

## A type of sarcasm that current theories fail to explain – evidence from *sarchasm*

**Overview.** In this work in progress, I examine multiple instances from my experiment data that fall under the category of *sarchasm*, an utterance that is intended to be sarcastic but missed by the listener or overhearer (Fox Tree et al., 2020). Missed instances of sarcasm provide a unique window for thinking about the use and interpretation of sarcasm. I show that there is a particular subtype of sarcasm found in real data, which current theories of sarcasm (or verbal irony<sup>1</sup>) fail to explain. I propose a new framework that can address this phenomenon.

- (1) *Context:* Your friend was sure it would not rain today but you realize it is raining.  
*Response:* What a great day.

**Theories of sarcasm.** In the **Gricean theory**, sarcasm is identified when there is a blatant violation of maxim of quality. The response in (1) is therefore is sarcastic because the speaker is being untruthful. In **Echoic theories**, a speaker “echoes” (as opposed to “uses”) an utterance to convey a negative attitude. An echoic utterance alludes to the thoughts or utterances of others, which reminds the listener of norms or failed expectations and allows for the interpretation of sarcasm. (1) is sarcastic since the listener would know that the speaker is merely ‘echoing’ the previous thought that it was not going to be rainy, in order to express her negative attitude towards it. In the **Pretense theory**, a speaker (S) ‘pretends’ to be an alternative speaker (S’) speaking to an alternative listener (H’). S poses a negative attitude towards the utterance of S’, and H’ is ignorant and takes the utterance literally, while H understands it all. In (1), the speaker thinks that the weather is bad but pretends to be a person who thinks that the weather is good, and has a pretend-listener who would believe it and intends for the actual listener to understand all of it. In the **Implicit display theory**, sarcasm occurs if the speaker has an unmet expectation and conveys a negative attitude toward the failed expectation through the utterance. The speaker in (1) had an expectation that her friend’s belief would be true but expresses her negative attitude when the belief turned out to be wrong. It is not the focus of this work to discuss the limitations of individual theories. Instead, I show data that suggest that there is another type of sarcasm that current theory as a whole cannot explain.

**Data.** I use data collected from four (two production and two comprehension) online experiments. In each production experiment, participants (N=60 and N=128) were provided with contexts (N=32 and N=40), responded freely, and rated how sarcastic their responses were from 1: *not at all* to 6: *completely*. In each comprehension experiment, new participants (N=360 and N=512) rated how sarcastic they found the same responses as external evaluators. Neither speakers nor evaluators were given sarcasm definitions in order to obtain natural data. I selected the instances to which the speakers gave the highest sarcasm rating (6), which I consider as having sarcastic intent. Of 584 such instances, I selected the ones that external observers gave lower than 4 on average (*sarchasm*). I have identified 251 such instances and show examples below.

**Limitations of previous theories.** In (2), the speaker points out how blind Steve is to his own flaw by bluntly pointing it out to him.

- (2) *Context:* Steve has a brother Bill. Bill often feels annoyed by his friend. The reasons that Bill finds his friend annoying are the same as the reasons why you find Steve annoying (for example, both Steve and Bill’s friend always ask for money and never pay it back). Steve says, “why is my brother even friends with that guy? I don’t get it.”  
*Response:* Well you should know, shouldn’t you?

---

<sup>1</sup>I treat *sarcasm* and *verbal irony* synonymously following recent work. The default terminology is *sarcasm*. See Fox Tree et al. (2020) and references therein.

The maxim of quality has not been violated (Gricean theory), nor does the speaker provide any echoic utterance (Echoic theory). If the speaker was engaged in a pretense (Pretense theory), the real listener (H) would have to figure out that the utterance is sarcastic, but given how true and direct the response is, the listener would be faced with a garden-path situation at best. We do spot a failed expectation (Implicit display theory), which is that the speaker expects Steve to be aware of his own flaws at the presence of a similar example. But it is not obvious whether a direct remark would embed a negative attitude at the failure of expectation, which is required for an utterance to be sarcastic.

- (3) *Context:* Steve gives you a watering can on your birthday while smiling at you with a strange expression. But you don't even have a single plant.  
*Response:* Umm?? What's this for?

(3) provides a similar type of sarcasm: no violation of the maxim of quality, no echoing, no pretense. The failed expectation is also not clear in this case because even if the listener knows that the speaker does not expect a watering can, it is still possible that the speaker is just being unassuming and asking a genuine question. But the speaker still meant for the response to be sarcastic even though it is unlikely to provide the listener the cues necessary to interpret sarcasm, violating the cooperative principle (Grice, 1975). So how do we explain that such utterances are sarcastic?

**Proposal of a new framework.** I argue that sarcasm has a variant in which the speaker makes a reasonable remark in a direct manner but actually suppresses her desires to be more emotive, which often leads to *sarchasm*. The reason for muting emotion could be, among others, to save face (Jorgensen, 1996), avoid being rude (Dews et al, 1995), or keep the amicable relationship to the listener (Gibbs, 2000). The intentional suppression of attitude is deemed sarcastic by the speaker because she knows the underlying emotion behind the utterance, but the listener often misses it unless obvious or external cues are available. This type of sarcasm could be considered as 'reverse sarcasm', in which the speaker wishes to convey an attitude but (ironically) does so by being direct instead of choosing the literal/straightforward (emotionally strong) reaction.

- (4) *Context:* You are having a small party at your house. Steve, a little tipsy, starts mixing ketchup, mustard, potato chips, and orange juice and says "hey, look, I made something delicious!"  
*Response:* As long as you eat it buddy, you do you, and don't make a mess!

Then we can interpret the response in (4) as sarcastic. The speaker wishes to point out the silliness of Steve's behavior and does it by making reasonable requests, therefore muting her emotional reaction to him. If Steve also understands the silliness of his own behavior, he might get the sarcasm in the speaker's remark. Otherwise, it will likely become an instance of *sarchasm*.

**Implications.** The new proposal aligns with prior work that discusses the communicative functions of sarcasm (muting of criticism & face-saving). Sarcasm is used to subdue the criticism embedded in a message (Dews et al., 1995) or to save face by appearing less rude and fairer (Jorgensen, 1996). A new finding that emerged from the data I showed is that the muting of the negative message can go as far as turning an utterance into a direct remark that is reasonable given the context, and thus create a garden-path-like utterance for the listener. But as long as there is intentionally suppressed emotion behind the utterance, it will still count as intended sarcasm, but it will be missed by some listeners. The proposal I made in this work suggests that theory of sarcasm may need to separate intended and perceived sarcasm to thoroughly grasp the complexity of the phenomenon.