of meaning have been identified—e.g. semantic or sentence meaning vs. pragmatic or utterance meaning in context—the idea persists that for any one of these types of meaning each utterance has one overall meaning. Grice divides an utterance meaning into two parts: in uttering a sentence, (1) a speaker intends to *say* something, and (2) to *implicate* something else, or something more. Saying comes *before* implicating, and thus functions as a central part of the supporting evidence for working out what else, or what more, the speaker may have implicated in uttering the sentence.

Grice was careful to distinguish how meaning supervenes on the speaker's communicative intentions, as opposed to how hearers try to work out what it is that the speaker meant. The former question concerns a *metaphysical* determination of meaning—i.e. where meaning comes from the mental states of the speaker and these mental states involve higher order communicative intentions to get something across. The latter question concerns an *epistemological* determination of meaning of how hearers recognise the speaker's communicative intention and work out what it is that they are trying to communicate.

This distinction pertains to the kinds of information and processes that hearers use as evidence to form a hypothesis about what the speaker has said and/or implicated in uttering a sentence. Though the metaphysical and epistemological questions are separate, they constrain one another. Speakers want (and expect) to be understood, and hearers seek

(and expect) to understand. So, in forming communicative intentions the speaker relies on the hearer's ability to grasp those intentions. Vice versa, in interpreting an utterance, the hearer relies in turn on the speaker's capacity to exploit this ability.[3]

When things go well and communication succeeds, there is a fair overlap between what the speaker means to get across and what the hearer is able to work out the speaker to have actually got across. This is because in uttering a sentence, the speaker typically means to convey one overall message. But what if the speaker intends that the hearer should understand (or at least contemplate) *multiple meanings*? What if communicative acts in fact rely for their meaning on the hearer's ability to hold several contradictory interpretations in their mind, and to draw inferences from each of those?

Once we begin to contemplate more complex forms of communication, we quickly realise that many utterances have multiple meanings at the same time, or even across time. Garden-path monologues are the stock in trade of stand-up comedians, and derive their humour precisely from the intention to convey one meaning at one point in time, and then subsequently to convey a different meaning of the same words later on.

Finally, meaning is not the sole provenance of words. Words are reified because of print. But communication often goes beyond words. Our communication repertoire includes many more signals other than linguistic ones: e.g. facial expression, vocal intonation, gesture, body posture, etc. convey meaning via multiple communicative channels. Notably, pupil dilation, tears, a flushed face, or an eyebrow flash are not signals from parallel communicative channels entirely independent of speech, nor are they the handmaidens of words, there only to elucidate the correct interpretation of the utterance. In fact they weave with words in sequences of signals that are at any one moment coming in a coordinated way through multiple channels. Sometimes the meanings conveyed by non-word channels convey different meanings to words, and the speaker means to convey those multiple meanings in parallel. This requires in turn an ability to pool multiple sources of signals and integrate them together in an attempt to recover the speaker's meaning.

The fact that most linguists and philosophers talk about the things we do with *words* shows the limited part of the problem we choose to study. We are like the monk who

---

[1] Some authors have reduced this structure to (*i1*) and (*i2*). Grice himself proposed various forms of this definition to accommodate various objections (see Neale 1992).

[2] See Millikan (1984), Sperber and Wilson (1995), Wilson and Sperber (2012), Carston (2002), Green (2007), Bar-On (2013), Moore (2017).

[3] Despite this distinction being prominent in Grice's writing, his theory has often been (mis-)construed as primarily a speaker-focused account. The interpretation side of his account has been taken to be limited to a rational reconstruction: i.e. uncovering the paths that hearers would need to follow to recover the speaker's meaning by showing how such meaning is calculable rather than how hearers actually process it.

holds the elephant's ear and claims that the elephant is an animal round and thin and flat like a pancake. Humans do use words in sentences in various conversational settings. That is special. And sentences convey very rich meanings. But humans and our ancestors have communicated for millennia with other means,[4] and these deeply embedded ways of communicating meaning are not so obvious to us, perhaps because they require so little conscious effort. Indeed, they suffer in the presence of it. Thus, they are easy to ignore.

It's thus important to acknowledge that theorising about speakers, hearers, and utterances, is an artefact of our theoretical idealisations, and it would be desirable to broaden our repertoire with multi-modal communicative acts, signalling, and the intersubjectivity of meaning. For simplicity sake, we shall continue to use the more familiar notions, though a change in vocabulary is required to free us from this box. This paper is thus an attempt to sketch a theory of communication that takes account of these facts, rather than of merely the meanings of utterances. We now try to encapsulate the principles that a good theory of communication might follow.

(1) Communicative acts are inherently *multi-modal*. Rarely is human communication unimodal. The spaces of meanings signalled by different channels of communication are separate, but they can be intertwined, i.e. meanings in one channel of mode can indicate or modify meanings in another mode.

(2) Since communicative acts are multi-modal, the signals transmitted in each channel convey their own meanings, but typically the information on each separate channel is combined and integrated to form a model of the speaker's meaning. This can come in three possible ways:

(a) Speakers sometimes intend to convey multiple contradictory meanings at the same time, i.e. for the hearer to concurrently hold multiple conflictual meanings from different signals. An instance of this occurs in irony and humour.

(b) Speakers sometimes intend to convey one meaning at one time and another meaning at a later point, thus evolving over time. An instance of this occurs in garden-path sentences and jokes.

(c) Speakers sometimes intend to convey two different meanings at the same time, which requires embedding one inside another. An instance of this occurs in what we call *compound figures*.

In what follows we outline the main ingredients for a richer theory of multi-channel communication by focusing on *compound figures*.

3 Compound Figures: Logical Order of Interpretation

Thinking and speaking in figures is key to our creativity. With metaphor we describe how things are in the world by presenting them in a new, evocative light. We re-imagine one thing as another by evoking similarities between them. With irony we pretend to believe what we appear to say—*P* (for some contextually sentence meaning *P*)—in order to give a ridiculing-portrayal of someone believing, asserting *P*, or behaving like *P*, thereby drawing attention to how *P* falls short of our expectations.[5] In speaking ironically we thus express ridicule towards someone who would believe *P*, thereby conveying a belief that [*Invert-P*] is the case. Both metaphor and irony thus enable us to do one thing in order to achieve another: say one thing to mean another. This comes easily because of our inferential abilities both to work out meaning in context and to re-purpose words for new uses.

Linguists and philosophers of language have typically focused either on uses of self-standing figurative utterances (e.g. metaphorical, ironic), or embedded inside a more complex utterance such as conditionals, belief reports, modals, etc. (see Levinson 2000; Camp 2006, 2012; Wearing 2013; Bezuidenhout 2001, 2015; Popa-Wyatt 2009, 2018; Barker and Popa-Wyatt 2015).[6] There is however a sub-class of ironic utterances in which the irony builds upon another figurative utterance to form an ironic compound. This is different from cases in which some part of the utterance is metaphorical, while another is ironic, as when saying about someone who is bullying their friend to get what they want: "*Oh yes, the meeting went brilliantly, she flayed them alive*". The metaphor and irony are used here disjointly (i.e. applying to different segments of the same utterance).

In this paper we shall be focusing on cases where two different figures are used *conjointly*: i.e. they combine together to form a compound figure which draws on both meanings, though it cannot be reduced to either. Grice (1989) was the

---

[4] See Sterelny (2017), Planer (2017).

[5] See Wilson (2006) for a paradigm echoic account; Currie (2006) and Walton (1990) for a paradigm pretence account; for a hybrid account, see Popa-Wyatt (2014). A key ingredient of these accounts is the focus on attitude-expression. The specific details are not critical for our purposes here: for simplicity we include both attitude and inverted content under the more generic ironic content.

[6] For example, metaphor embeds in a conditional's antecedent: ''If music be the food of love, then play on'' (Shakespeare). Also, irony embeds in a conditional's antecedent: "If Bill has been such a fine friend, you shouldn't speak to him again".

first to point out the existence of such compound figures when he gives the example of an angry wife saying to her husband:

(1) You're the cream in my coffee.

to convey that her spouse has fallen short of the speaker's affection. Grice also notes that such cases require a *determinate order of interpretation* in that the hearer has to reach first the metaphor 'You are my pride and joy' and then calculate an ironic interpretation 'You are my bane' on the basis of metaphor. However, Grice does not give an argument for this preferred order of interpretation. Nor does he explain how the passage from metaphorical to ironical meaning is negotiated.

Stern (2000), Bezuidenhout (2001, 2015), Camp (2006, 2012), have provided various arguments for why we should prefer an order of interpretation in which metaphor has priority over irony. Here are some examples where the metaphors are quite novel:

(2) [*of an illegible handwriting*] What delicate lacework! (Stern 2000)
(3) [*of an unattractive woman*] She is the Taj Mahal. (Bezuidenhout 2001)
(4) [*of a terrible orator*] Norman really is God's fountain pen, isn't he. (Soames 2008)
(5) [*of an old lady*] The fountain of youth is waiting in line for her pension. (Adapted from Camp 2012)

(1)–(5) are examples where metaphor and irony are combined into a compound figure that contains both figures, though it cannot be reduced to either. We will use (2) as our toy example. Imagine a student essay with a messy piece of handwriting, illegible and covered in ink blotches. In uttering (2) the speaker is not making a serious remark, but is rather ridiculing the idea that one might be thinking of the handwriting as exhibiting artistic value. The utterance is both metaphorical and ironic. Note that the same sentence might be uttered on another occasion to say that the handwriting shows care and carefulness, craft and training, a wonderful attention to subtle calligraphic flourishes. In this case, the utterance is a non-ironic metaphor. Also, when (2) is uttered about a pair of curtains that your dog has just shredded to pieces, the utterance is a non-metaphoric irony.

Here we shall focus on compound figures containing both metaphorical and ironic meanings. How should we describe utterances as in (1)–(5)? Is it an ironic metaphor, or a metaphorical irony? For Stern (2000, p. 235), the issue concerns the "*logical order of interpretation*". Does the metaphoric interpretation depend on the ironic interpretation, or vice versa? Stern argues that the metaphor has priority over irony in the structure of what is communicated in that the ironic

content is *logically dependent* on the metaphorical content. The issue concerns "whether one interpretation is conditioned on the other" (Stern 2000, p. 235). This makes sense.

Let's refer to this priority claim as *Metaphor Priority Thesis* (MPT). Stern and Bezuidenhout are careful to avoid claiming that logical priority has anything to do with the temporal order in which processing operations occur as part of the hearer's actual interpretation, though it may have implications for these. So, we can distinguish two possible priority claims: *Logical-MPT* (i.e. one interpretation is conditioned on the other) and *Temporal-MPT* (i.e. one is typically processed before the other). Here we are primarily concerned with the psychological order of interpretation, though we shall first rehearse the arguments provided for the logical order of interpretation. Logical-MPT can be expressed thus:

> **Logical-MPT:** Metaphor is *prior* to irony in the sense that in the logical order of interpretation, the metaphorical content *must* come first.

Logical-MPT is a claim about the structure of the content being communicated. It states that we first derive the metaphorical content, and then use that to derive the ironic attitude/content. This means that the metaphorical content is logically prior to the ironic content in that the latter builds on, and is conditioned, by the former. Both Stern and Bezuidenhout use Logical-MPT as the starting point for arguments that irony and metaphor are markedly different types of figurative speech, thus conveying distinct content-types—i.e. metaphor is truth-conditional, irony is non-truth-conditional in that the former is open to dispute in way that is not available for the latter.

For Stern, metaphor and irony belong to two distinct families of figures (M-type and I-type) such that I-type figures depend on M-type figures. If M-type and I-type were the same, he argues, then we should expect freedom as to how they might be logically ordered. Stern's explanation of this logical dependence is as follows: ''(M)-type figures are semantic interpretations, interpretations determined by the semantic structure of the language; whereas (I)-type figures are post-semantic'' (2000, p. 238). By "post-semantic'' Stern means that irony is pragmatic (i.e. implicature). Essentially, metaphor priority is motivated by the fact that metaphor and irony employ different interpretive functions. For Stern (2000, p. 237), metaphorical interpretations are semantic operations on sentences that yield propositional contents in their contexts. Ironic interpretations, in contrast, are post-semantic operations on propositional contents to yield (different) propositional contents. Since semantic operations are logically prior to post-semantic operations, Logical-MPT follows.

Bezuidenhout (2001) agrees with Stern that metaphorical content contributes to what is said or asserted (rather than

being implicated), but offers a pragmatic explanation. She argues that Logical-MPT is a consequence of how ironic metaphor is interpreted. It follows from a natural criterion according to which interpretations that serve as *input for launching further implicatures*, belong to asserted-content, rather than implicated-content. Only metaphor meets this criterion since in ironic metaphor, the metaphorical interpretation is first generated from the particular expressions employed in a sentence, and then launches an ironic implicature.[7] Since asserted-content is determined prior to, and inferentially warrants, the implicature-calculation, Logical-MPT follows.

Camp (2006, 2012) explains the distinctiveness between metaphor and irony in terms of scope. Metaphor operates *locally* on expressions (before the whole utterance is computed), whereas irony operates *globally* on propositional contents to determine new contents. Since local operations work prior to global operations, this supports Logical MPT. Thus, in those cases in which the metaphoric interpretation is local, the irony swings into play only after all interpretations involving words have been calculated.

The content-based priority claims by Stern, Bezuidenhout, and Camp for the distinctiveness between metaphor and irony seem correct, as far as they go.[8] They explain the logical priority in terms of the content-types of the two figures: i.e. such that the content-type of irony depends logically on the content-type of metaphor. This content-type dependence thus explains the data, and the data is taken to be supportive of the particular content-type distinctiveness between metaphor and irony.

4 Compound Figures: Psychological Order of Interpretation

None of Stern, Bezuidenhout, or Camp make any claims to the psychological plausibility of what Logical-MPT tells us about how hearers actually process the compound utterance to infer its meaning. A straightforward translation of Logical-MPT into a sequence of computational steps requires that the metaphor be processed *temporally* prior to irony. Let's call this *Temporal-MPT*. Since this is a claim about how hearers process the utterance in practice, it is thus a psychological claim. What model might be able to capture this claim?

In order to evaluate candidate models it is worth distinguishing between two adequacy criteria—*descriptive* and *theoretical adequacy*. (Neo-)Gricean theories have been taken to be only descriptively adequate in that they offer an account of the mechanisms of interpretation and rules via which our successful communications *could* be arrived at. They do not claim that it is via these rules that successful communication *does* in fact occur. This offers a rational reconstruction instead of a psychologically grounded model. Many critics have suggested that a (neo-)Gricean model fails to fit with psychological evidence concerning various forms of linguistic understanding (including figurative speech), and that we should look for models that better comply with such evidence.

Theories of communication more integrated into cognitive sciences such as relevance theory and other contextualist accounts seek to have their predictions on the goals of communication and reasoning testable and empirically validated. They thus aim to be not only descriptively adequate but to also correctly identify the real processes via which the interpretation of an utterance is actually achieved. Though these accounts have moved the field towards an empirically-driven paradigm, it is fair to note that Grice was not concerned to investigate the psychological reality of the mental processes that hearers carry on in interpreting utterances. He was not however entirely insensitive to the criticism that because the inference procedures he identifies are not conscious, they are unreal. He sought to deflect the criticism by distinguishing between explicit inferential reasoning to a conclusion and a "quick way" of reaching the same conclusion. Inferential transitions between thoughts/propositions are performed via fully explicit rules, but they give rise to mental shortcuts if the transitions are repeated. As Grice (2001, p. 17) comments:

> "We have… a 'hard way' of making inferential moves; [a] laborious, step-by-step procedure [which] consumes time and energy… A substitute for the hard way, the quick way, … made possible by habituation and intention, is [also] available to us, and the capacity for it (which is sometimes called intelligence and is known to be variable in degree) is a desirable quality."

Grice's idea, then, is that if an inference procedure is pursued explicitly (consciously) sufficiently often, it can become second nature: sub-personal shortcuts can replace the fully explicit steps via which a conclusion might be derived.

With this caveat aside, a naïve implementation of Temporal-MPT requires first deriving semantic operations and then pragmatic operations. This then corresponds to a three-stage processing of compound figures: first deriving the literal

---

[7] Camp (2006) objects to this argument to the extent that implicatures too can work as input for other implicatures. Though Camp (2012) defends metaphor as truth-conditional.

[8] Explanations of this ilk are also proposed by Carston and Wearing (2015) who extend the analysis from ironic metaphor to hyperbolic compounds. See also Popa-Wyatt (2020a, b).

content, then metaphorical content, and then building the ironic content on top of it[9]:

semantics > pragmatics
utterance > **said-content** > **implicated content** > speaker meaning
metaphor > **irony** > ironic metaphor

In this model the inference is implemented sequentially as proceeding from left to right: thus the ironic implicature builds on what is said metaphorically. Let's call this the *uni-directional processing model*. This pipeline view corresponds to a naïve way of carving up the semantics/pragmatics divide in that general pragmatic inference (i.e. implicature) is not supposed to intrude into semantic (compositional) processing.[10] It follows then that it should not be the case that an ironic implicature can affect the processing of what is said metaphorically.

This uni-directional model is unsatisfactory, however. First, it reflects an idealised logical order of interpretation in which semantic and pragmatic operations run sequentially and independently one of another. The motivation for this is to ensure that communication is grounded in reason and inference, i.e. such that only one type of content can work as a reason or piece of evidence that logically warrants the other type of content in a valid reasoning. But actual processing need not follow this exact order. There is no reason to expect that human mind is bound by such requirements. On the contrary, we might expect that cognitive efficiency requires pooling together all sorts of relevant information as they come to our attention in a bid for integration.[11]

Second, this naïve view predicts a computationally inefficient processing of compound figures. What's special about compound figures is that by conjoining together two figures of speech the speaker is arranging them in a particular structure where one serves as means for another. This is important because it can guide the processing itself by revealing what function each figure plays in the compound and how they put constraints on one another. An immediate consequence of this is the streamlining of the metaphorical processing. Because the utterances in (1)–(5) are not used as self-standing metaphors, there is no need to process an open-ended range of similarities between two conceptual domains: that would be computationally too costly. This is because not all metaphorical properties retrieved in a first-stage are suitable for further ironic processing in a second-stage. Instead the overall processing is arguably more integrated than a sequential order predicts. (More details below.)

Computational considerations of this sort do not figure much in debates on language, yet they should. This is because in theorising about language uses we are not describing a wholly abstract entity, but a system of communication that is used in practice. This practice requires efficiency in inference, so it's important to define clear constraints that render processing tractable and efficient. Work by pragmaticists and relevance theorists has moved into this direction by looking at ways of rendering their claims psychologically plausible. For example, Wilson and Sperber (2012), Carston (2002) develop a notion of "mutual adjustment" between said- or asserted-content (i.e. "explicature") and implicated-content. Essentially, expectations of relevance can warrant specific implicatures, the derivation of which requires the explicit content to be adequately enriched. This involves *backwards inferences*: when a contextual implication is derived, the hearer treats it as a potential implicature of the utterance, which may in turn help adjust some word's meaning into a modulated meaning, thus shaping the truth conditional content which can then warrant the expected implicature.[12] Hence, the contextual assumptions get their inferential warrant entirely from backwards confirmation. Along similar lines, Bezuidenhout (2015) argues for a more fluid derivation of said/asserted-content and implicated-content—i.e. where implicatures can take as input not only said-content but also other forms of implicit (pragmatically derived) content (including metaphorical content).

This suggests a modified model where information can travel in both directions, such that given a hypothesis about a *prospective implicated-content* the hearer can backtrack to modulate what is said more efficiently. This violates the isolation of said-content from effects of implicature-inference. Nevertheless, if we accept a more fluid processing between semantic and pragmatic information, we can begin to see the basis for a more psychologically plausible account. A *bi-directional implementation* would allow inferences to pass back and forth along the pipeline. Thus, the moment we have a hypothesis about what the implicature might be, this hypothesis could be used to guide processing towards that communicative goal. If the speaker gives additional information about that goal, then this indication can enable

---

[9] Items in bold indicate processing stages.

[10] This is a simplification of the semantics/pragmatics debates, in particular with regards to how much contextual supplementation is needed, and what kind of pragmatic processes should be allowed to determine what is said or asserted with a given utterance. However, these debates remain relatively unconcerned with processing questions.

[11] This is very much in the spirit of relevance theory accounts. More on this below.

[12] For example, if the conclusion—implicature—seems a promising hypothesis about the speaker meaning, and the entire interpretation is consistent with the expectations of relevance raised by the utterance, then the hearer has good reason to adjust the premises so that they warrant that conclusion.

an efficient search for said-content, implicatures, and thus adjust speaker meaning.

Applied to compound figures, we want to argue that recognising the compound as primarily ironic comes *first* in the *order of communicative intentions*, serving as an indication to constrain the processing order in which the sub-component meanings are to be retrieved. This leads to a more psychologically plausible processing ordering claim than a naïve Temporal-MPT that mimics Logical-MPT. Central to our argument is that by taking into account computational considerations we come to an adjusted view of how compound figures are actually processed, insofar as compound figures force us to confront the priority problem in a way that is absent in self-standing figurative utterances.

The argument proceeds as follows. First, we describe the inferential pattern corresponding to a naïve Temporal-MPT implementation of the processing underpinning compound figures. Then, we show a revised inferential pattern. This depends on an *indication*, at the commencement of the inference, that makes salient that the speaker's overall communicative intention is not metaphorical but ironic. Thus, compound figures forcefully show the difficulty of drawing an inference without such indication.

5 Naïve Temporal-MPT

Let us return to our toy example in (2). The speaker displays a piece of handwriting, perhaps a student essay, covered in blotches and ink stains, with poor lettering, and utters (2). What are the processing steps that a hearer goes through in reconstructing the speaker meaning? A naïve Temporal-MPT predicts three processing stages.

First, syntactic-semantic processing must extract the literal (sentence) meaning that there is a piece of lacework that is delicate. The hearer must then match this linguistic understanding to the salient referent. However, the referent is not a piece of lacework. There is a mismatch between linguistic information and salient contextual information (here visual information).[13] One hypothesis is that the utterance is not literal but metaphorical. This is a quick local fix of word meaning ("lacework") to remedy the contextual mismatch. To establish the metaphor requires seeing the handwriting through the lens of a piece of embroidery, i.e. matching similarity properties of delicacy, craft, artistry, etc. onto properties of the handwriting in the essay. But such a straightforward match also initially fails because there can

be no match between the two sets of properties. Rather, the contrast between the delicacy of lacework and the indelicate visual appearance of the handwriting guides the hearer to infer that the utterance is not only metaphorical but also ironic. It's ironic about a metaphorical characterisation of the essay. We can thus characterise the processing order implemented by a naïve Temporal-MPT:

> What delicate lacework > What delicate handwriting > What indelicate handwriting
> Literal > Metaphorical > Ironic Metaphorical

This sequence shows two problems. First, the inferential process to derive speaker meaning is computationally costly. It involves many computational steps which make processing effort taxing: the hearer first posits that the utterance is metaphorical, then tries to find matching properties, but instead finds contrasting properties; then posits that the metaphor is not used for its own sake but rather put in the service of making an ironic point.

Second, the hearer cannot establish the exact metaphorical content until they know that this content is to be further inverted. Thus, finding *contrasting matches*, rather than straightforward matches, is what guides the metaphorical processing such that the metaphor can function as a suitable target for ironic ridicule. This means that positing an ironic intent is necessary in order to draw specific conclusions about the metaphorical content. This is because the metaphor understanding is filtered through the hypothesis that the metaphor is only instrumental to achieving ironic purposes. If we are to take the view that metaphor is said-content, this is odd because the inference of the metaphorical content we settle on depends on the prospective knowledge that the implicature is ironic. But this in turn is supposed to depend on knowing what is said/asserted. We have a chicken and egg problem.[14]

We may draw the following conclusions. First, to complete the derivation of the metaphorical content requires that the hearer make a hypothesis about a *prospective ironic implicature*. Second, the ironic implicature depends on what is said metaphorically, so it will only be derived after the metaphorical content is settled on. Therefore, naïve Temporal-MPT is not psychologically plausible.

---

[13] I'm not suggesting that metaphor always works by finding literal mis-matches. There are well discussed cases of truth-apt metaphors that involve no mismatch (see Hills 1997). Though this example relies on visual information being salient, that need not be the case: other contextual information may make the target properties salient.

[14] This is a familiar problem at the heart of semantics/pragmatics debates concerned with the nature of pragmatic intrusion intro truth-conditional content. For models circumventing this problem, see Levinson (2000), Carston (2002), Recanati (2004), Bezuidenhout (2015), Popa-Wyatt (2017, 2018), among many others.

6 Signalling and Communicative Channels

There is an alternative model. This is based on the observation that communication uses multiple channels from various sources of signalling.[15] When making an utterance a speaker uses intonation, facial expression, gesture, and body-posture to communicate. These signals are neither independent of the sequence of words used, nor merely the handmaidens of words, there only to disambiguate the correct interpretation of the utterance. Information sent through different channels is instead coordinated in order to increase the salience of relevant information. Your intonation, facial expression and gesture are coordinated both with one another and with the sequence of words you utter, thus guiding a computationally efficient recovery of speaker meaning.

The fact that we should allow communicative channels other than words alone to inform our linguistic behaviour has already been noted by Grice (1989, p. 216). His theorizing of "utterance" as encompassing "any candidate for [non-natural] meaning", i.e. including non-linguistic behaviour produced for communicative purposes, a major step in the right direction, despite philosophers' reluctance to draw on computational and psychological considerations. We should be able to remedy this.

Thus, the hypothesis is that it is not only reasonable to suppose that multiple communicative channels may be recruited during utterance production, but also that it plays a critical role in utterance interpretation. This makes sense insofar as speakers and hearers take turns in their roles of signals producers and interpreters. They are, in a sense, co-communicators and the meaning production and interpretation is in fact a dove-tailing process. As we learned from Grice, in forming communicative intentions the speaker relies on the hearer's ability to infer those intentions, and in interpreting an utterance the hearer relies in turn on the speaker's capacity to exploit this ability.

Our hypothesis is that multi-modal signalling can in fact boost this dove-tailing. Speakers may signal, through various cues, salient information guiding the hearer to retrieve her communicative goal. Hearers in turn are attuned to such cues which enable them to integrate various bites of information from multi-modal signals, as they become available. This is likely to increase computational efficiency, as information is integrated early on. Such considerations are often taken for granted, and thus underplayed in discussions of linguistic behaviour, including those concerning ironic utterances.[16]

This is despite the fact that irony and sarcasm in particular recruit a distinctive intonational contour (i.e. marked by slow rate, exaggerated stress, and nasalisation; see Rockwell 2000; Bryant and Fox Tree 2005). The purpose of intonation is thus to flag the speaker's *ironic intent* in a way that is independent of the full-fledged derivation of *ironic content*.

To bring the point home, it is also worth stressing evidence from the developmental literature on understanding of metaphor and irony. Winner (1988), Winner and Gardner (1993, p. 442) argue that whereas with metaphor it is possible to recognise the utterance as metaphorical, while being unsure about the exact content of speaker's meaning, nevertheless recognising the utterance as ironical is ensured merely by the recognition of the speaker's ironic intention:

> "[In metaphor] it is as if the listener said, "I know the speaker is being metaphorical, but I do not know what [s]he is getting at." In the case of irony, once one recognizes that the speaker is being ironic, there is a click of comprehension and the speaker's meaning is grasped. It is difficult to imagine thinking, 'I know the speaker is being ironic but I just don't know what [s]he means.' Rather, one is more likely to think, 'Oh, now I understand. He was being ironic!'"

We contend it is thus useful in this regard to distinguish between (a) *grasping the communicative intent/point* of the utterance—i.e. determining *that* an utterance is intended to be interpreted in a specific way (e.g. metaphorically or ironically), and (b) *content-derivation*—i.e. determining *what* full-fledged content of the utterance is conveyed. From a computational point of view, it is also reasonable to suppose that (a) can precede and be independent of (b). Applied to compound figures, we can hypothesise that grasping the ironicity of the utterance (i.e. that it's used to make an overall ironic point) may *temporally precede* the determination of both the metaphorical and ironic contents.

This changes the models that we might consider computationally efficient and thus psychologically plausible. If we suppose that supplementary channels (e.g. intonation, gestures, rolling eyes and other forms of behavioural mimicry) provide additional cues that signal the speaker's communicative intent, then armed with this information the hearer's processing may be significantly eased.
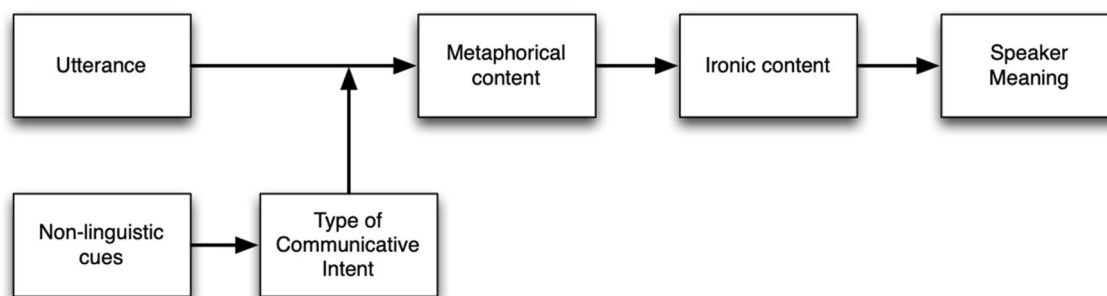
The claim is that if an indication of the type of communicative intent is signalled upfront (here ironicity), this can then be used to constrain the subsequent processing of the component meanings in a compound figure. We submit that a multi-modal system of communication can help reduce

---

[15] For game-theoretic models of signalling, see Skyrms (2010) as well as recent developments in game-theoretic pragmatics (Benz et al. 2006).

[16] In obviously ironic utterances—where the utterance clearly sits in contrast to the known facts—there is less need for cues such as sarcastic intonation, and skilled users of irony may minimise these extra

---

Footnote 16 (continued)

linguistic cues in order to slow the listener's inference of the ironic implicature and heighten the effect.

the inferential process in that information about a prospective ironic implicature can work *backwards* to modulate the metaphorical processing to a smaller sub-set of metaphorical properties which can then serve as object of ironic ridicule.[17] It is arguable that the metaphorical search is constrained to only those matching metaphorical properties which can in turn yield relevant contrasting properties under a further ironic interpretation. Thus, rather than seeking matches, the metaphorical processing is instead prompted to settle on *contrasting matches*. Critically, searching for contrasting matches is not the same as searching for mismatches.[18]

On this hypothesis, the recognition of ironicity that is being signalled through non-linguistic channels *precedes* the content-determination captured by Logical-MPT. In other words, ironicity is *prior* in terms of the structure of speaker's communicative intentions, and thus can constrain in a more efficient way the derivation of the full-fledged metaphorical and ironic contents. This gives us the following model:

Suppose we accept this modified model of processing ordering. What does it buy us? Essentially, it buys us a computationally efficient processing in that we do not have to employ a complex abductive reasoning to infer the speaker's communicative intent. In short, we need not infer *what* specific ironic content is communicated prior to inferring *that* the utterance is primarily geared to achieving ironic purposes. This is not incompatible with Logical-MPT: the logical priority of metaphorical content over ironic content is indeed retained. However, by adding an additional level of ironicity signalled through various communicative channels, we gain in computational efficiency.

7 Communicative Acts Structure

What do these computational considerations tell us about the process of forming communicative intentions? As you recall, Grice's definition of communicative intentions (see Sect. 2)

has been criticised as being psychologically implausible.[19] We've also seen that the inferential abduction required by a naïve implementation of Temporal-MPT is computationally too complex to be plausible. Such complexity is increased in compound figures because the speaker has not just one communicative intention but two. But intentions are not atomic entities; they have a structure. So it's important to clarify how those intentions are related and what role each figure plays in the compound.

One hypothesis is that when metaphor and irony merge into a compound the speaker is not committed to being metaphorical. She is not using the metaphor to make a full-fledged metaphorical act, as with self-standing metaphorical utterances. Instead she is using the metaphor merely as a

---

[17] This contrasts with processing of self-standing metaphors which often invite an open-ended search of metaphorical properties, especially when the metaphor is novel as in (1)–(5). An interesting approach which stresses the imaginative dimension of metaphorical thinking is proposed by Lepore and Stone (2015). There it is argued that metaphor and irony are not aspects of communication but involve instead an "imaginative engagement that needn't have a single goal" (2015, pp. 4–5), so "any conclusions the audience thereby discovers are implicit and tentative suggestions, rather than transparent and public contributions" (2015, p. 39). While it is true that open-ended creative exploration is what makes metaphors powerful conceptual tools, when they are used for ironic purposes the processing would arguably be needlessly complicated.

[18] There are many non-salient but possible mismatches between the properties of lacework and the handwriting. Lacework contains holes and is made of thread, but those properties do not match in this case. In another case they might. Suppose a person mockingly held up a piece of poor sewing, full of holes. Here these properties would provide matches, with which the term 'delicate' would then contrast to the visual referent. This would be non-metaphorical irony.

[19] An alternative, "minimally Gricean" model of communication is proposed by Moore (2017, p. 8) where he recasts Grice's publicity requirement in terms of a "functional constraint" on intentional communication. For more psychologically-oriented accounts, see Millikan (1984), Bar-On (2013), Green (2007).

basis for achieving other goals, i.e. being ironic about commitments characteristic of metaphorical acts. What role does the metaphor then play within the compound? We argue that the metaphorical act is not made in a serious (committal) mode but under pretence (or used to echo someone else's metaphorical claim or act), in order to draw attention to the fact that it falls short of expectations. Thus, the metaphorical act is recruited as an object for ironic ridicule. This explains that the metaphorical act (and its underlying communicative intention) is in fact nested inside an ironic act and intention. We can thus give the structure underpinning an ironic metaphoric act as follows:

In uttering (2), the speaker intentionally:

(i) engages in the behaviour characteristic of someone committed to "The handwriting is delicate lacework" (literal act);

(ii) engages in the behaviour characteristic of someone lacking the intention in (i), but who performs (i) with the intention to assert that the handwriting is beautiful, shows skill, etc. (metaphorical act);

(iii) engages in the behaviour characteristic of someone lacking the intention in (ii), but who performs (ii) with the intention to ridicule someone who would assert that the handwriting is beautiful (ironic act);

(iv) communicates that she has the intention presented in (iii) (full-fledged ironic act).

The ironic metaphoric act involves a layering of acts. First, there is a literal assertion (i), which is incorporated in a metaphorical act (ii), which in turn is incorporated in an ironic act (iii). Whereas the intentions underpinning the acts under (i)–(iii) are not committal,[20] the speaker's communicative intention pertains only to the outermost act (iv), which is a full-blooded ironic act having a metaphorical act nested inside it. This explains that making an ironic metaphor compound amounts to giving a ridiculing portrayal of a metaphorical act, by showing that it is inappropriate in the context. The speaker is thereby expressing two kinds of commitments:

(a) a ridiculing attitude towards anyone who would be committed to a metaphorical act,

(b) a belief that some form of inversion of the metaphorical act is the case.

The speech-act structure underlying ironic metaphor compounds also explains that in terms of communicative intentions, the recognition of ironicity is prior and that a metaphorical act is merely instrumental to achieving a more complex ironic act whose target is the metaphor itself. How does the speaker signal this structure? As we've argued in Sect. 6, multiple communicative channels such as intonation, gesture, etc., can guide processing by flagging early on the speaker's ironic intent and thereby signalling that the speaker lacks the commitments characteristic of a metaphorical act and is using it instead for ironic purposes.

8 Conclusion

Where does this leave us? Our contention is that if we are prepared to incorporate multi-channel information this can then simplify the processing of compound figures. What we have shown here is that the introduction of multiple communicative channels shows that there is a psychologically plausible solution that is computationally feasible. Moreover, the structure of the communicative acts underpinning the compound figure is also captured. Finally, given that irony is prior in terms of communicative intentions, it can ease processing by constraining the metaphorical processing to contrasting metaphorical matches which are then more easily interpreted ironically. An open question is whether this phenomenon is more widespread. Is it plausible that in many utterances a separate channel conveys information that gives strong clues as to the type of communicative intention?

References

Barker SJ (2004) Renewing meaning. Oxford University Press, Oxford

Barker SJ (2017) Figurative speech: pointing a poisoned arrow at the heart of semantics. Philos Stud 174:123–140

Barker SJ, Popa-Wyatt M (2015) Irony and the dogma of force and sense. Analysis 75(1):9–16. https://doi.org/10.1093/analys/anu104

---

[20] One possibility is to think of the acts in (i)–(iii) as various kinds of proto-acts (see Popa-Wyatt 2017, Barker 2004, 2017).

Bar-On D (2013) Origins of meaning: must we 'go Gricean'? Mind Lang 28(3):342–375

Benz A, Jäger G, Rooij RV, Van Rooij R (eds) (2006) Game theory and pragmatics. Palgrave Macmillan, London

Bezuidenhout A (2001) Metaphor and what is said: a defense of a direct expression view of metaphor. Midwest Stud Philos 25:156–186

Bezuidenhout A (2015) The implicit dimension of meaning: ways of "Filling In" and "Filling Out" content. Erkenntnis 80:89–109

Bryant G, Fox Tree J (2005) Is there an ironic tone of voice? Lang Speech 48:257–277

Camp E (2006) Contextualism, metaphor, and what is said. Mind Lang 21(3):280–309

Camp E (2012) Sarcasm, pretence, and the semantics-pragmatics distinction. Noûs 46:587–634

Carston R (2002) Thoughts and utterances. Blackwell, Oxford

Carston R, Wearing C (2015) Hyperbolic language and its relation to metaphor and irony. J Pragmat 79:79–92

Currie G (2006) Why irony is pretence. In: Nichols S (ed) The architecture of the imagination. Oxford University Press, Oxford, pp 111–133

Davidson D (1982) Rational animals. In: Davidson D (ed) Subjective, intersubjective, objective, 2001st edn. Oxford University Press, Oxford, pp 95–105

Green MS (2007) Self-expression. OUP, Oxford

Grice P (1989) Studies in the way of words. Harvard University Press, Cambridge

Grice P (2001) Aspects of reason. In: Richard W (ed). Oxford University Press

Hills D (1997) Aptness and truth in verbal metaphor. Philos Top 25(1): 117–153. https://doi.org/10.5840/philtopics199725118

Lepore E, Stone M (2015) Imagination and convention: distinguishing grammar and inference in language. Oxford University Press, Oxford

Levinson S (2000) Presumptive meanings: the theory of generalized conversational implicature. MIT Press, Cambridge

Millikan RG (1984) Language, thought, and other biological categories: new foundations for realism. MIT Press, Cambridge

Millikan R (2004) Varieties of meaning. Massachusetts Institute of Technology, Cambridge

Moore R (2017) Gricean communication and cognitive development. Philos Q 67(267):303–326

Neale S (1992) Paul Grice and the philosophy of language. Linguist Philos 15:509–559

Origgi G, Sperber D (2000) Evolution, communication and the proper function of language. In: Carruthers P, Chamberlain A (eds) Evolution and the human mind: language, modularity, and social cognition. Cambridge University Press, Cambridge, pp 140–169

Planer RJ (2017) Protolanguage might have evolved before ostensive communication. Biol Theory 12(2):72–84

Popa-Wyatt M (2009). Figuring the code: Pragmatic routes to non-literal. University of Geneva, Geneva

Popa-Wyatt M (2017) Compound figures: priority and speech-act structure. Philos Stud 174(1):141–161

Popa-Wyatt M (2018) Embedding irony and the semantics/pragmatics distinction. Inquiry 62(6):674–699

Popa-Wyatt M (2020) Hyperbolic figures. In: Athanasiadou A, Colston H (eds) On the diversity of irony. John Benjamins, Amsterdam and Philadelphia

Popa-Wyatt M (2020) Mind the gap: expressing affect with hyperbole and hyperbolic figures. In: Gargett A, Barnden J (eds) Figurative thought and language. John Benjamins, Amsterdam and Philadelphia

Popa-Wyatt M (2014) Pretence and echo: towards an integrated account of verbal irony. Int Rev Pragmat 6(1):127–168

Recanati F (2004) Literal meaning. Cambridge University Press, Cambridge

Rockwell P (2000) Lower, slower, louder: vocal cues of sarcasm. J Psycholinguist Res 29:483–495

Skyrms B (2010) Signals: evolution, learning, and information. OUP, Oxford

Soames S (2008) Drawing the line between meaning and implicature—and relating both to assertion. Noûs 42(3):529–554

Sperber D, Wilson D (1995) Relevance: communication and cognition, 2nd edn. Blackwell, Oxford

Steedman M (2000) The syntactic process (language, speech, and communication). The MIT Press, Cambridge

Sterelny K (2017) From code to speaker meaning. Biol Philos 32(6):819–838

Stern J (2000) Metaphor in context. MIT Press, Cambridge

Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: the origins of cultural cognition. Behav Brain Sci 28:675–691

Walton K (1990) Mimesis as make believe. Harvard University Press, Cambridge

Wearing C (2013) Metaphor and the scope argument. In: Penco C, Domaneschi F (eds) What is said and what is not said. CSLI Publications, Stanford, pp 141–157

Wilson D (2006) The pragmatics of verbal irony: echo or pretence? Lingua 116:1722–1743

Wilson D, Sperber D (2012) Meaning and relevance. Cambridge University Press, Cambridge

Winner E (1988) The point of words: children's understanding of metaphor and irony. Harvard University Press, Cambridge

Winner E, Gardner H (1993) Metaphor and irony: two levels of understanding. In: Ortony A (ed) Metaphor and thought, 2nd edn. Cambridge University Press, Cambridge, pp 425–445