Fat Jokes and Their Measurable Implicature:
An Analysis of Grice

JOSHUA HENSTROM

> Descartes invites his date, Jeanne, to a Michelin-starred restaurant for her birthday. The sommelier hands them the wine list, and Jeanne plumps for the most expensive Burgundy on the list. "I think not!" exclaims an indignant Descartes, and *POOF* he disappears (Burton).

Jokes, witticisms, and humor are pervasive linguistic acts that make communication enjoyable. Without them, the wells of friendship and human connection would dry up since all communication would be excruciatingly bland and solely informational. Punchlines are the vibrant color to the canvas, the delectable dessert after dinner, or the jazzy solo at a concert. These linguistic delicacies are what make conversation interesting. Though philosophically speaking, the literature surrounding humor is not very substantial, it is a growing field that has struck many modern thinkers

Joshua Henstrom graduated from Brigham Young University in August of 2024 with a BA in Philosophy and a minor in Legal Studies. He is currently applying to law schools and plans on working in malpractice, AI, and criminal litigation realms of law. His philosophical interests include legal, political, ethical, and metaphysical studies. This essay took third place in the David H. Yarn Philosophical Essay Contest.

building upon the ideas of older philosophers such as Kant, Aristotle, and Plato.[1]

In this paper, I will focus mainly on the linguistic applications of humor. I will look at how they may apply to Grice's Cooperativity Principle and maxims in section I. In section II I will outline Raskin's and Attardo's arguments who have claimed that humor has no place in the Gricean *bona fide* model of language. They argue that a separate sphere of argumentation is necessary to even begin to analyze humor, the *non bona fide* model. I will argue that humor is compatible under the Gricean model. However, to give credence to Raskin's and Attardo's arguments, I will argue that humor is compatible under the Gricean model by giving an example of a humorous statement that does not violate or flout any of the maxims nor Grice's Cooperative Principle. In section III, a non-violating calculably implicated joke will be assessed as following Gricean principles. I will then visit some possible rebuttals from Raskin and Attarado in section IV which include the possibility that either a) the joke is out of the possible realm of truth or b) the purported joke is not a calculable joke under the *non bona fide* model. These rebuttals will be addressed in sections V and VI respectively.

I. Grice on the Cooperativity Principle

Grice aims to separate that which is implicated and that which is said within his language model. This model presupposes that conversation is a cooperative endeavor. As such, whenever two interlocutors engage in conversation, they base all conversational presuppositions and their possible implications on the Cooperativity Principle (CP): "Make your conversational 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 1975, 26). In short, when we are engaged in a conversation with another individual, we are aiming for our transmitted ideas to be understood by the other party as clearly as possible. When one engages with CP, they follow certain maxims found under the categories of Quantity (#), Quality (Q), Relevance (R), and Manner (M) (see Grice 1975, 26–7). Each category has features that will be outlined below.

---

[1] See Hobbes (1996); Descartes (1989); Scruton and Jones (1982); Lambert Deckers (1993); Spencer (1911); Freud (1960); Beattie (1779); and Schopenhauer (1818).

Under #:

> 1. Make your contribution as informative as is required (for the current purposes of the exchange).
> 2. Do not make your contribution more informative than is required.

Under Q:

> 1. Try to make your contribution one that is true (the super maxim).
> 2. Do not say what you believe to be false.
> 3. Do not say that for which you lack adequate evidence.

Under R:

> 1. Be relevant.

Under M:

> 1. Don't be perspicuous (the Supermaxim).
> 2. Avoid obscurity of expression
> 3. Avoid ambiguity.
> 4. Be brief (avoid unnecessary prolixity).
> 5. Be orderly.

However, it is quite clear that when one engages in conversation, not all of these maxims need to be followed. Someone can: a) violate a maxim quietly so the other party does not notice the violation, b) opt out of a maxim or cooperative principle, c) resolve a clash of maxims by choosing one more pertinent than the other, or d) flout the maxim in such a way that the hearer understands that the maxim is being flouted, leading to implicature, or those conversational statements that exist implicitly above the *explicit* conversation. In the case of flouting, when one obstinately fails to follow a maxim, the hearer, who is still under the assumption that CP is still being followed, may question why maxim A was flouted, or how a statement *P* is relevant to the conversation. This connection between a seeming lack of relevance leads the hearer to obtain implicit knowledge through conversational implicature rather than conventional meaning. For example,

> (1)   A: *Smith doesn't seem to have a girlfriend these days.*
>   B: *He has been paying a lot of visits to New York lately*
>   (Grice 1975, 32).

> (2)   A philosophy professor is writing a testimonial about a pupil who is a candidate for a philosophy job, and his letter reads as follows: "*Dear Sir, Mr. X's command of*

*English is excellent, and his attendance at tutorials has been
regular. Yours, etc.*" (Grice 1975, 33).

In the case of (1), no maxim was violated nor flouted. However,
under the assumption that B is following the maxims under CP, specifically
R.1, when speaking of Smith not having a girlfriend, it follows that *B*'s
contribution is implying that Smith might have someone of interest in
New York. In the case of (2), #.1 is clearly flouted. The lack of information
in the letter of recommendation indicates to the reader that either a) these
are the pupil's only good attributes, or b) the teacher has nothing good
to say about the pupil but wants to remain civil. As shown, and as Grice
argues, implicature, especially under this model, is calculably separate (i.e.,
found separately from the explicit language by finding maxim flouts and
violations) from conventional or explicit meaning (39).

II. Raskin's and Attarado's Non-Bona-fide model

Grice makes no mention of humor within his model. As such, Raskin
and Attarado aim to fill in the possible holes formed by the use of humor.
Humor is a strange occurrence that, according to Raskin and Attarado, does
not follow CP. They determine that Grice's CP works only under the realm
of "no nonsense mode as a mode in which the speaker and the hearer are
mutually committed to the truth" which requires straightforward answers
(Raskin and Attarado 1994, 32). While CP uncomfortably accommodates
violations to the maxims, conversations allow for the individual to violate
some maxims as long as they keep the other maxims. From this they
determine that in general there are three attitudes towards all language in
CP: a) superficial and essential abidance of all of the maxims, b) essential
abidance of some of the maxims while flouting/violating others, and c) an
utterance is devoid of cooperativeness altogether. Gricean's *bona fide* model
follows (a) and/or (b), metaphor follows only (b), lies are only uncooperative
or (c), and humor violates (a)–(c) (33–4).

Humor falls outside of the *bona fide* model because it fails to hold
any truth value. When one tells a joke, the hearer, initially believing the
speaker is following CP, will feel confused because the hearer's attempts
to apply the *bona fide* model to the joke seemingly failed. The hearer then
looks for alternative means to interpret the proposition put forth, namely
under the *non bona fide* model (36). However, there is nonetheless a sense
of cooperativeness when it comes to humor, as humor is not the same as
lying. As such, Raskin and Attarado create *The Cooperative Principle for Joke
Telling*:

1. Maxim of Quantity: Give exactly as much information as is necessary for the joke.
2. Maxim of Quality: Say only what is consistent for the world of the joke.
3. Maxim of Relation: Say only what is relevant to the joke.
4. Maxim of Manner: Tell the joke efficiently (37).

Raskin and Attarado do concede that it is tempting to try to include humor within the Gricean model as a "possible world of a joke." However, "*possible worlds* is a well-defined technical term wholly subordinated to the logical concept of truth . . . [humor] operates in its own world, which is not a possible world in the strictly philosophical sense . . ." (37; 65). Truth in humor is not necessary; however, truth is necessary only for the *bona fide* model as Raskin and Attarado define it. Therefore, humor must be considered as a separate entity from CP in hopes of computationally understanding it. We will go into those computational assessments in section VI.

III. Norm Macdonald: The Non-Violating Calculable Joke

Before we get into any analysis of Raskin and Attarado's claims, let us first take a brief intermission[2] and consider the following joke:

(3)   Norm: *Do you know what my doctor told me? I went to my doctor, right? I say to my doctor, "Doc, what's going on?" The doc responds, "You are fat." And I say, "How fat am I?" And he says, "You are so fat, that if you don't stop you are going to have a heart attack or a stroke."*

Person B: *"That's not funny."*

Norm: *"Well he's not a comedian!"*

Person B: *"I was expecting a joke."*

Norm: *"From my cardiologist?"*

In the metaphysical sense, the joke of this joke is that *there is no joke*. In short, it is an *implicit joke*. However, conventionally, this statement is not considered funny *per se*. It seems as if the joke is not what is explicitly

---

[2] Philosophy, though enjoyable, can be tiring to read. I hope the joke will give the reader a quick respite from the technical philosophy.

said but what is implicated and only what is implicated. If there is any sort of implication in language, then it will be calculably separate from the conventional meaning under Grice's model. What is truly peculiar, however, is that this joke does not violate or flout any of the maxims, as Raskin and Attarado argue all humor does. This should also be calculable.

Under #, it provides the necessary information to get the point across while also not providing too much information. For example, Norm does not say, "The triglycerides will clog your arteries, stopping your heart and all other blood flow to your body, leading to clinical death." We have no reason to believe that he is violating or flouting Q for he is adequately fat enough for his doctor to make such a statement. We also have no reason to believe that Norm is saying something that he believes to be a lie. Norm follows R because what he says is relevant to the conversation at hand. Lastly, he follows M, for Norm is "clear [as to] what [his] contribution is... and [executes] his performance with reasonable dispatch" (Grice 1975, 28). This is because he is fat and he will die if his current lifestyle continues. There is no obscurity nor lack of orderliness because Norm is clear and straight to the point. So, if none of the maxims were violated, then can conversational implicature be pulled from such a clear statement?

We must look at the initial setup of the joke. As is commonly understood, "How fat am I?" is a common start to a regular joke. By the phrasing of his statement, and Person *B*'s affinity to believe that Norm is keeping CP, it is assumed that Norm is going to tell a joke by flouting maxims found in Q. For example, a possible doctor's Q violation response to Norm's question could be: "You are so fat that when you skip a meal, the stock market crashes." Clearly, this example exaggerates to the point that the statement is no longer true, which would fall into Raskin's and Attarado's *non bona fide* territory. However, Norm does not do that in (3), but in fact says something that could be true and fails to flout any maxims. As such, there is an *incongruity* as to what Person *B* thinks will occur and what does occur. The joke's language *implicates* a "You're so fat" joke and yet, it is not completed in that *regular* way. This follows into a very mainstream theory about humor with which Raskin concurs with: the Incongruity Theory. The Incongruity Theory dictates that we laugh at "something that violates our mental patterns and expectations," and Norm's "joke" violates Person *B*'s expectations while also following all the

maxims (See Morreall 2020[3]). Therefore, humor is possible within CP, proving Raskin and Attarado wrong.

IV. Rebuttal by Attarado and Raskin

There are two possible objections that Raskin and Attarado may consider after my assessment of the implicit-joke: either a) that my assessment did not include the presuppositions of the conversation or b) the implicit-joke mentioned is not a joke as they have outlined in their own computational theory.[4] In the case of (a), though a truth may have been spouted by Norm to Person B from a recent doctor's visit, those truths do not add to the current common ground (Dynel 2008, 176–179). It could be that the person, in light of the current topic, does not take the information seriously, or takes the information to be within the *non bona fide* realm. As such, even though (3) did not flout the maxims in any way, they would argue that the information provided does not add to the conversation. Norm was not being serious, and his language was not to be taken nor was computed to be taken that way. So, even though the joke does follow all the maxims, Person B was already "switched" to a *non bona fide* calculation of Norm's statement. In this sense, CP was violated before any maxims were flouted.

In the case of computational humor, or objection (b), Raskin and Attarado outline six Knowledge Resources. They might argue that the very joke outlined in (3) is not in fact a joke as expressed within their *non bona fide* model. As such, (3) should be considered a mere statement of information in the real world that made someone laugh. The Knowledge Resources are Script Oppositions, Logical Mechanisms, Situations, Target, Narrative Strategies, and Language, which outline their General Theory of Verbal Humor (GTVH) from the top of the hierarchy to the bottom of the hierarchy (54).[5] By breaking down common jokes, such as the Lightbulb

---

[3] I do not wish to go into the novelty of the incongruity model nor expand upon what has already been expressed. Raskin already agrees with the model and so there seems no reason to explore every model at the moment.

[4] The potential objections are just that, potential. (i) and (ii) seem to contradict, and if Raskin and Attarado took both positions, they would be standing on unequal footing. I am arguing against these rebuttals not by claiming they would take both sides, but that either side is an equally viable option for them to take.

[5] In a hierarchy of language, changes to those nodes at the top, for example script opposition, will affect what actions can be taken lower in the hierarchy, for example language. If I switch from a Script Opposition that is actual/nonactual to sexual/nonsexual, I have changed what type of jokes I am able to successfully make.

Joke, they can distinguish key items that make a joke a joke (45–50). These variables are an attempt to mathematically create jokes. In this sense, a joke that has outputs in every category determines whether this is verifiable *non bona fide* humor. Definitions for the nodes are as follow:

> *Script Oppositions*: When a bit of text "is compatible with its opposite in a special predefined sense" (50). It is when a text violates the expectations of the hearer through twisting reality to its opposite, for example, taking a possible statement and turning it into an impossible one with a punchline.

> *Logical Mechanisms*: Makes it possible to combine two bits of opposing text together. This can be through "false analogies," "juxtapositions," or "garden path phenomena" (52). This is the clear designator that something is or isn't in the *non bona fide* sphere of language.

> *Situations*: "Every joke must be about some story," and that story "provides the props for the joke" that can either be relied upon to tell the joke or be completely ignored.

> *Target*: The butt of the joke or the target of which the speaker wishes to make fun. However, this KR can have an empty value if the joke is considered to be "non-aggressive."

> *Narrative Strategies*: The very framing of the joke, such as a question/answer format, a narrative, a riddle, an aside in a conversation, etc.

> *Language*: The very language of the text that separates a serious statement in the text from the punch line of the text. The distinction here between *bona fide* and *non bona fide* language is the very punchline found usually at the end of the text (51; 53).

Under these new nodes, Raskin and Attarado would argue that at face value, (3) is merely a serious statement. They would argue that (3) fails to fill some of the necessary Knowledge Resources values, namely Script Opposition, Logical Mechanism, and Language. It is only by coincidence or by misinterpretation that Person B finds the statement to be humorous.

1. No value for Script Opposition because the script presented by Norm only speaks in actual or possible language and fails to switch towards its opposites.

2. No value for Logical Mechanism because there is no language necessary to combine opposing scripts as there is no Script Opposition value.

3. No value for Language as there is no punch line. Therefore, the language here is *bona fide*.

V. Counterargument: Revisiting the Calculability of Norm's Joke

So, it has been shown that (3) does not violate any of the maxims as laid out by Grice. However, as argued by Raskin and Attarado, it does not matter. Anything deemed humorous violates the very CP that Grice's theory depends upon. The main reason Raskin and Attarado do this is to separate serious language from non-serious language. When a joke is stated, I as the listener am not supposed to consider the language as something truthful or factual. However, there is an issue when CP is not honored: "[the hearer] no longer has any clue as to the speaker's commitment of truth…and is limited to the literal of the speaker's utterances…and is barred from making any inferences from these meanings" (Raskin and Attarado 34). Now, this does not mean there are no rules whatsoever as one must take into consideration Raskin's and Attorado's Humorous CP, but it does mean one has left the realm of "reality" or the realm of literal communication.

However, this leaves Raskin and Attarado in a hole. If one cannot be humorous while also following CP, then one cannot implicate in one's joke. By making this distinction, one either can implicate in the world of the *bona fide* or be unable to implicate in the world of the *non bona fide*. It then follows that these worlds of language must be mutually exclusive. If they are mutually exclusive, then merging the two modes of language is impossible. However, it seems one can implicate and script-switch (or have opposing scripts connected through logical mechanisms and punchlines) at the same time as shown with conversational humor and quips (i.e., *Friends*).

(4)   Chandler: There is something different about you.

Ross: I went to that tanning place your wife suggested.

Chandler: Was that place the SUN?!

Initially, the conversation starts off normally, maybe how someone would expect a conversation to go among friends. However, the script switches to 'abnormal' in the end when Chandler invokes the sun as what tanned Ross.[6] However, this is all within the conversational setting that they have. Ross is the one that came in looking burnt to a crisp and uneven in all places. It is through Chandler's exaggeration that Ross finds out his tan is not good at all. What Chandler says can be described as a hyperbole, which Grice uses as a Q flouting example:

(5)   *Every girl loves a sailor* (Grice 1975, 34).

It is not true that *every* girl in the world loves a sailor, very much how Chandler stating that Ross looks like he went to the sun doesn't actually mean he went to the sun. However, what is implicated (that Ross looks terrible) gets across. However, under the *non bona fide* model, implication is not allowed as the hearer is not able to make heads or tails of what the speaker is saying. But there is a switch to the *non bona fide* model nonetheless. This is implicated by a script opposition, which is clearly marked out by Chandler's punchline. In this sense, (4) is *both non bona fide* and *bona fide* language, for how else are we to know the implications of Chandler's quip? Clearly, it is possible to linguistically exist in both worlds.

VI. Counterargument: Measuring Norm's Implicit-Joke in GTVH

Clearly, I have my work cut out for me. Though Raskin's and Attarado's GTVH is consistent with formulaic humor, it seems on a face-value approach, (3) is simply a serious statement within the Gricean *bona fide* realm. And this is true: if I were to take the text of (3) to be the only linguistic information available, the statement above seems more depressing than actually funny. However, what failed to be part of a face-value computational assessment is the underlying feature of the joke, or the implicit joke, that is stated in the background. It is this that leads Person B to correctly laugh at a joke rather than mocking Norm's potentially life-threatening diagnosis. Understanding this will also help us not only interpret jokes that are humorous at face-value, but also explain why statements intended to be funny can have very plain language.

As proved before with (4), jokes can both implicate and script-switch. This is what makes (3) possible. To consider (3) void of these implications frames Person B as an insensitive individual. So, in order to calculate

---

[6] I take no joy in explaining jokes. This is absolute torture.

(3), we must take the underlying implicature of the joke as well into the equation:

> Script Opposition: As Raskin and Attarado have argued, (3) has no script opposition and therefore is not a joke. However, this analysis is void of Norm's implicature. By setting up his phrasing, he is setting up an SI that is to tell a fat joke. In a *non bona fidian* sense, Norm is setting up to be non-cooperative and indicating to the hearer that a joke is about to occur. In the process of not completing the expectation of an explicit script opposition, Norm creates through implication a script opposition between the expectations of Person B and what was explicitly said. Though this type of script opposition is not listed, it is nonetheless valid under the definitions given by Raskin and Attarado. I will name it the **non-meta/meta** SO with a **sub-type of non-serious/serious**.

> Logical Mechanisms: Raskin and Attarado have argued that (3) has no logical mechanism because there is not an SO to combine. However, as argued above, there is one that is implicated. But what is the logical mechanism that combines these oppositional ideas? **Implicated juxtaposition.** Combining the ideas of an implicated joke with an explicit non-joke. Clearly, they are at odds with one another, and they are combined through **garden path phenomena**, in which Person B defines the content of the Norm's question as a joke, and then is confused at the end by the lack of one.

> Situation: **Fat joke**. This is a very common phrase used to insinuate a joke, very much like "How many X people does it take to change a lightbulb?" sets up a lightbulb joke.

> Target: **Himself**. The joke is a fat joke that targets himself.

> Narrative Strategies: **Question/answer**. Norm poses both the question and answer to his joke.

> Language: As Raskin and Attarado would argue, there is no punchline. That is because the punchline is implicated. As stated before, by phrasing the question "How fat am I?" Norm sets up Person *B* with the

> expectation that there is a punchline. With no punchline
> at the end of the statement, Person B is confused as to
> why there wasn't a punchline. The punchline is not
> explicitly stated. It is implicitly stated in that the joke of
> (3) was that there was no joke.

Clearly, (3) is a joke, and as shown before, it follows CP. By being both a joke and following CP, I can conclude that Raskin and Attarado were wrong in their assessment of the Gricean model.

VII. Conclusion

In this paper, I have argued that Raskin and Attarado's *non bona fide* language model is both unnecessary and too restricting. Their attempts to separate serious speech from non-serious speech leaves their own assessment of humor lacking, failing to consider humor that is solely based in implication. I first showed an example of a computational joke that follows all the Gricean maxims while also remaining consistent with CP. I also addressed and successfully argued that it is possible to follow *non bona fide* language as well as *bona fide* language, disproving the mutual exclusivity of either model. I then analyzed (3) under GTVH to prove in fact that it was a joke.

However, that does not mean their model is without merit. GTVH is clearly a working theory that allows one to plug in new points and create novel jokes. However, it is limited to that which is only explicitly stated within the *non bona fide* world. As has been shown, implicature is central to many different jokes, such as a quip or a joke pretending to be a serious statement. This leaves me wondering if GTVH can be successfully integrated into a Gricean model without leaving the realm of CP, which would allow implications created by maxim flouting to be included in GTVH's joke assessment. I would go even as far as to rename the whole theory and name it as the General Theory of Verbal and *Implicated* Humor (GTVIH) to emphasize that point. However, that is clearly another topic for another paper.

Works Cited

Beattie, James. *Essays: on poetry and music, as they affect the mind: on laughter, and ludicrous composition; on the usefulness of classical learning.* London: Printed for E. and C. Dilly and W. Creech, Edinburgh, 1779.

Burton, Neel. "Top 10 Philosophy Jokes." Psychology Today, Sussex Publishers, 2013, www.psychologytoday.com/us/blog/hide-and-seek/201303/top-10-philosophy-jokes.

Deckers, Lambert. "On the validity of a weight-judging paradigm for the study of humor." (1993): 43–56.

Descartes, René. *Passions of the Soul.* Hackett Publishing, 1989.

Dynel, Marta. "There is Method in the Humorous Speaker's Madness: Humour and Grice's Model." *Lodz Papers in Pragmatics*, vol. 4, no. 1, 1 Jan. 2008, pp. 155–89, https://doi.org/10.2478/v10016-008-0011-5.

Freud, Sigmund. *Jokes and their relation to the unconscious.* WW Norton & Company, 1960.

Grice, Herbert Paul. "Logic and conversation." Speech Acts, 12 Dec. 1975, pp. 41–58, https://doi.org/10.1163/9789004368811_003.

Hobbes, Thomas. *Leviathan.* Richard Tuck, ed. New York: Cambridge University Press 3 (1996): 415.

McDonald, Norm. "Norm Receives a Warning from His Doctor." YouTube, Norm McDonald Live, www.youtube.com/watch?v=NVt9HynzXW4.

Morreall, John. "Philosophy of Humor." Stanford Encyclopedia of Philosophy, Stanford University, 20 Aug. 2020, plato.stanford.edu/entries/humor/.

Morreall, John. "Verbal humor without switching scripts and without non-bona FIDE communication." *Humor: International Journal of Humor Research*, vol. 17, no. 4, 23 Jan. 2004, https://doi.org/10.1515/humr.2004.17.4.393.

Raskin, Jonathan D., and Salvatore Attardo. "Non-Literalness and Non-Bona-Fide in Language." Pragmatics and Cognition, vol. 2, no. 1, 1 Jan. 1994, pp. 31–69, https://doi.org/10.1075/pc.2.1.02ras.

Schopenhauer, Arthur. *The world as will and idea.* Vol. 3. K. Paul, Trench, Trübner & Company, Limited, 1906.

Scruton, Roger, and Peter Jones. "Laughter." *Proceedings of the Aristotelian Society, Supplementary Volumes* 56 (1982): 197-228.

Spencer, Herbert. "On the Physiology of Laughter: Essays on Education and Kindred Subjects." *London: Dent* (1911).