Grice’s Maxims in *Town of Salem*: Rules Made to be Broken

by Gabrielle Baum, Cañada College  
Mentor: Sarah Harmon

In the party game “Mafia,” a group of players sits or stands in a circle; one person is assigned the role of the Mafia, and another is assigned the role of the Guardian Angel. Everyone else is an average towns-person. While the townspeople and Guardian Angel have their eyes closed, the Mafia picks one person to kill, then closes his/her eyes. The Guardian Angel then opens his/her eyes and picks a person to save. If the Mafia and Guardian Angel picked the same person, that person lives. If not, that person dies (is out of the game), and everyone gets a chance to discuss who the Mafia is. After several rounds, if the Guardian Angel dies or if everyone else has died, the Mafia wins; if the Mafia is found out anytime beforehand and accused by at least two other people, the townspeople win.

Why is this relevant, anyway? Well, a similar diversion, Town of Salem (henceforth abbreviated ToS), is an online game that is different from most others and quite fascinating from a linguistic perspective. It resembles a more complicated version of Mafia with a little bit of the Salem Witch Trials thrown in. Since it is online, nobody has to close their eyes. Discussion and accusation take place via text chat; most roles cannot chat at night when they are doing their jobs, and even those who can chat can only talk with their targets, the dead, or their allies (depending on their assigned role). And there are over thirty roles, some of whom are townspeople (green), some Mafia (red), and the rest neutral figures (other colors) with no obligation to help the other factions, so nobody is just an average towns-person. In typical Mafia fashion, the game is practically made for lying; in fact, your survival may depend on it. Bad at lying? Well, your game partners are unable to see your nervous giggling and blushing as you claim to be a member of the town when you are actually a Serial Killer. *Everyone* lies in this game, which makes for hilarity, surprises, copious amounts of salt… and some interesting linguistic devices. As such, the purpose of this paper is to investigate these peculiarities, particularly the art of breaking linguistic rules - like Grice’s Maxims.
Previous research by others into this topic

Town of Salem is only two years old, so there has been no previous research regarding it. There has, however, been some linguistic research into other, notably older games: the most recently released games in most published research came out in 2010 (Drachen, 2011; Ensslin, 2012). Anders Drachen analyzed and compared player communication in three similar RPGs, one tabletop, one console, and one online (Drachen, 2011). While Drachen’s findings are not very relevant to this paper, Astrid Ensslin briefly discussed the role of lying and the “deliberate [flouting] of rules” (Ensslin, 2012) in various online games, including an online version of poker; she then compared this phenomenon to intentional rule breaking in language. “Although a large percentage of ludic and communicative behaviour follows the rules set out by any given language or game, human beings often bend or [flout] them, for a variety of reasons” (Ensslin, 2012).

Data collection

As I found out the hard way, video game developers do not easily relinquish chat logs, even with names effaced. For this reason, I chose ToS, which has a “Trial System” that displays chat logs from games in which people have been reported for bad behavior. This is so that experienced players can judge whether they deserve punishment (for example, a temporary ban from the game). I do not want to write a term paper filled with examples of bad apples, however, so I specifically picked games in which people were either wrongly accused or committed an offense irrelevant to gameplay (like picking an offensive nickname - and there is already a filter, so that is very much relative). These are essentially the same as the average game and so provide a wide array of situations and utterances to analyze.

A linguistic picture of Town of Salem

Before we continue, here are some observations about the linguistics of ToS as a whole. Many of these are shared with other online games.

Very short phrases, sentences, and moves are the norm, as can be seen in some of the snippets seen in this paper. This can be attributed to the fact that ToS’s phases of the day are on a timer. Players can only chat for up to 1 minute and 15 seconds at a time. Turn-switching is also frequent; this is because multiple players are typing furiously at the same time. This is not something you see in many other games; it is more common in text chat tools like Skype and
Discord, even though there is nothing resembling a time limit on these. Essentially, the in-game text chat is a lot more like actual speech than your average email.

There are universal gaming, texting, and meme-speak terms (“gg” for “good game,” clippings of common words like “u” for “you”, and coinages like “kappa” which means “just kidding”), as well as game-specific terms, like clipped or abbreviated names of roles: “Invest” for Investigator, “BG” for Bodyguard, “maf” for Mafia/Mafioso, “GF” for Godfather. There are also terms like “N1” for “night 1.” This game-specific lexicon does not spill out into real life; it stays in-game.

Here is one example of meme-speak in action, given by the Serial Killer who is on the verge of winning the game. However, the last townie has just been acquitted, as he would die the following night anyway. This “just for the heck of it” decision stalls the game unnecessarily. And so the Serial Killer says:

*McCree: y u do dis*

This phrase was originally posted to a meme site as part of the image below in Figure 1:

![Image of a kitten with “y u do dis?” text](https://example.com/meme.png)

*Figure 1: Who hurt him?!*

Meme speak, game-specific lexicon, and texting lingo figure prominently throughout the game chat, which begs the question of whether ToS players might constitute a veritable speech community. This thought will be revisited later, after examining a multitude of examples of “ToS-speak” in action in the indented, italicized chat excerpts below. In some cases, unclear terms or abbreviations have been explained in brackets.

**An introduction to Grice’s Maxims**

The philosopher Paul Grice proposed a concept called the Cooperative Principle, which says that “... when we communicate we assume, without realising it, that we, and the people we are talking to, will be conversationally cooperative - we will cooperate to achieve mutual
conversational ends. This *conversational* cooperation even works when we are not being cooperative *socially*. So, for example, we can be arguing with one another angrily and yet we will still cooperate quite a lot conversationally to achieve the argument” (University of Lancaster, 2005).

This concept can, according to Grice, be narrowed down into a brief list of linguistic “rules” that comprise the assumptions made by all parties in a conversation:

- Maxim of Quantity: be as informative as possible but no more
- Maxim of Quality: tell the truth
- Maxim of Relation: keep your reply relevant
- Maxim of Manner: be clear and brief, avoid ambiguity, and react appropriately

(University of Pennsylvania)

In Gricean (1975) terms, … a conversational maxim such as ‘be truthful’ may be violated on purpose to express irony. Saying ‘what a wonderful day!’ when the weather is absolutely terrible, for example, [flouts] Grice’s maxim of quality, but may cause [others] to chuckle or moan in agreement. They do so because they assume the speaker has followed the cooperative principle … and deduce from the context that the untruthfulness … is intended to be understood ironically (Ensslin, 2012).

Grice didn’t actually expect these to be hard, fast rules, like grammatical rules. He knew that they’d be bent or broken constantly, and that they’d serve as more of a framework for all conversational partners to deduce the real meaning of an ambiguous utterance - for example, sarcasm, or answering a question with a question.

Let’s take a look at each of the maxims in action - or rather, being violated - in ToS. In each of the examples below, all errors are theirs, not mine.

**Maxim of Quantity: Dodging the question, or indirectly answering it**

The Maxim of Quantity is an interesting rule to look at in ToS, basically because nobody has time to break it by overspeaking unless they try very, very hard. Here is an example of one Witch’s attempts to stay alive through reverse psychology while under pressure from the Jailor.

*Glitch*: role pls
*moisterraiserxd*: gf [Godfather] xdddd
*moisterraiserxd*: kappa ["just kidding"]
*moisterraiserxd*: im invest [Investigator] :)
*Glitch*: dont joke like that
...
*moisterraiserxd*: im invest
“Moisterraiserxd” says way more than he needs to - he could have just said something resembling “I’m the Investigator” and been fine. But he is trying to wackily engineer his way into being mistaken for a Jester in order to stay alive, and so he is purposefully jocular and repetitive in his speech.

In other cases, people may not respond with enough information in text, but they still manage to communicate their message. The situation below occurs at the very end of a game. “Jman” the Doctor and “Lights out” the Werewolf are the last ones alive. Jman has no hope of winning, as he cannot lynch Lights Out alone, and his two fellow town members went out in a blaze of idiotic glory the night before. (The Vigilante shot the confirmed Spy, and Lights Out attacked them both.)

Lights out: do you have any self heals left?
Jman: you’ll find out
...
Jman: will this answer your question
Jman has left the game.

Grice’s Maxim of Quantity specifies that communicators give the right amount of information - not too much or too little - but this maxim is continually challenged in ToS. Purposefully saying too much can be useful when pretending to be a role one is not, but even in the complete absence of words, communication can still be crystal clear.

Maxim of Quality: Deception and bluffing

Deception - the art of violating Grice’s Maxim of Quality - is a major mechanic in ToS. It might be outright lying about one’s role to avoid dying, or it might be something a bit more complex, like bluffing. In ToS, bluffers rarely achieve their goal of keeping themselves alive to win the game. Their strategy usually only causes temporary confusion, but it is still a fascinating thing to analyze.

These players are Mafiosi on trial because they have been discovered, yet try to pass themselves off as something else:

everytime Im TP people lynch me

TP refers to Town Protective; this player claimed to be a Bodyguard, a Town Protective role that can save a target from death by killing their would-be murderer, at the cost of the Bodyguard’s life.
“what kind of BG keeps a will”

This one also claimed to be a Bodyguard but could not provide his last will, containing the names of his protectees, as evidence. He is trying to roll with it and attempt to dupe the other players into thinking he does not need one.

You can claim a role, or just straight-up bluff, implying that you are a member of the town:

“town will lose [if you lynch me]”

This one is a bit different - and brazen. “Bones” is evil, in this case an Arsonist, but is not on trial. Knowing that sometimes people do not notice who is saying what, he throws this out:

“what was bones?”

How do we interpret this bluffing linguistically? Those who lie tend to do so in one phrase that references the fake role in an indirect way. They throw the lie out there timidly, perhaps thinking they're being clever, so as not to call attention to themselves. By contrast, true role declarations are much more direct and sometimes given without anyone asking:

Saladass: By the way I am sheriff
McCree: poppy, role?
Poppy: jailor

The role claim below was given by a prisoner. People who have been jailed usually claim their roles to the Jailor, so the Jailor knows whether to execute them or not.

“yo im vigi” [Vigilante]

The prevalence of bluffing and straight-up lying makes the Maxim of Quality among the most violated of all of the maxims in respect to ToS, as one might expect. As Ensslin stated, it’s certainly the same way with poker, online or not, or any other bluffing-dependent game whose very essence revolves around deliberately breaking the Maxim of Quality.

Maxim of Relation: Going way, way off topic

The Maxim of Relation recommends that one’s reply to another’s remark be relevant to the original conversation. As with the other maxims, it is not always obeyed. However, its flagrant violation results in some rather interesting exchanges.
BirSat: surv claim
BirSat: evils dont kill me
waifu: kk
animeisgay: anime is trash and if you watch it then you are trash aswell
Cthulu: im just gunna put this out there... i'm a hufflepuff

BirSat claimed to be the Survivor at the very beginning of the game, which is a typical thing to do because the Survivor’s goal is to survive the whole game; the conversation somehow went from there to Cthulu saying that they’re a Hufflepuff (as in Harry Potter) through a series of non-sequiturs.

Robin: random lynching tiiime
Willow: wow
McCree: rip Tls [Town Investigative - the Sheriff and Investigator just died]
Robin: i love you all

Robin’s last comment seems like it might lack context. It could be her way of saying goodbye to the townspeople because as McCree notes, they are dropping like flies. She could also be taking her role as the Escort - the role that literally loves everyone - quite seriously.

Maxim of Manner: ALL CAPS
Robin: I CAUGHT UMBERTO HAVING A WEAPON
...
UMBERTO HAS A WEAPON GUYS
...
AND HE HAS BLOOD
...
HOW DO YOU EXPLAIN THE BLOOD
HOW DO YOU EXPLAIN THE BLOOD
HOW DO YOU EXPLAIN THE BLOOD
...
IT SHOWED ME THAT YOU HAVE A GUN AND A BLOOD WTF THIS IS THE GAME,MAYBE YOU GOT FRAMED I DUNNo

Robin, as the Escort, used her night ability to “spend the night” (yes, this is what you probably think it is) with Umberto, distracting him from doing anything (in his case, killing anyone). However, the Mafia did kill someone that night, because if the Mafioso is hindered from killing their target, the Godfather will finish the job. As a result, it’s not quite clear how Robin knows what she knows. She could be completely making it up as she goes, but she
happens to be right! Regardless, since typing in all caps equates to yelling, this is definitely an overreaction which is inappropriate for the situation. By contrast, this is what non-hotheaded townies’ findings look like:

“John Proctor the sheriff - n1. Fayette - member of mafia”

“yo im TP [Town Protective] so protector is mafia from invest[igator]'s will”

Whereas Robin goes on a practical diatribe, repeatedly voicing her suspicions about Umberto in all caps, most everyone else just simply states their findings: “oh yeah, this person is mafia, I’m the Sheriff.” The yelling isn’t unique - other players do type their findings in all caps for emphasis - but most people clearly aren’t screaming “HOW DO YOU EXPLAIN THE BLOOD” multiple times. Others who yell at the town to get their attention say things like this:

Poopoo Head: VOTE 6 HE’S MAFIOSO

It’s the sheer quantity and the manner that make Robin’s example stand out.

**Breaking Grice’s Maxims 101**

I will leave you, the reader, with an example of a player who broke every rule in the book in an attempt to stay alive.

K: role? [K is a member of the town trying to root out shady role claims.]
moisterraiserxd: im just a moisterraiser
moisterraiserxd: xd
moisterraiserxd: witch
moisterraiserxd: i mean invest [Investigator]
moisterraiserxd: i mean gf [Godfather]
moisterraiserxd: xd
Lights out: jest [he implies that Moister is a Jester]
...
moisterraiserxd: guys im not a jester
moisterraiserxd: im a jäster

“Moisterraiserxd”, who is in fact a Witch, is pretending to be a Jester. The Jester wins the game along with any other winners if he/she is lynched, so many Jesters try to accomplish their goals by not definitively claiming a role, being obnoxious, and generally acting shady. "moisterraiserxd" is no exception. The issue with letting a Jester win is that if he/she does, he/she gets to kill one of the people who just voted to lynch them. And obviously, nobody wants to die in ToS, so nobody lynches the Jester.
Nearly all the examples above violate at least one of Grice’s Maxims of Conversation. The usual liars only violate the Maxim of Quality, but “moisterraiserxd” violates all of them in this exchange and throughout the entire game. He (can we assume it’s a he?) is not being informative, truthful, relevant, or concise, in an effort to reverse-engineer his way into surviving. This is by no means unique - players pretend to be Jesters all the time, sometimes convincingly, sometimes not. However, it is relatively rare to see someone breaking every maxim for any reason, let alone successfully.

**Conclusion**

Town of Salem is obviously a game where all of Grice’s Maxims are repeatedly broken, yet the vital communication between players remains intact. The Cooperative Principle assumes that the speaker and hearer “share some procedures of interpretation, of drawing the appropriate inferences from what is actually encoded” (McGregor, 2015). McGregor continues to explain that “in applying the Gricean maxims to [an example conversation] we had to appeal to background knowledge shared by the conversational participants, in this case information known generally to members of the speech community. In some cases, the shared information is specific to the conversation” (McGregor, 2015). So essentially, for Grice’s Maxims to be a valid framework for interpretation, the parties communicating need to be members of a speech community who share information providing a frame of reference.

Are ToS players a speech community? Are they a coherent group of people who share the same language(s) and similar norms of language use? Yes, loosely; even though they come from all over the world and not everyone speaks fluent English, let alone the same dialect, they use roughly the same linguistic norms and ToS-specific lexicon because the game essentially demands it. There is little room for regional slang, and there is no need for it. Furthermore, the players share game-specific rules, strategies, vocabulary and methods, which provide a coherent framework under which their often incomplete and sometimes erratic utterances are understood by the other players. Therefore, players of Town of Salem are indeed a weird and wonderful speech community, who revel in violating Grice’s Maxims for the sake of the game.
Sources


"Town of Salem Trial System." *Town Of Salem.* BlankMediaGames, 2016. Web. 02 Nov. 2016. <http://blankmediagames.com/Trial/>. This is where I obtained chat logs. The link leads to the login screen; you must have an account and have played a number of games before you can use the Trial System.