Jürgen Bohnemeyer*

Demaximizing Grice?

DOI 10.1515/zfs-2016-0004

1 The trouble with maxims

Grice's (1975, 1989) theory of conversational implicatures is one of the most celebrated contributions of the 20th century to the philosophy of language, linguistics, and the neighboring disciplines. Contemporary theorizing of the relationship between language and logic arguably relies as much on it as on the works of Frege and Montague. Almost from the moment a wider academic audience began to embrace Grice's theory and realize its significance, one of its cornerstones began to inspire intense efforts to overhaul or replace it: Grice's set of conversational maxims (along with the classification of implicatures that derives from it). The maxims are one of two central explanatory tools of the theory, the other being the Cooperative Principle. From a logical perspective, a conversational implicature is a highly complex non-monotonic inference that involves five or seven premises, depending on the type of implicature (see below):

(i) The truth-conditional meaning p_{nn}[1] of the utterance U that triggers the implicature;
(ii) The sets of lexical items L = {I_1, I_2, ... I_n} and morphosyntactic constructions C = {C_1, C_2, ... C_n} of U used to express p_{nn};[2]
(iii) The (linguistic and situational) context *c* of U, modeled as a set of propositions;
(iv) A discrete set P = {p_1, p_2, ... p_n} of interpretations that are compatible with p_{nn} and *c*, but that are neither entailed nor presupposed by p_{nn};
(v) Sets of alternative lexical items L′ = {$I_{1'}$, $I_{1''}$, ... $I_{2'}$, $I_{2''}$, ... $I_{n'}$, $I_{n''}$, ...} and alternative constructions C′ = {$C_{1'}$, $C_{1''}$, ... $C_{2'}$, $C_{2''}$, ... $C_{n'}$, $C_{n''}$, ...} that seman-

1 *nn* here stands for 'non-natural', Grice's way of designating the conventional meaning of the utterance.

2 Or the proposition that U employs these constructions and lexical items, if one wants all premises to be propositions.

* **Corresponding author: Jürgen Bohnemeyer,** University at Buffalo – SUNY,
E-mail: jb@buffalo.edu

tically overlap with the elements of L and C and that could be used to modify U so as to express the elements of P;

(vi) The content of the conversational maxims;

(vii) Speaker's and hearer's assumption of mutual compliance with the Cooperative Principle, ensuring the application of the maxims to the interpretation of the utterance.

For the sake of convenience, the maxims and the Cooperative Principle are listed in (1):

458 zweimal (1)

(1) a. **The Cooperative Principle**
"Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." (Grice 1989: 26)
b. **The maxims of Quality**
"Try to make your contribution one that is true," specifically:
"Do not say what you believe to be false"
"Do not say that for which you lack adequate evidence" (Grice 1989: 27)
c. **The maxim of Relation (Relevance)**
"Be relevant" (Grice 1989: 27)
d. **The maxims of Quantity**
Q1: "Make your contribution as informative as is required (for the purposes of the exchange)"
Q2: "Do not make your contribution more informative than is required" (Grice 1989: 26)
e. **The maxims of Manner**
"Be perspicuous," specifically:
M1: "Avoid obscurity of expression"
M2: "Avoid ambiguity"
M3: "Be brief (avoid unnecessary prolixity)"
M4: "Be orderly" (Grice 1989: 27)

The inference resulting from the premises takes the form of a selection of that element of P that optimally satisfies the maxims. Grice distinguished two types of conversational implicatures, which he called (arguably somewhat unhelpfully) "particularized" (PCIs) and "generalized" (GCIs). Only GCIs, but not PCIs, make use of premise (ii) and (v). In other words, GCIs are sensitive to the particular constructions and lexical items employed by the utterance, whereas PCIs are not. Example (1) illustrates PCIs:

(1) It's cold in here.

Application of Relevance gives the hearer license to infer that the speaker, in uttering (1), was pursuing a communicative purpose related to the propositional content of (1). Some propositions in P that would satisfy this condition are those in (2) (using *S* to denote the speaker and *H* for the hearer):

(2) a. S uttered (1) to make small talk
    b. S uttered (1) to politely suggest to H to close the window
    c. S uttered (1) to encourage H to make a pass on S

GCIs take into account the form of the utterance in a broad sense. Consider (3):

(3) Floyd ate some of the cookies.

Suppose H knows that there were six cookies in the jar before Floyd had a go at it, but does not know how many Floyd ate. The following propositions are all in P:

(4) a. Floyd ate two of the cookies
    b. Floyd ate three of the cookies
    c. Floyd ate four of the cookies
    d. Floyd ate five of the cookies
    e. Floyd ate many of the cookies
    f. Floyd ate most of the cookies
    g. Floyd ate all of the cookies

All of these are more informative than (3) (in the sense that they are compatible with a smaller set of possible worlds than [3]). And in each case, the informativeness contrast is attributable to the selection of a semantic quantifier, each of which has a specific expression that would replace *some* in (4). Let us assume, for the sake of argument, that all of these alternatives are equally salient.[3] If H now assumes that S applied Q1 (1d), it follows that S did not intend to convey any of the propositions in (4) – because otherwise (3) would not have

3 Alternatively, if *all* is in fact a significantly more salient alternative than the other competitors of *some* – as Franke's (2014) modeling of the data presented in Degen and Tannenhaus (2015) and van Tiel (2014) suggests – this will boost the probability of the 'not all' inference. Textbook discussions of scalar implicature typically only consider *all* as the sole relevant alternative to *some* anyway.

been sufficiently informative for realizing S's communicative goals. But since (3) entails (given H's prior knowledge) that the number of cookies eaten must be between two and six, H cannot simply assume that all of them are false. The interpretation that optimally satisfies the maxims in this case – and specifically Q1 – is that S remained purposely vague about the number of cookies Floyd ate. This makes (4a)–(4d) and (4g) implausible candidates for the best estimate of what S had in mind, since they narrow down the intended number of cookies to a single specific number. Which leaves (4e) and (4f). Again, Q1 licenses the inference that these are not intended by S. Therefore, H infers that the scenario S intended to describe in uttering (3) is one in which the number of cookies eaten remains vague, but is most likely less than half of the total number. And S can count on H arriving at this interpretation.

If H does not know how many cookies were in the jar at the outset, the result is the same, except H will now arrive at the conclusion that the likeliest intended referent is less than half of however many cookies were in the jar. Comprehension data reported in Degen and Tannenhaus (2015) and van Tiel (2014) (according to the description in Franke 2014) are in line with this prediction.

The informal computation of the scalar implicature I just sketched for (3) is different from that suggested in the target article (Franke and Jäger this volume; henceforth, F&J) and in Franke (2014). It makes different assumptions. I discuss the differences in Section 4. But it also suggests that the computation of conversational implicatures does indeed rely on something akin to the rationalistic probabilistic reasoning proposed by F&J and the literature they summarize.

The attractiveness of this approach (or this family of approaches, given the variation in assumptions across the existing efforts) lies in two properties:

- The ability to support quantitative modeling of experimental data and to generate testable quantitative predictions;
- The prospect of being able to account for conversational implicatures, and to correctly predict their occurrence, without the need to postulate conversational maxims.

My commentary here is centrally concerned with the second feature. No other aspect of Grice's theory has received as much critical attention as Grice's proposed set of maxims.[4] Prominent reconstruction attempts have proposed reduc-

4 The runner-up is perhaps the psychological plausibility – or lack thereof – of the reasoning mechanism implied by the theory. Speaker and hearer are assumed to weigh a potentially considerable number of semantic and expressive alternatives within an interval measured in milliseconds. But since the conclusion does not change for a given expression as far as GCIs are concerned (except for the effect of context) as long as the semantic and morphosyntactic tion of the set of maxims – or, in the case of Levinson (2000), at least the subset that licenses GCIs – to three principles (Levinson 2000), two (Horn 1984), or even just one (Sperber and Wilson 1995). Skepticism vis-à-vis Grice's set of maxims is readily motivated. To begin with, the large number of maxims seems unparsimonious. It is difficult of find another example of a scientific theory that postulates a set of nine fundamental principles of near-axiomatic status, all of which are entirely domain-specific. Occam's Razor favors theories that can account for the same set of observable phenomena – in this case, the implicatures – equally well invoking a smaller number of principles of more general currency.

Secondly, the composition of the set of maxims is not derived from first principles within Grice's theory, but instead from the table of categories Kant (1929 [1781]) postulated in his *Critique of Pure Reason*. Categories in the Aristotelian and Kantian sense are types of predicates (alternatively, in more contemporary terms, types of information, or types of representations). It appears that Grice turned to Kant in a bid to capture any aspect of the composition of an utterance through which the speaker could make his or her communicative intentions guessable to the hearer. However, the maxims express neither this idea that unifies them – the goal of making implicated meanings recoverable – nor the cognitive and computational mechanisms by which this goal can be achieved. As a result, the theory is unable to offer compelling arguments for the exhaustiveness of the set. In the case of the submaxims, Grice (1975: 46) himself observes that "one might need others" as far as the Manner maxims are concerned. Similarly, Grice himself points out that principles similar to his conversational maxims govern nonverbal interactions as well (Grice 1975: 47). A great example is Norman's (1988) cognitive theory of industrial design, whose principles bear striking resemblance to those of Grice's theory (although Norman does not cite Grice). Yet, the maxims are stated exclusively for verbal interactions.

Lastly, the acquisition of conversational implicatures by children would seem a daunting task if it required learning a set of maxims that is large and not straightforwardly derived from first principles. There is indeed evidence suggesting that the adult-like computation of conversational implicatures emerges relatively late by linguistic standards, although this question continues to be a matter of much debate (Guasti et al. 2005; Noveck 2001; Papafragou and Musolino 2003; *inter alia*). However, there are obvious cognitive reasons for this lateness independently of the content of the maxims, including the required

properties of the expression and its competitors remain constant, it can presumably be stored in long-term memory.

computational effort and the dependence of the concept of communicative intentions on theory of mind. In any case, the learner's task should be proportional to (or at any rate a function of) the relative simplicity or complexity of the theory of conversational implicatures. This alone justifies attempts to push for a simpler theory than the one sketched by Grice in a bid to understand the true extent of the demands on the learner.

In the remainder, I first consider why probabilistic pragmatics seems a more promising approach than the most acclaimed previous attempt at doing away with the maxims, Sperber and Wilson's (1995) relevance theory (Section 2). I then address the arguably most significant obstacle for explicitly Bayesian approaches to implicatures such as that sketched by F&J: the extension from scalar to other types of implicatures (Section 3). Finally, I discuss two rather minor objections I have to F&J's proposal.

2 Why probabilistic pragmatics may outperform relevance theory

Relevance theory (RT) models implicatures as inferences that result from deductive reasoning guided by a single general principle, the "principle of relevance". In any deductive reasoning attempting to uncover the intentions underlying a communicative act, the principle of relevance favors those inferences that maximize information gain while simultaneously minimizing computational effort. Sperber and Wilson present the principle of relevance, not as a maxim rational agents will follow in order to realize their goals, but as an inherent part of the cognitive processes involved in what they call "ostensive-inferential communication", which they characterize as follows: "The communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a certain set of assumptions" (Sperber and Wilson 1987: 700).

The most important consequence of dispensing with conversational maxims in RT is the lack of a specific mechanism for deriving implicatures from the linguistic properties of the utterances that trigger them. Consequently, there is no distinction between GCIs and PCIs in RT, and there are in some sense only PCIs. There is thus for example no context-independent account of scalar implicatures. A relevance theorist might point out that sentences identical to (3) except for minimal contextual adaptations may trigger rather different inferences in different contexts. For instance, in (5), where the context establishes universal quantification as an explicit alternative, the use of *some* will indeed implicate the negation of universal quantification. That is the most informative, and thus "relevant", interpretation in this context. Not so, however, in (6), where the question whether Floyd ate all of the cookies is not at issue. What matters in (6) is whether Floyd ate any cookie. Universal quantification not being a salient contextual alternative, its negation plays a much less prominent role in the implicature (cf. Noveck and Sperber 2012 and the literature summarized there). If Floyd turns out to in fact have eaten all cookies, the speaker's utterance will likely be considered misleading in (5), but not in (6).

(5) A: Did Floyd eat all of the cookies?
B: He ate SOME of the cookies.

(6) A: The cookies were poisoned! Anybody who ate even one cookie needs medical attention immediately.
B: FLOYD ate some of the cookies.

While it unquestionably makes an important observation, this RT account still does seem to understate the generality of scalar implicatures. For example, it seems to predict that B's response in (7) should be perfectly inconspicuous even if Floyd actually ate multiple cookies. Yet, this is not so, at least not in my intuition. The so-called "upper-bounding" scalar implicature, according to which B's answer implies that Floyd ate exactly one cookie, still seems to take effect here, despite the apparent low contextual relevance of this interpretation.

(7) A: The cookies were poisoned! Anybody who ate even one cookie needs medical attention immediately.
B: FLOYD ate a cookie.

Consider also the neutral context in (8):

(8) A: What did people eat at the party last night?
B: Sally had a burger. Floyd ate some of the cookies. Harriet ate only grapes.

It seems that it will still be understood in (8) that Floyd did not eat all of the cookies, and in fact probably did not even eat most of the cookies, despite these alternatives not being particularly salient in this context.

Probabilistic pragmatics seems the superior theory here. It correctly predicts the relative generality of scalar implicatures. Unlike RT, it is able to do so despite its disuse of maxims, because it has a mechanism that captures the impact of the relative informativeness of a given trigger vis-à-vis its narrow expressive competitors. Q1 is effectively built into the Bayesian mechanics of F&J's account, which explicitly evaluates alternative expressions for their relative informativeness. And the Cooperative Principle translates into the assumption that rational speakers will strive to maximize the "expected utility" of their communicative choices, by choosing expressions that will make their communicative intentions optimally recoverable (i.e., inferable) for the hearer. At the same time, probabilistic accounts also seem optimally suited to model context-induced gradation effects such as the contrast between (5) and (6).

3 But does it generalize?

F&J's Bayesian evaluation engine is designed to estimate the communicative utility of expressions by comparing them to plausible alternatives in terms of their relative informativeness. It is able to predict scalar implicatures without explicitly referring to Grice's Q1 because Q1 stipulates the goal of optimizing informativeness for the purposes of the exchange, and F&J's algorithm is designed to do just that. But this immediately raises the question whether – and if so, how – a probabilistic pragmatics (PP) account can be extended to other types of conversational implicatures. Contemporary pragmatic theory recognizes three such types: stereotype implicatures based on Q2, Manner implicatures based on the Manner maxims, especially M1 and M3, and PCIs. Four scenarios must be distinguished:

(i) F&J's algorithm predicts all four types of implicatures without adaptation.
(ii) While F&J's algorithm cannot in its present form account for all four types, a modified version of it can.
(iii) No modified version of F&J's algorithm correctly predicts all four types of implicatures. However, it is possible to construct different PP algorithms for each of the four types.
(iv) A PP account of some or all of the other three types of conversational implicatures beside the scalar ones is impossible.

I have sorted these in the order of the implications they hold for the goal of constructing a maxim-free theory of conversational implicatures. Scenarios (i) and (ii) would clearly realize this goal; scenario (iv) would fail it entirely; and scenario (iii) would mean that a reflex of the maxims would survive in the form of the distinction among the different algorithms. My view is that the reality most likely lies somewhere in between (ii) and (iii).

Scenario (i) appears to be clearly false. The mechanism F&J describe, following Franke (2014), is specifically designed to predict that interpretation of a given utterance which a rational speaker is most likely to have intended in uttering it assuming she were attempting the utterance to be optimally informative given the intended meaning. And it is making this prediction based exclusively on one kind of evidence: the relative informativeness of the utterance vis-à-vis its closest competitors. None of the other types of implicatures can be derived in this fashion.

Stereotype and manner implicatures trade off on one another. Despite stereotype implicatures being attributed to Q2 (Atlas and Levinson 1981), the dimensions along which they do this are the domain of the Manner maxims: length, complexity, and frequency.[5] Stereotype implicatures assign defeasible stereotype interpretations to simple, short, high-frequency expressions, whereas manner implicatures assign non-stereotypical interpretations to expressions the production of which is more than minimally effortful. Consider for example the contrast between lexical and syntactic causative constructions: the former pick up stereotypically direct causation (9a), the latter the absence of stereotypicality and thus indirectness (9b) (McCawley 1976, 1978).

(9) a. Sally stopped the car.
b. Sally caused the car to stop.

To account for this distribution, one would need an algorithm similar to the one F&J describe, which however would evaluate expressions, not in terms of their relative informativeness, but in terms of their length, complexity, and frequency. On the semantic side, typicality and atypicality could presumably be assessed in terms of frequency alone. These models would assign optimal expected utility to high-frequency (i.e., stereotypical) interpretations of relatively simple, short, high-frequency expressions, but also to low-frequency (i.e., atypical) interpretations of non-minimally long and complex expressions and to uncommon ones. Lacking even the most passing familiarity with the niceties of Bayesian models, I see no obvious reason to assume that such an algorithm would hold significantly more challenges for the modeler compared to the one described in Franke (2014).

As for PCIs, non-conventional indirect speech acts such as those discussed in Section 6 of F&J are in fact an example of PCIs (compare the discussion of

5 I follow Levinson (2000) closely in assuming three major types of GCIs and the trade off between stereotype and manner implicatures.

[1]–[2] above). I assume, then, that a solution to modeling PCIs is not only possible, but may in fact already exist.

To determine in which world we are vis-à-vis the distinction between scenarios (ii) and (iii), a single system would need to be built that combines the capabilities of predicting the four types of implicatures. In the best-case scenario from the perspective of the goal of a maxim-free implicature theory, this is possible without the integrated algorithms interfering with one another and degrading each other's performance.

However, even if this relatively ideal outcome can be achieved, we would still be left with a system that has been designed specifically to derive exactly the kind of inferences licensed by the currently assumed set of maxims – and no others. Should a previously unrecognized type of conversational implicatures be discovered at some point in the future, the system would have to be redesigned again to accommodate it. Thus, a reflex of the maxims is perpetuated as part of the design of the PP inference engine. The theoretical goal of dispensing with the maxims by deriving them from first principles will have been met only partway.

4 Quibbles: Rationality, experimental paradigm, and modeling assumptions

Here I would like to briefly register two minor disagreements with F&J. In no particular order:

F&J endorse a notion of rationality according to which it applies to any behavior that is "optimally adapted to solve [sic!] a particular purpose" (p. 6). While I am all in favor of demystifying rationality, this characterization strikes me far too weak. Any evolved process is by definition (somewhat) optimally adapted. Thus, any evolved instinctual behavior would be rational by F&J's definition. I would prefer to restrict the notion of rationality to strategic choices made in pursuit of goals. This would not deny nonhuman animals rationality – but it would exclude from rationality purely instinctual behavior where evolution makes the choices for the individual, so to speak. At the same time, however, we should of course guard against the trap of overlooking the continuity between evolved instinctual and rational non-instinctual strategic behavior.

My second point of contention concerns the experimental paradigm F&J and Franke (2014) rely on to model the derivation of conversational implicatures. In the experiments of Degen and Tannenhaus (2015) and van Tiel (2014), participants rate the goodness of fit between utterances and scenarios on Likert scales. However, we should probably assume that an actual Gricean hearer derives her non-monotonic inferences in the face of uncertainty about the scenario. For Gricean interactions to regularly succeed, the speaker must be able to more or less reliably anticipate the inferences the hearer will most likely derive without knowing the actual scenario the speaker intends to communicate. The (informal and yet admittedly cumbersome) derivation of the scalar implicature I sketched in Section 1 tries to take this into account.

5 Conclusions: The glass is more than half full

For the goal of a maxim-free theory of conversational implicatures, the news out of F&J is mixed. On the downside, the PP model does not succeed in removing the content of the maxims entirely from the theory by deriving them from first principles. On the upside, F&J make a compelling case that the reasoning engine that generates Gricean implicatures is nothing other than the very general ability to reason about conditional probabilities. And in doing so, they at the very least minimize the role of the maxims, by reducing them to predictions concerning the types of evidence Gricean speakers and hearers should apply their Bayesian reasoning powers to.

References

Atlas, Jay David and Stephen C. Levinson. 1981. It-clefts, informativeness, and logical form. In Peter Cole (ed.), *Radical pragmatics*. 1–61. New York: Academic Press.

Degen, Judith and Michael K. Tanenhaus. 2015. Processing scalar implicatures: A constraint-based approach. *Cognitive Science* 39(4). 667–710. DOI: 10.1111/cogs.12171.

Franke, Michael. 2014. Typical use of quantifiers: A probabilistic speaker model. *Proceedings of CogSci* 2014. 487–492.

Franke, Michael and Gerhard Jäger. This volume. Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft* 35(1). 3–44.

Grice, Paul Herbert. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics III: Speech acts*. 41–58. New York: Academic Press.

Grice, Paul Herbert. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Guasti, Maria Teresa, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini and Luisa Meroni. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20(5). 667–696.

Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q- and R-based implicature. In Deborah Shiffrin (ed.), *Meaning, form, and use in context*. 11–42. Washington, DC: Georgetown University Press.

Kant, Immanuel. 1929 [1781]. *Critique of pure reason*. Transl. and ed. by Norman Kemp Smith. London: Macmillan.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: Cambridge University Press.

McCawley, James. 1976. Remarks on what can cause what. In Masayoshi Shibatani (ed.), *Syntax and semantics VI: The grammar of causative constructions*. 117–129. New York: Academic Press.

McCawley, James. 1978. Conversational implicature and the lexicon. In Peter Cole (ed.), *Syntax and semantics IX: Pragmatics*. 245–258. New York: Academic Press.

Norman, Don. 1988. *The design of everyday things*. New York: Basic Books.

Noveck, Ira A. 2001. When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition* 78. 165–188.

Noveck, Ira A. and Dan Sperber. 2012. The why and how of experimental pragmatics: The case of 'scalar inferences'. In Dan Wilson & Deirdre Sperber (eds.), *Meaning and relevance*. 307–330. Cambridge: Cambridge University Press.

Papafragou, Anna and Julien Musolino. 2003. Scalar implicatures at the semantic-pragmatics interface. *Cognition* 80. 253–282.

Sperber, Dan and Deirdre Wilson. 1987. Précis of *Relevance: Communication and cognition*. *Behavioral and Brain Sciences* 10(4). 697–754.

Sperber, Dan and Deirdre Wilson. 1995. *Relevance: Communication and cognition*. *Second edition*. Oxford: Blackwell.

van Tiel, B. 2014. *Quantity matters: Implicatures, typicality, and truth*. Nijmegen: Radboud Universiteit dissertation.