

András Benedek / Kristóf Nyíri (eds.)

**Image and Metaphor
in the New Century**

Hungarian Academy of Sciences
Budapest University of Technology and Economics

András Benedek / Kristóf Nyíri (eds.)
Image and Metaphor in the New Century
Budapest:
Hungarian Academy of Sciences /
Budapest University of Technology and Economics
2019

© MTA–BME Open Content Development
Research Group, 2019
© The authors, 2019

The papers here collected have been accepted after a strict
double-blind peer-review process.

Cover design: István Ocztos

ISBN 978-963-313-307-1



Tibor Barany

Pictorial (Conversational) Implicatures

1. SS, SSSM, and Pictorial Implicatures

Considering the widespread use of pictures in ordinary communicative acts nowadays, it might be natural to think that our well-proven pragmatic theories developed for explaining how we communicate with linguistic expressions can be easily applied to pictorial communication. What might support this idea is the firm conviction that by and large, pictorial and linguistic communication proceed along basically the same lines.

On the other hand, there seems to be a deep asymmetry between the communicative use of words and of pictures. While it is reasonable to talk about *syntactic rules* operating on linguistic items (by which the syntactic structure of the sentence will be generated) and posit *compositional semantic content* or *conventional meaning* of sentences derived from conventionally encoded lexical word- (or morpheme-) meaning and conventional semantic rules for composition (which correspond compositionally to conventional syntactic rules), pictures appear to lack any kind of syntactic structure or encoded meaning-constituents. According to the Standard Model for Meaning Comprehension of Linguistic Utterances (for short: Standard Model, SM), the “total significance” of a speech act emerges as the outcome of four different types of processes:

- I. Working out the CONTEXT-INDEPENDENT CONVENTIONAL MEANING of the sentence uttered. The cognitive process of linguistic interpretation consists in *decoding* lexical meanings and the *computation* of semantic content as a function of the contents of the constituents and the way they are syntactically combined.

II. Working out the CONTEXT-DEPENDENT CONVENTIONAL MEANING of the sentence uttered. This process involves the mandatory contextual assignment of semantic values to various types of expressions – pronouns and indexicals, relative adjectives (like *tall*), and incomplete predicates (like *ready*), etc. – without which the utterance would remain semantically incomplete, lacking any determinate propositional content. Since the elaboration of this type of context-dependent meaning is governed by linguistic constraints, the cognitive process of interpretation is partly based on *linguistic competence*, because it involves knowing how to use conventionally certain expression-types (independently of whether the tokens of these types are overtly occurring in the sentence uttered or merely appearing in its logical structure while unpronounced), and partly on *pragmatic reasoning*.

III. Working out the PRAGMATICALLY ENRICHED MEANING of the sentence uttered, which is the result of optional pragmatic processes (free enrichment, broadening/loosening, predicate transfer etc.). These processes do the work of *truth-conditionally relevant* contextual modification of the proposition expressed; they can serve to “fine-tune” the proposition in order to ensure correct fit between the meaning-intentions of the speaker, the meaning-expectations of the audience, and the information conveyed by the utterance.

IV. Working out the CONVERSATIONAL IMPLICATURES (and other types of implicit meaning) of the sentence uttered. One can describe this process as the *derivation of additional propositions* (supposedly meant by the speaker) from the proposition expressed by the utterance, which is triggered by particular features of the conversational context and the speaker’s communicative behaviour (or on hypotheses thereof formed by the audience).¹ It should be noted that whilst it seems appropriate to

¹ Conversational implicatures are not inferences, as Kent Bach emphasizes (“The Top 10 misconceptions about Implicature”, in B. Birner & G. Ward [eds.], *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Sem-*

Pictorial (Conversational) Implicatures

say that these elements of meaning are conveyed by the utterance, they are not part of what the utterance literally means, so the elaboration of them does not affect the truth-conditions of the proposition generated as the outcome of the processes subsumed under I–III.

I dubbed this model “standard” because almost everyone agrees in the literature (with few exceptions) that by and large these kinds of processes – decoding, pragmatic reasoning: obligatory/optional enrichment, derivation of implicatures – take place when a language user understands a verbal utterance. However, there’s profound disagreement among philosophers of language over how to match particular types of meaning-elements with particular types of processes (e.g. are “scalar implicatures” pragmatically enriched meanings or generalized conversational implicatures?, etc.); how to define key concepts (e.g. in the case of conversational implicatures should what is meant be “completely separate” from what is said? [Kent Bach], as I characterized in the previous paragraph, in the spirit of the original Gricean program, or can conversational implicatures have truth-conditional consequences? [Stephen C. Levinson]; etc.); how to explain the relationship among various conceptual distinctions that underlie our pragmatic theories, such as “semantic”–“pragmatic”, “literal”–“non-literal”, “explicit”–“implicit”, and so on (we see massive proliferation of terminology in the literature, e.g. “explicature” [Relevance Theorists], “implicature” [Bach], etc.); and where to locate the theoretical boundary of *what is said* (between context-dependent conventional meaning and pragmatically enriched meaning, or somewhere else?). In short: if we fill in the gaps, SM will no longer remain quite as standard as it might have initially seemed.

antics in Honor of Laurence R. Horn, Amsterdam: John Benjamins, 2006, pp. 21–30), because they are defined as elements of speaker’s meaning, not hearer’s interpretation. However, how conversational implicatures are grasped by the audience, should be understood as some kind of inferential process.

However, for our present purposes, a much more schematic (let’s say, grossly oversimplified and highly contestable) model works well:

The Super Simplified Standard Model for Meaning Comprehension of Linguistic Utterances (SSSM): whenever an audience *A* understands a linguistic utterance *U* performed by a speaker *S*, *A* must grasp

- (i) *what is said* by *U*, which contains *decoding* lexical items, *compositional derivation* of sentence-meaning, *contextual completion* of sentence-meaning (viz. mandatory enrichment of semantically underspecified items in order to get complete proposition), and *contextual expansion* of sentence-meaning (viz. optional enrichment of the proposition expressed);² and
- (ii) *what is implicated* by *U*, which is triggered by *S*’s verbal behaviour (by the supposed tension between *S*’s being cooperative and her apparent violation of conversational norms or *U*’s lack of relevance at first sight).

Now, why should we think that SSSM can be applied to pictorial utterances, first and foremost, to cases of *genuine pictorial communication* (communicative exchanges performed solely by pictures, without any explicit “verbal accompaniment”)? Should we think it at all? I think the answer is “yes”. The fact that pictures do not have syntax and in interpreting pictures the audience normally does not decode visual items does not make the concept of *what is shown* (as analogous to what is said) theoretically useless or unintelligible. We can reasonably speak of “meaning (implicating) one thing by showing (a picture of) something else” – so there are no theoretical obstacles to maintain the distinction which underlies SSSM and the whole pragmatic machinery. How should we imagine such a theory? I sketch one version.

² By and large this is what Jennifer Mather Saul calls “constrained conception” of what is said. (*Lying, Misleading, and What Is Said*, Oxford: Oxford University Press, 2012, pp. 31–33.)

Pictorial (Conversational) Implicatures

First step: let's endorse some version of the resemblance theory of depiction, and assume that in normal cases visual features of the picture determine what it depicts. Resemblance theorists contend that the fact that pictures have a certain content (i.e. they represent, or are used by some people to represent objects, persons, events, actions, etc. from a particular perspective) should be explained by appeal to the visual resemblance relation between the picture and its object (rather than by appeal to the fact that pictures belong to some *conventional* representational system, as Nelson Goodman or more recently John Kulvicki suggested).

Second step: let's endorse (some version of) the theory of pictorial speech acts, which was developed as early as the 1970s by David Novitz and Søren Kjörup.³ (Interestingly, this work has not been pursued further. We had to wait until the 2010s for in-depth discussion of pragmatic phenomena connected to pictorial communication – mainly in relevance-theoretic terms; see for example the analyses of various types of “multimodal communication” by Charles Forceville.) According to this theory, the main difference between speech acts and pictorial acts lies in the role that conversational context plays in determining the “total significance” of the acts. In most cases performing a pictorial act can be described as expressing a proposition pictorially with some illocutionary force; in other words: someone who uses a picture for conversational purposes by showing or sharing the picture performs a particular pictorial propositional act and a particular illocutionary act (where the latter corresponds appropriately to the former).⁴ What makes a propositional act *pictorial* is the nature of predication: the utterer attributes to the referent some property that is visually represented in the picture. However, when performing a pictorial propositional act, the utterer cannot use convenient pictorial equivalents of linguistic devices such

³ See David Novitz, *Pictures and their Use in Communication*, The Hague: Martinus Nijhoff, 1977 (especially chapter IV and V: “Pictorial Illocutionary Acts”, “Pictorial Propositions”, pp. 67–107); Søren Kjörup, “Pictorial Speech Acts”, *Erkenntnis* 12 (1978), pp. 55–71.

⁴ See Novitz, p. 89; Kjörup, pp. 62–64 (propositional act), pp. 64–68 (illocutionary act).

as proper names, pronouns; negation, disjunction (and other Boolean operators); quantifiers; tense indicators – moreover, explicit illocutionary force indicators are also missing from the visual communicative apparatus. The explanation for this lack lies in the nature of depiction: elements of pictures are not endowed with syntactic structure, and they can contribute to the content expressed only via some sort of visual resemblance.⁵ But of course utterers *do* express quantified, tensed, negated propositions with a certain illocutionary force by showing/sharing pictures. In the cases of *multimodal communication* the missing pictorial elements are often substituted by linguistic elements (or other types of conventionally encoded devices); see e.g. internet memes (captioned pictures or videos), traffic signs (the red border as a shape signaling danger), etc. In the cases of *genuine pictorial communication* the “missing parts” of meaning become parts of the expressed proposition by contextual inferences drawn by the audience. These inferences are based on assumptions about the speaker’s communicative intentions and on her knowledge of the situation.⁶

⁵ As Alex Grzankowski puts it, “the semantic facts about depiction are determined at least in part by visual resemblance relations and since nothing *looks* to be disjunctive and nothing *looks* to not be the case, it’s unsurprising that pictures do not depictively express disjoined or negated propositions.” (“Pictures Have Propositional Content”, *Review of Philosophy and Psychology*, vol. 6, no. 1 (2015), pp. 151–163, the quoted passage on p. 159.)

⁶ Philosophers who deny that pictures have propositional content argue that pictures cannot be true or false, just more or less accurate (similar to perceptual experience), and we cannot assert or negate anything (or express disjoined propositions) by using only pictures; if we want to do all of this, we need to employ some kind of non-pictorial symbol, or to stipulate some *ad hoc* convention which relates aspects of a picture to negative facts, etc. See Tim Crane, “Is Perception a Propositional Attitude?”, *The Philosophical Quarterly* 59 (2009), pp. 452–469, esp. pp. 457–461; Richard M. Sainsbury, *Reference Without Referents*, Oxford: Clarendon Press, 2005, p. 242. Alex Grzankowski, *op. cit.*, points out that from pictures’ expressive limitations it does not follow “that all contents pictures *do* express fail to be propositions” (p. 158). Grzankowski is right: from pictures’ expressive limitations *concerning negation* (and disjunction etc.) the general inability to express certain propositions pictorially simply does not follow. But the problem, I think, lies deeper. If we consider *all types* of expressive limitations

Pictorial (Conversational) Implicatures

As Kj rup puts it picturesquely: “a picture may be construed as a heap of adjectives and other characterizing or predicating verbal phrases: ... is a man, ... is middle-aged, ... wears a tie, etc.”⁷ From these adjectives and other predicative VPs the audience should work out the propositional content of the pictorial act. Moreover, visual properties of pictures are also *filtered* by the context: the audience should pick out those properties that are relevant with respect to predication.⁸ For example by using this picture (Figure 1) the utterer might predicate different properties about different referents (a) in a Madrid travel guide, (b) in a book analyzing the European Union’s refugee policy, or (c) in a fashion blog. The three different tokens of the same picture-type share very few properties with respect to the proposition expressed – if they share any properties at all.

pictures *do* have, taking pictures in themselves – not only their inability to express negation, but the lack of pictorial quantifiers, pronouns, tense indicators, illocutionary force indicators, etc. – we arrive at this view: pictures *do not express propositional functions* (functions from contexts of utterance to propositions) *in a context-independent way*. To express complete propositions (with some illocutionary force) by conventional means requires not only conventional signs for representing objects in the world (and for indicating the illocutionary force of the utterance), but elements which perform functions such as quantification, forming tenses, building up syntactic structures – and that’s what is missing from pictures. To express complete propositions by conventional means, in a context-independent way, requires *more*, and pictures lack the adequate resources to perform the functions in question. On the other hand, Crane and Sainsbury are, it seems, adopting an untenable stance if they want to deny the possibility of genuine pictorial communication. Showing a photo or sharing it in the social media without any linguistic “accompaniment” sometimes does constitute a pictorial act. Non-propositionalist philosophers systematically underestimate the role of context in understanding pictorial communicative acts.

⁷ Kj rup, *op. cit.*, p. 63.

⁸ “... even though it is true that a picture must look like whatever it is a picture of, it does not follow from this that a picture attributes these looks to anything unless it is used to do so” (Novitz, *op. cit.*, p. 103).



Figure 1: City council of Madrid  Juanjo Martin, El Mundo.

It seems natural to think that this process is strictly analogous to the disambiguation of sentences, which is also based on pragmatic reasoning; both takes logical priority over interpreting context-independent compositional meaning.⁹

⁹ Catharine Abell in a recent paper draws a similar distinction between a picture’s *visible* content (viz. the content we would attribute to the picture on the basis of its perceptual properties) and its *depictive* content (the content we actually communicate by the picture), see “Pictorial Implicature”, *Journal of Aesthetics and Art Criticism*, vol. 63, no. 1 (2005), pp. 55–66. She proposes that the difference between visible content and depictive content be explained based on the difference between *what is said* vs. *what is implicated* when uttering a sentence. I am not sure her stance is tenable. What triggers an implicature (even a pictorial one) is something that has certain semantic properties on the basis of which (and with the help of other contextual clues) the implicature can be worked out by the audience. However, the visible content of pictures *by definition* lacks any determinate semantic property. Contrary to what the content-talk suggests, visible content is nothing more than the sum of the picture’s perceptual properties *perceived by* a potential spectator – without any representational nature. According to Abell, even “nondepictions” may have visible content (p. 56). Perhaps she *mislocates* the point from which pictorial implicature is derived: the audience of pictorial acts *do* employ pragmatic reasoning in picking out the relevant properties or in construing the depictive content of the picture – but this is not the same as deriving a completely new proposition from what is expressed by the utterer (because otherwise

Pictorial (Conversational) Implicatures

In sum: according to the theory of pictorial acts, in interpreting a pictorial act, the context (and our assumptions thereof) plays a much more significant role than in the case of verbal communication; “the context in which a pictorial illocutionary act is performed is to some extent a determinant of, or is responsible for, the propositional content of the act.”¹⁰

Third step: let’s apply SSSM to genuine pictorial utterances!

Super Simplified Standard Model for Meaning Comprehension of Pictorial Utterances (SSSM): whenever an audience *A* understands a genuinely pictorial utterance *U* performed by a speaker *S*, *A* must grasp

- (i) *what is said* by *U*, which contains *visual recognition* of properties on the basis of pictorial resemblance, *deducing* from the context (and from *A*’s prior knowledge, her assumptions concerning *S*’s meaning-intentions, etc.) the missing parts of the proposition – namely reference, tense, quantification, etc., and *contextual expansion* of the proposition expressed; and
- (ii) *what is implicated* by *U*, which is triggered by *S*’s communicative behavior (by the supposed tension between *S*’s being cooperative and her apparent violation of conversational norms or *U*’s lack of relevance at first sight).

Perhaps this solution to the problem of the application of our pragmatic explanation to pictorial communication strikes readers as a bit too easy; and they are right. The crucial point is whether we can hold on to the distinction, on the one hand, *between literal propositional content* expressed by showing or sharing a picture and, on the other hand, *conversational implicature* as a cancellable, non-detachable and calculable proposition (or set of propositions) triggered by the utterer’s apparent violation of conversational norms or the utterance’s apparent lack of relevance – despite the fact that pictures lack

the audience cannot secure the conviction about the utterer’s cooperative behaviour), which is characteristic of conversational implicatures.

¹⁰ Novitz, *op. cit.*, p. 92.

syntactic structure, therefore we cannot make sense of context-independent compositional picture meaning(-type).

The problem lies in the fact that “pragmatically enriched meaning” and “conversational implicature” are conceptually linked to “conventionally encoded compositional meaning”: one would think this is what can be pragmatically enriched and this is what can trigger a conversational implicature in appropriate circumstances – and this is what is unavailable in the case of genuine pictorial communication. In the next section I will phrase accurately the philosophical problem of pictorial conversational implicatures.

2. The Problem of Pictorial Conversational Implicatures

We have three propositions that are independently plausible and jointly inconsistent; I will call this “the problem of pictorial conversational implicatures”.

(Non-P) Anti-propositionalism: pictures do not have context-independent, conventionally encoded propositional content (propositional function).

(C) Only those representations can be used to convey conversational implicatures which have associated with them a context-independent, conventionally encoded propositional content (function).

(I) Pictures can be used to convey conversational implicatures.

There are three ways of responding to the problem: affirm (Non-P) and (C) while denying (I); affirm (C) and (I) while denying (Non-P); or affirm (Non-P) and (I) while denying (C).

Strategy I: *denying (I)*. This solution disallows the possibility of pictorial conversational implicatures. Communicating by pictures involves certain kinds of inferential processes, but strictly speaking these derived meanings are not conversational implicatures.

But I can see no clear reason for disavowing pictorial conversational implicatures. When I interpreted A’s photo in the chat win-

Pictorial (Conversational) Implicatures

dow in a particular way, I did the same as somebody would who is working out a conversational implicature on the basis of a particular verbal utterance and of some contextual clues. Without further arguments the proposed Strategy I is nothing more than denying the obvious.

Strategy II: *denying (Non-P)*. Employing Strategy II is the most popular way of solving the problem of pictorial conversational implicatures in the literature. Philosophers and linguists analyzing “multimodal discourse” often relativize the concept of “conventionally encoded content” to the conventions of the *visual genre* to which the picture belongs (for example: cartoon, commercial image, selfie, etc.). As Charles Forceville and Billy Clark puts it: “Understanding pictures requires knowledge of conventions of depiction as well as of genres that, even though pictures do not have a grammar or a vocabulary, suggest that we should broaden the concept of ‘encoding/decoding’.”¹¹

But I think there is a considerable difference between linguistic conventions to which we automatically and directly conform and conventions of pictorial genres the adherence to which is usually preceded by a conscious decision. Genre-conventions do not serve as adequate substitutions or analogies for linguistic conventions – at least they are not adequate for our theoretical purposes.

What I propose is the adoption of Strategy III: I think *we should deny (C)*. Nothing supports the idea that the propositional content which triggers a conversational implicature in a particular situation *must be* (partly) encoded conventionally by the representational device. All that we need is the propositional content that the audience might use as a clue (among other contextual clues) for working out

¹¹ Charles Forceville and Billy Clark, “Can Pictures Have Explicatures?”, *Linguagem em (Dis)curso* 14 (2014), pp. 451–472, the quoted passage on 469. See also Charles Forceville, “Relevance Theory as Model for Analyzing Visual and Multimodal Communication”, in David Machin (ed.), *Visual Communication*, Berlin: Mouton de Gruyter, 2014, pp. 51–70.

the conversational implicature in question – independently of the “origin” of that content.¹²

Consider the case of “serial implicatures”! We can put some conversational implicatures in sequences: that the hearer is expected to infer x from what is said, while the hearer can also reasonably infer a further conversational implicature y from x . This means a conversational implicature can serve as a vehicle to another conversational implicature. But there is no expectation that the vehicle implicature x be compositionally obtained! (For example: the speaker says that *It’s almost 7*. She conversationally implicates: “we’d better leave to catch our 8 o’ clock film at the movie theater”. This in turn implicates: “if you don’t make a decision fast as to what you’ll wear, we’ll be late for the movie”. This in turn implicates: “stop staring at your closet’s contents and put something on already”. Or: “we’d better call a cab soon so it’s here in time to get us to the movie theater”.)

In sum: conventionally encoded meaning need not serve as a departure point for conversational implicatures – and this holds for language and pictures alike.¹³

¹² Marcello Frixione and Antonio Lombardi also embrace this strategy, see “Street Signs and Ikea Instruction Sheets: Pragmatics and Pictorial Communication”, *Review of Philosophy and Psychology*, vol. 6, no. 1 (2015), pp. 133–149.

¹³ I am deeply grateful to Zsofia Zvolenszky for very helpful comments on earlier drafts of the present paper. This research has been supported by Grant No. K-116191 “Meaning, Communication; Literal, Figurative: Contemporary Issues in Philosophy of Language” received from the Hungarian Scientific Research Fund – National Research, Development and Innovation Office (OTKA–NKFIH).