Truth and Falsity in Communication: Assertion, Denial, and Interpretation

Kensuke Ito[1]

Abstract

Our linguistic communication is, in part, the exchange of truths. It is an empirical fact that in daily conversation we aim at truths, not falsehoods. This fact may lead us to assume that ordinary, assertion-based communication is the only possible communicative system for truth-apt information exchange, or at least has priority over any alternatives. This assumption is underwritten in three traditional doctrines: that assertion is a basic notion, in terms of which we define denial; that to predicate truth of a sentence is to assert the content it expresses; and that one should, in the context of radical interpretation, try to maximize the truth of what foreigners believe or utter. However, I challenge this assumption via a thought experiment: imagine a language game in which everyone aims to exchange only falsehoods. I argue that information exchange is possible in this game, and so truth-guided communication and falsity-guided communication are conceptually on a par. As a consequence, we should reject the three doctrines, based as they are on the conceptual priority of assertion-based communication.

1 Introduction

In our linguistic communication, we aim to convey information. Our typical vehicles are declarative sentences, uttered with assertoric force. And we operate with the background assumption that, in general, we aim at truth. But can't we imagine that we English speakers can successfully convey truth-apt information by uttering only false English sentences, just as we do with true ones? Our vehicles for conveying information are also declarative sentences, but the presumption is that we all aim at falsehood, not truth.

The question is *not* to imagine a language that is semantically different from English. Wittgenstein considers a language such that "by '*p*' we mean $\sim p$" (1922: 4.062). Lewis also considers "a language like English but with all the truth values reversed"

✉ Kensuke Ito
  kensuke.ito@uconn.edu

1   University of Connecticut, 13 Old Schoolhouse Rd, Storrs, CT 06268, USA

([1980]: 80). Both suggest essentially the same language different from English, so let us call it 'Reverse English'.[1] In our English, by "Snow is white" we simply mean that snow is white. But a speaker of Reverse English means by "Snow is white" that snow is *not* white. Thus, the same string of symbols becomes true (or false) when tokened in English and false (or true) when tokened in Reverse English. But communication is still possible with Reverse English; instead of saying "Snow is white" in English, one can say "Snow is *not* white" in Reverse English. She thereby means that snow is white; and snow is in fact white. So, Wittgenstein says that the Reverse English sentence "Snow is not white" is *true* "construed in the new way".[2] Lewis says that the speaker is *truthful* in Reverse English.[3] Thus, if the present question is tackled along the lines of Wittgenstein and Lewis, the answer is negative; Reverse English speakers exchange truths in Reverse English, not falsehoods.

However, the present question is to ask whether, for a given language (e.g. English) whose truth-values are fixed, it is possible for the speakers to communicate with false sentences of that same language. In this paper, I argue that this is a conceptual possibility.

What consequences, if any, does this conceptual possibility have for our central notions of truth, assertion and denial, and interpretation? There are, I will argue, three significant consequences. Given the conceptual possibility in question,

1. Assertion and denial are conceptually on a par—neither is more basic than the other. So, we should reject the so-called Frege-Geach account of denial, according to which the notion of denial is defined in terms of assertion and negation.
2. We should reject one popular form of deflationism, which I shall call *illocutionary deflationism*. According to this form of deflationism, the role of the truth predicate is characterized in terms of the notion of assertion or endorsement. We cannot conflate its logical role with the alleged illocutionary role.
3. We should reject Davidson's Principle of Charity, understood, as Davidson understands it, as a necessary constraint on radical interpretation. It is not a matter of conceptual necessity that speakers communicate with true sentences.

In Sect. 2, I argue for the conceptual possibility of what I call "the F-game", played by speakers who successfully convey information by uttering only falsehoods. In Sect. 3, I argue that the possibility of the F-game shows that, contra Frege and Geach, assertion is not conceptually prior to denial. In Sect. 4, I argue that the possibility of the F-game refutes illocutionary deflationism. In Sect. 5, I argue that the

---

[1] Lewis calls it "Liarese" and characterizes it as spoken by a tribe of "Liars", who, unlike ordinary liars, never intend to deceive (Lewis [1980]: 80). I thank John Mackay for drawing my attention to Lewis's Liarese example.

[2] Wittgenstein says, "[A] proposition is true if we use it to say that things stand in a certain way, and they do; and if by "p" we mean $\sim p$ and things stand as we mean that they do, then, construed in the new way, "p" is true and not false" ([1922]: 4.062).

[3] "To be truthful in L is […] to try never to utter any sentences of L that are not true in L", where "L" can be any natural language or Liarese (Lewis [1975]: 167).

possibility of the F-game shows problematic consequences from the Principe of Charity taken as the a priori constraint for radical interpretation.

2 The F-game

Suppose that I utter:

It's false that zero is odd.

If you trust my sincerity and reliability, you can know that zero is even. Likewise, if I utter:

That's wrong!

as a response to someone's utterance "Zero is odd", you can draw the same information. Now, consider this case: you know that by uttering the next sentence, I am going to tell you a falsehood, and that I know that you know this.

(Z)   Zero is odd

Again, you can attain the same content from this utterance. Here I am not asserting the content that is expressed by (Z); rather, I am *denying* that content—presenting the content as *false*, just as I do when I said "It's false that zero is odd", and "That's wrong!". However, in the case of (Z), my negative attitude towards that content is not expressed by (Z) itself. There is nothing negative about (Z) itself: the content of (Z) is just that zero is odd.[4] So, if you are unaware that I was intending to tell a falsehood, you would mistake my denial for an assertion. Nevertheless, communication of truth-apt contents is possible even in the case of (Z), provided that you understand that I aim to produce a falsehood.

2.1 Grice's Maxim of Quality, and a Variant

Let me start by drawing on four maxims of Grice's as an articulation of conversational norms governing our communication. I shall call those maxims "the Truth rules", or just "the T-rules". Also, I shall call the communicative practice governed by the T-rules, "the T-game". Now, let us focus on the maxim of quality. It has one super maxim and two sub-maxims (Grice 1989: 45–6).

Maxim of Quality

Supermaxim	Try to make your contribution one that is true
Submaxim 1	Do not say what you believe to be false

---

[4] We should be careful to distinguish the case of (Z) from the case where I assert: "It is false that zero is odd". In the latter I aim to produce a truth. In the former I aim to produce a falsehood. These two cases are very different, even though the same information is conveyed.

Submaxim 2	Do not say that for the truth of which you lack adequate evidence

Maxim of Quantity	Don't say too much or too little
Maxim of Relation	Be relevant
Maxim of Manner	Be perspicuous

So what should the rules for the opposite game be? We replace every occurrence of the truth/falsity predicate by the falsity/truth predicate in the quality maxim. Then, we obtain what I call *the variant maxim of quality*.

Variant Maxim of Quality

Variant Supermaxim	Try to make your contribution one that is *false*
Variant Submaxim 1	Do not say what you believe to be *true*
Variant Submaxim 2	Do not say that for the *falsity* of which you lack adequate evidence

Call this variant maxim, together with the other three maxims, which remain unchanged, "the F-rules". Following the F-rules, one should try to put forward a sentence that one believes to be false. For instance, one might utter, "Snow is black", "Snow is yellow", "Snow is gray", and so on, while knowing that snow is in fact white. One can also utter, "Snow is *not* white"; and, in general, one can utter the negation of a sentence *s*, where one believes, with evidence, *s* to be true. I shall argue that the F-game is *playable* in the following sense: if every player of the F-game sincerely follows the F-rules, knowing that all other players do so, they can exchange truth-apt information.[5]

2.2 Playability

In the usual situation in which people are following the T-rules, everyone knows that everyone follows this set of rules. Suppose that Spencer utters a sentence *s*, and thereby expresses the propositional content *p*. Suppose also that the utterance is proper in terms of the T-rules. Then we may assume that *s* and *p* are in fact true, setting aside the question as to which is the primary truth-bearer. Now Audie hears *s* and understands that it expresses *p*. If she trusts Spencer's sincerity and reliability, she may infer:

(a) *p* is true.

---

[5] In a nutshell, I'm using the Maxim of Quality for capturing the normative aspects in truth-apt information. But I'm not arguing that the Gricean approach is the best account for capturing the normative aspects of assertion. In particular, my argument is not committed to *the truth account*—"assert *p* only if *p* is true"—espoused in Weiner (2005) and Whiting (2013). They employ the Gricean model to challenge *the knowledge account*—"assert *p* only if one knows that (it is true that) *p*"—championed by Williamson (2000). If one prefers the latter approach, one can use, instead of the Variant Maxim of Quality, this formulation: "utter *p* only if one knows that *it is false that p*".

Audie assumes that Spencer tries to put forward true utterances (by Supermaxim), and avoids conveying what he believes to be false (by Submaxim 1), based on certain evidence (by Submaxim 2). From those assumptions, Audie infers:

(b) The speaker believes *p* to be true.

What about the case of the F-game? We start with the alternate assumptions: Spencer utters *s*, which expresses *p*; and the utterance is proper *in terms of the F-rules*. Then, we may assume that *s* and *p* are *false*. Now Audie hears *s* and understands that it expresses *p*. If she trusts Spencer's sincerity and reliability, then she is likely to infer:

(c) *p* is false;
(d) The speaker believes *p* to be *false*.

Again Audie knows that Spencer is intending to put forward a falsehood (by Variant Supermaxim), and is avoiding conveying what he believes to be true (by Variant Submaxim 1), based on certain evidence (by Variant Submaxim 2). So Audie can infer that *p* is false, and not-*p* is Spencer's belief content.

It might be objected that an utterance can be false in too many different ways. Suppose that in some context of the T-game the appropriate conversational contribution is to utter "Snow is white". In this context, the following utterances are inappropriate: "Snow is black", "Snow has a color", "A swan is white", "Snow is white and snow is white and …", and so on. They are all inappropriate in one way or another: either false, or conveying too little information, or irrelevant, or unnecessarily prolix. In the T-game, the number of appropriate utterances is nicely constrained compared with the number of inappropriate utterances—which seems to be infinitely many. But the F-game, the objection goes, allows the players to utter what are inappropriate to utter in the T-game—so the ratio is reversed. This may seem to compromise information exchange via the F-game.

But this line of objection is based on confusing inappropriateness with falsity. In the F-game, speakers are also constrained by the maxims of quantity, relevance, and manner. For example, if in a situation the most appropriate content to convey is that snow is white, then the most appropriate utterance in the F-game is also fixed, that is, "snow is not white". These three maxims are still operating, so the number of permissible utterances does not proliferate any more than in the T-game. The three maxims are as effective in the F-game as they are in the T-game. Given the three maxims, it would, in the given situation, be inappropriate for the T-game player to say 'Snow is not black', and for the F-game player to say 'Snow is black'. Players of both games are equally constrained by the three maxims.

One might object further. In the above context, an F-game player should utter "Snow is *not* white". Still, it is true that the last utterance is more roundabout, compared with uttering "Snow is white" in the T-game. The F-game player has to use a negative linguistic term. Here, the F-game is more complicated, compared with the T-game. As to this objection, I agree that F-game communication often requires additional labor to convey the same information. But, first, notice

that sometimes the F-game is more efficient, as when in the T-game one wants to produce a negation: "Joe is not on time". In the F-game, one can say more simply "Joe is on time". To get more to the heart of the matter: efficiency as a communicative tool is not my concern here. My point is that the F-game players can exchange the same truth-apt information as the T-game players. It is precisely in this sense that the F-game is *playable*. The substantial impediments to conveying truth-apt information are eradicated in virtue of the maxims of Quantity, Relation, and Manner. So the concern about inefficiency has no effect on the claim that the F-game is an alternative practice for exchanging truth-apt information.[6]

Accordingly, it is as viable to exchange truth-apt information in the F-game as it is in our ordinary situation. The result forces us to re-examine our concept of *sincerity*, or 'truthful'-ness. For sincerity need not require that a speaker presents only truths or says only what she believes; rather it requires that speakers follow a shared set of coherent rules.[7] The F-game player is not insincere. Consider again the case in which I tell you whether zero is odd or even. If we are playing the F-game and I utter "Zero is even", I am deceiving you since, trusting me, you would infer that zero is odd, which is false. Intentionally uttering a truth in the F-game counts as a form of deception. Rather, sincerity is maintained in the F-game by everyone's aiming at falsity.

3 The Conceptual Parity of Assertion and Denial

There are several approaches to a characterization of the illocutionary act of assertion. Some may focus on the mental states that an assertion expresses, or norms governing those states; and others may emphasize the delivered information that expands a 'common ground', or the commitments one takes on by making an assertion. However we characterize assertion, what an F-game player performs is the

---

[6] There is a claim that denial is evidentially less specific than assertion, which has a long history in philosophy (see Horn 2001: ch1). When one asserts that a is F, it must be based on the particular evidence that the individual constant a satisfies the predicate F. In contrast, when one denies that a is F, there are several possibilities for justifying this utterance: a does not satisfy F, a is empty, to predicate F of a is a category mistake, etc. (see Dickie 2010; Incurvati and Schlöder 2017). However, such "messiness" of denial is orthogonal to the possibility of the present thought experiment. For the F-game brings about the change in the way of speaking, not the way of how we deal with evidence. To utter "a is not F" in the F-game requires the same specificity of evidence that a is F as to utter "a is F" in the T-game does. Instead, uttering "a is F" becomes messy in the F-game.

[7] The norm that one should assert what one believes is considered by some to be a norm of sincerity (e.g. Lowlor and Perry 2008; Green and Williams 2010; see also Goldberg 2015: 1.3). In a discussion of deflationism, Price in (2011) takes it to be the norm of sincerity that one should manifest what one has in mind. According to Price, this norm is not a genuine norm of assertion because this claim also applies to the case of other mental states such as desire. One should follow the norm of sincerity just described not only in the case of expressing beliefs or propositional contents one has in mind, but also in the case of expressing desires or other non-propositional contents. I agree with Price that manifesting what one actually believes is not the norm of assertion for this reason. But I think, as against Price, that trying to manifest what one *disbelieves* may count as a sincere act in certain cases, as, for example, in the F-game.

opposite illocutionary act, that is, a denial.[8] If people play the F-game, then they give expression to a negative attitude towards certain content.[9] A proper F-game utterance expresses the speaker's *dis*belief in the expressed content. The speaker will be responsible for the *falsity* of the uttered sentence; otherwise the audience will be misinformed. Thus, in the F-game, the primary illocutionary act is not assertion, but denial, the act of presenting a propositional content as false. And denials in the F-game do not necessarily involve negation.[10] This leads us to question an account of denial derived from Frege and Geach.

3.1 The Frege-Geach Account of Denial

According to Frege's account of denial in (1919), elaborated by Geach in (1965), denial does not form a *sui generis* category, since it is definable in terms of assertion and negation:

denying *p* = df. asserting not-*p*.

---

[8] According to MacFarlane's classification, there are largely four types of theories of assertion – *the attitudinal account*, *the common ground account, the constitutive rule account*, and *the commitment account* (MacFarlane 2011; the labels are from (Goldberg 2015: 9)). My claim that an utterance in the F-game counts as a denial is compatible with any of these accounts; for, whatever characterization one gives to assertion, I may ask for the corresponding account of denial, and then argue that the denial with that characterization becomes the normal move in the F-game.

[9] In his "Truth" (1959), Dummett famously compares our linguistic practice with playing a game of chess, claiming that this normative aspect is part of our concept of truth, but is not captured by Frege's theory:

 [I]t is part of the concept of truth that we aim at making true statements; and Frege's theory of truth and falsity as the references of sentences leaves this feature of the concept of truth quite out of account. Frege indeed tried to bring it in, afterwards in his theory of assertion — but too late; for the sense of the sentence is not given in advance of our going in for the activity of asserting, *since otherwise there could be people who expressed the same thoughts but went in instead for denying them.*
(Dummett 1959: 2; emphasis added).

 To reconstruct this passage as a simple *reductio* form*,* suppose that Frege's theory is correct. It follows that the proposition or thought expressed by a sentence would be identifiable independently of our activity of assertion. Then we should be able to express that content while performing the opposite speech act, i.e., denial. But this is absurd, as contended in the italicized clause. So it is concluded that the starting supposition is wrong.

 The playability of the F-game shows that Dummett's reductio argument doesn't work; the F-game shows that it is possible that everyone expresses thoughts in order to go in for denying them. Dummett's argument does not reach an absurdity.

[10] My treatment of denial is different from *bilateralism*: "the view that meanings in general are to be given via conditions on assertion and denial" (Ripley 2020), the view also held in (Price 1990; Smiley 1996; Rumfitt 2000; Restall 2005). For Bilateralists, assertion and denial are both primitive notions and so conceptually on a par. In contrast, my treatment of denial does not take assertion and denial to both be primitive notions. I establish parity in a different way. In the context of conversation and truth-apt information exchange, the F-game shows that we can exchange information via denials just as well as we can via assertions. Neither assertion nor denial has priority over the other. It is in this way that I argue for the conceptual parity of assertion and denial.

I will call this explanation the Frege-Geach account of denial.[11] Given this account of denial, denial is a derivative notion, to be characterized by assertion and negation.[12] So all we need for constructing a logic system is assertion and negation.[13] Geach writes:

> [I]f we regarded [denying] a proposition as different from asserting the negation of a proposition, we should have here two quite different logical forms; we might write these as follows, using Lukasiewicz' sign ⊣ for a [denial] [as] opposed to Frege's assertion ⊢:
>
> ⊢If not q, then r; ⊣q; ergo ⊢r.
> ⊢If p, then r; ⊢p; ergo ⊢r.
>
> Plainly this is a futile complication. All we need in logic for assertion and negation is two signs—the assertion sign, and a negation which does not convey [denial] (as in "if not q…"); whatever is more than these, as Frege says, cometh of evil.
> (Geach 1965: 455).[14]

With the special assertion sign together with the logical operation of negation, we do not have to consider denial as a primitive in its own right. This way, assertion is conceptually prioritized over denial.

3.2 A Problem with the Frege-Geach Account of Denial.

Recall that an F-game player denies that zero is odd by uttering "zero is odd". In this act of denial, there is no use of negation. Moreover, in the F-game, in order to inform her audience that zero is even, the player can utter "Zero is *not* even". Thus, in the F-game, the primary illocutionary act is denial, rather than assertion. So, assertion in turn becomes the definiendum:

asserting p = df. denying not-p.

---

[11] Rumfitt notes that early Frege held a theory of "content of possible judgment" (Rumfitt 2000). According to the theory, when we judge "p" we grasp not only "p" but also "~p", and so "[t]he rejection of the one and the acceptance of the other are one and the same" (Frege 1879–81: 8). Here Frege suggests an account that treats acceptance and rejection conceptually on a par with respect to negation. Later, Frege dropped the terminology "content of possible judgment" when he discovered the sense/reference distinction, but he suggested no substantial change to the underlying idea in the previous account (e.g. Frege 1892: 186). Thus, there is room for asking, as some bilateralists do, how strongly Frege himself is committed to what I call the Frege-Geach account of denial (see, for example, Ripley 2020, fn. 2). I won't here decide the textual question of what exactly Frege's view was. The Frege-Geach account of denial that is my focus is certainly a standard interpretation of Frege; in fact, Geach calls it "the Frege point" (Geach 1965: 449).

[12] For the bilateralist position concerning this definitional account, see (Ripley 2011, 2020).

[13] Smiley (1996) characterizes this approach in terms of Ockham's razor.

[14] A terminological point: where I use "denial", Geach would employ the term "rejection".

where assertion is characterized in terms of denial and negation. Thus, given that assertion and denial can each be characterized by the other *plus* negation, we should not draw the conclusion that assertion is more primitive in the definitional or explanatory order.

Furthermore, the possibility of the F-game exposes a tacit assumption in Geach's remarks, that introducing a denial sign in a standard logic system is "a futile complication". His point can only be true when we have already introduced the assertion sign. What we do by uttering "Snow is white" in the T-game is achieved by uttering "Snow is *not* white" in the F-game. If we regard the former as an assertion, then the same conversational effect is achieved by using negation in the F-game. So, one can describe the F-game practice employing *only* Lukasiewicz's reversed turnstile sign and negation. Certainly, it is useless to introduce two turnstiles that are interdefinable by negation. However, we should notice that, just as we can dispense with ⊢ in the T-game, the F-game players can successfully dispense with ⊣.[15] Therefore, theoretical parsimony does not support the Frege-Geach account of denial.

4 The Truth Predicate in the F-Game

The T-game and the F-game can be played in the same language, say, English. In both games, "snow" remains the name of the same referent, "being white" outputs truth when it takes a name of a white substance as an input, and "not" switches the truth-values. However, these names, predicates, and functions may be used for

---

[15] Here's how such an 'F-logic' would look in comparison with the standard one, 'the T-logic'. First, let's stipulate that if a sentence follows the reverse turnstile $\$\$\dashv\$\$$, it means that the sentence is produced while following the F-rules. So,

⊣Snow is black,

represents my *denying* the content that snow is black.

The F-game allows the players to generate only falsehoods. So the rules of the F-logic must be *falsity preserving* in the sense that they allow the F-game players to draw from known falsehoods some new falsehood. Let's start with conjunction. Consider the sentence "Snow is white and snow is not white". It is false, and so we can state it in the F-game. However, we may not eliminate the conjunction to present each conjunct separately, since the first conjunct is true in isolation. So the standard Conjunction Elimination rule does not obtain. Instead, we know that the whole conjunction is false whenever one conjunct is false. So, given that p is false, we can infer any conjunction $p \& q$ with an arbitrary q. Accordingly, the F-logic has the following introduction rules for conjunction: $\frac{\dashv p}{\dashv p \& q}$[& Intro1] $\frac{\dashv q}{\dashv p \& q}$[& Intro2]

These are the 'upside-down' versions of the standard Conjunction Introduction. And similar reasoning shows that the F-logic has the Disjunction Elimination as follows:

$\frac{\dashv p \lor q}{\dashv p}$[∨Elim1] $\frac{\dashv p \lor q}{\dashv q}$[∨Elim2]

Thus, the behavior of conjunction governed by its introduction rules. in the F-logic is just like that of disjunction governed by its introduction rules in the T-logic; and the behavior of disjunction governed by its elimination rules in the F-logic is just like that of conjunction governed by its elimination rules in the T-logic.

Now go back to a contradiction:

$p \& \sim p$

This is a logical falsehood, and so we can put it forward with the reverse turnstile. But, with the variant introduction and elimination rules, the sentence above does not generate explosion. This suggests that the F-logic is no more inconsistent than the T-logic.

different purposes in the two games. In the T-game, one combines them into a sentence and states it in order to manifest a positive attitude towards the expressed content. In the F-game, what's manifested is a negative attitude. In this section, I consider how the truth predicate operates in these two games.

4.1 Illocutionary Deflationism

According to one version of deflationism, the truth predicate is used for making assertions. Ayer, for example, writes:

> Thus, to say that a proposition is true is just to assert it, and to say that it is false is just to assert its contradictory. And this indicates that the terms "true" and "false" connote nothing, but function in the sentence simply as marks of assertion and denial.
> (Ayer 1946: 88–9).

Ayer reduces the use of the truth predicate to the making of an assertion. He goes on to say, because of the equivalence of asserting a sentence with predicating truth of its nominalization, "the phrase 'is true' is logically superfluous" (ibid. 88). Let us formulate this claim by the following equivalence:

Assertion & Truth   To predicate "true" of *s* is to assert that *p*

where "*p*" is the proposition expressed by the sentence *s*. This is the view I have called *illocutionary deflationism*, since it appeals to the notion of a speech act in order to develop the doctrine of deflationism about truth.[16]

An initial objection is raised by Frege and Geach. For example, in the sentence "If it is true that *p*, then *q*", the content that *p* is not asserted. So, a qualification is required: hereafter we only consider the predication applied in a non-embedded context.[17] Then, illocutionary deflationism is a position that the role of the truth predicate is explained in terms of the notion of assertion.

Assertion & Truth*   To predicate "true" of *s* in a non-embedded context is to assert *p*

The characterization of illocutionary deflationism is not yet complete, since, as it stands, it does not capture the very utility of having the truth predicate in a language. For instance, consider the case in which Anna says "What Bob said yesterday was true", even though she cannot precisely recollect the sentence he actually uttered. In general, one can predicate "true" of a sentential nominalization, even if one cannot repeat the corresponding utterance. In that case, the sentence of which one wants to predicate "true" is not explicitly given. But, given the ordinary conception of

---

[16] The terminology is from (Bar-On and Simmons 2007: 73). Illocutionary deflationism tries to explain the role of the truth predicate in terms of assertion. Unlike Bar-On and Simmons, here I am not arguing for (or against) the claim that the concept of truth is needed to explain the concept of assertion.

[17] There is a non-Fregean approach to truth predication in an embedded context (e.g. Hanks 2015: Ch 4), but, for reasons of space, I set this aside here.

assertion, it seems incoherent to say, "I assert some propositional content. But I cannot specify what it is—I can neither spell it out nor speak it aloud". So, *Assertion & Truth** does not fully explain the utility of using the truth predicate. We need to find another type of speech act, preserving the motivation of *Assertion & Truth**.

The special role of the truth predicate has been characterized as the logical role of *generalization*. Let us first look at a passage from Quine.

> We may affirm the single sentence by just uttering it, unaided by quotation or by the truth predicate; but if we want to affirm some infinite lot of sentences that we can demarcate only by talking about the sentences, then the truth predicate has its use. We need it to restore the effect of objective reference when for the sake of some generalization we have resorted to semantic ascent.
> (Quine 1970: 146).

For example, one can predicate "true" of the sentential nominalization "every sentence of the form '*p* or not *p*'", and thereby affirm infinitely many sentences of this form. But then, since we never complete the enumeration, we cannot assert them all. So, this is another type of case in which the truth predicate serves a purpose in the absence of assertion. Truth predication allows us to disquote such sentential nominalizations, and so we can affirm sentences that are not explicitly given but just "demarcated" by the nominalization.

This utility of disquotation has been widely recognized in the literature, whether one is for or against deflationism. Crispin Wright captures this role of the truth predicate in terms of *endorsement*.

> Since the defining thesis of deflationism is that "true" is merely a device of disquotation—a device for endorsing assertions, which we need only for the purposes of indirect ("Goldbach's Conjecture is true") or compendious ("Everything he says is true") such endorsements—since that is the very essence of the view, a deflationist must of course insist that the only substantial norms operating in assertoric practice are norms of warranted assertibility and that the truth predicate can indeed mark no independent norm.
> (Wright 1992: 18).

Consider also:

> As a device for semantic ascent, the truth predicate allows us to endorse or reject sentences (or propositions) that we cannot simply assert, adding significantly to the expressive resources of our language.
> (Williams 1999: 547).
> It is widely accepted that we use "true" to endorse propositions that we cannot assert directly.
> (Scharp 2013: 63).

The latter two passages clearly contrast assertion and endorsement.[18] In general, the utility of disquotation is explained by the notion of endorsement: disquotation enables one to endorse what one cannot explicitly assert. Of course, one can 'reject' what one cannot assert by predicating "*not true*" of the sentence nominalization, as Williams mentions.

Going back to illocutionary deflationism, it seems that the notion of endorsement is a good candidate for extending the idea that the role of the truth predicate can be explained in terms of its illocutionary role. The predicate enables us to endorse some content that cannot be explicitly articulated and so is un-assertible.

Endorsement & Truth   To predicate "true" of *s* in a non-embedded context is to endorse that *p*[19]

There are other candidates that seem close to endorsement, e.g., confirmation. These alternatives, however, might have implications with respect to temporality, whereas endorsement seems free from them. Consider Strawson, who is clearly committed to the idea that the role of the truth predicate is explained in terms of its illocutionary role:

> [T]he sentence "What the policeman said is true" has no use except to confirm the policeman's story […] the sentence does not say anything further about the policeman's story or the sentences he used in telling it. It is a device for confirming the story without telling it again. So, in general, in using such expressions, we are confirming, underwriting, admitting, agreeing with, what somebody has said; […] but we are not making any assertion additional to theirs; and are never using "is true" to talk about something which is what they said, or the sentences they used in saying it.
> (Strawson 1949: 93).

This passage suggests that the notion of confirmation, or any of Strawson's other candidates, intimates that somebody *(has) said* something. There are three concerns here. First, as long as there are infinitely many instances of "sentence of the form '*p* or not *p*'", we cannot incorporate such a past-tense implication. Second, there is no intrinsic problem with attaching the truth predicate to a phrase such as "What Bob will say". Third, the same problem may arise from empty descriptions, such as "the present King of France". That is, the sentential nominalization might be an improper description. Suppose that Anna says "What Bob said last night is true" but he made no utterance yesterday. This is a problem for the theory of reference, not for the theory of the truth predicate. I take it that the truth predicate may be applied to the sentential nominalizations that have not yet been produced or even lack denotations. For those reasons, we should avoid the view that, when one uses the truth predicate,

---

[18] Williams' characterization also states that the truth predicate allows us to reject some content. I take this to mean that we can reject what another person said by uttering, "What you said is *not* true", without repeating the actual sentence.

[19] The relation between assertion and endorsement is asymmetric: the former entails the latter, but not vice versa. This fits our ordinary conception of assertion and endorsement.

the object that she takes to be true is already given in the past. Endorsement, in the sense I have defined, is neutral with respect to such temporality or even the existence of the nominalized. Accordingly, I shall take *Endorsement & Truth* as the appropriate formulation of illocutionary deflationism.

4.2 A Problem for Illocutionary Deflationism

The objection is simple: *Endorsement & Truth* is false since F-game players do not endorse a proposition when they predicate "true" of the corresponding sentence. Suppose that one utters "Snow is black" in the F-game. She is not endorsing that snow is black, since she, as a proper player of the F-game, is trying to make her contribution one that is false, while avoiding saying what she believes to be true and what she lacks the evidence for the falsity thereof. Rather, she is denying that snow is black. But then she may also utter "'Snow is black' is true". In this utterance as well, she's denying that content. The point applies everywhere in the F-game: consider the sentences such as "What the suspect stated is true" or "Any sentence of the form '*p and* not *p*' is true". The generalizing role of "true" is as important to the F-game as it is to the T-game. But in the F-game, the predicate "true" does not provide a means of endorsement. An utterance of "'Snow is black' is true" is as much a denial of the content that snow is black as an utterance of "Snow is black". In these instances, "true" still retains its generalizing role. But, again, the F-game player does not endorse the contents expressed by those sentences.

So the idea that the role of the truth predicate is conceptually fixed in terms of the notion of endorsement is false. One might make the empirical claim that to predicate "is true" of *s* in a non-embedded context is usually, or more often than not, to endorse, to which I have no objection. However, illocutionary deflationism is a conceptual claim, and it is false.

Relatedly, the F-game shows that the claim that truth is a device for endorsement has to be relativized—the claim is correct as long as we're playing the T-game. Suppose a T-game player and an F-game player each say: "What Jane said yesterday is true". Both are in effect *repeating* what Jane said—thanks to truth's disquotational role. But the T-game player thereby endorses what Jane says, while the F-game player doesn't. If one insists on identifying some illocutionary aspect of the truth predicate, one should perhaps say that the predicate is a device to *repeat* truth-bearers. But the unrelativized claim that truth is a device for endorsement is false.[20]

---

[20] One might argue that there is a conceptual link between our *current* practices and the meaning of the truth predicate. To this claim, the F-game does not present any problem, because it only shows that, if we started the F-game, the meaning would change as the practice changed. I thank an anonymous referee for pointing out this claim.

5 Interpreting F-Game Players

Since the F-game is playable, let us imagine what if some other linguistic community that plays it. Suppose that German is a completely foreign language for an English interpreter Eddy, and also that, unknown to him, all Germany speakers started to play the F-game. Suppose further that the German speaker Greta utters, "Schnee ist nicht weiss". As we saw in Sect. 1.3, we meta-thinkers know that Greta, if she is sufficiently sincere and reliable, utters that sentence when and only when she believes that snow *is* white. But the question is how Eddy interprets Greta's utterance. For a Davidsonian truth-conditional theory of meaning, this is to ask what kind of T-sentence Eddy will produce for Greta's utterance. In this section, I shall describe how the case of Eddy and Greta poses a problem for the truth conditional theory of meaning applied to radical interpretation.

5.1 The Truth Conditional View in Radical Interpretation

Given a relevant T-sentence:

   s is true in L if and only if p,

   I take the truth conditional view to be composed of these three tenets:

   (I)   The T-sentence specifies the truth condition of *s*.
   (II)   The truth condition given via (I) in turn specifies the semantic meaning of *s*.
   (III) Based on (II), we can ascribe the content that *p* as the content of the speaker's belief.

(where "*s*" is replaced by the name of a foreign sentence, and "*p*" is replaced by an English translation of that sentence). Davidson would say that (I) is the basis for *both* (II) and (III), as he says "in interpreting utterances from scratch […] we must somehow deliver simultaneously a theory of belief and a theory of meaning" (1974: 144). In this section I argue against this truth conditional view.

Given the theory applied to radical interpretation, an interpreter, by producing a relevant T-sentence, specifies the truth condition of a given sentence. Then we are led to the second tenet: the relevant T-sentence gives the meaning of the sentence mentioned on the left hand side, which is specified by the sentence used on the right hand side. Finally, assuming that a foreign speaker sincerely expresses her belief, the interpreter ascribes to her a belief with the content expressed by the sentence used on the right. That is, based on the identification of the speaker's mental content with the meaning of the sentence she produced, we are led to the third tenet of the above theory.

Put in more concrete terms, suppose that Françoise, a French speaker who is sincere and reliable, utters the sentence "La neige est blanche". The *Principle of Charity* guides Eddy "to choose truth conditions that do as well as possible in

making speakers hold sentences true when […] those sentences are true" (Davidson 1974: 15). So, we expect that, given that Eddy recognizes that Françoise utters that sentence when and only when snow is white, he would produce the T-sentence:

"La neige est blanche" is true in French if and only if snow is white.

By (I), he is giving the truth condition of the French sentence. So, by (II), Eddy succeeds in uncovering its meaning. And, by (III), he can ascribe to Françoise a belief whose content is that snow is white.

5.2 A Problem for the Truth Conditional View

Going back to the case of Greta, what kind of T-sentence would Eddy produce in response to her utterance "Schnee ist nicht weiss"? As an F-game player, Greta utters this when and only when snow *is* white, the same conditions as those under which Françoise utters "La neige est blanche". Let us stipulate that Eddy can only avail himself of the resources that are observable to him, which presumably are reducible to the data concerning when and only when foreign speakers utter a sentence,[21] the ones of the same kind as he employs for interpreting French sentences. He presupposes that Greta is as sincere and reliable as Françoise, that is, that Greta by and large believes only truths, and expresses them by asserting them. Thus, as long as his interpretation is subject to the Principle of Charity as a necessary constraint, Eddy considers that the uttered sentences are largely true. Then, it is reasonable to expect Eddy to produce:

(T1)   "Schnee ist nicht weiss" is true in German if and only if snow is white

According to (I), Eddy specifies the truth condition of the sentence "Schnee ist nicht weiss" by this T-sentence. And so, by (II), Eddy is supposed to understand the *semantic* meaning of "Schnee ist nicht weiss".[22] But this is wrong; Eddy does not capture the semantic meaning of Greta's utterance. And the misinterpretation of the negation "nicht" may lead to the misinterpretation of compositionality in the target language.[23] Finally by (III), Eddy ascribes to Greta a belief with the content that

---

[21] Davidson himself later replied to his critics; that, although he sometimes used the term "radical interpretation" to refer to "the special enterprise of interpreting on the basis of a limited and specified data base", he "has never argued, specified, or assumed […] that the data on which the special enterprise is based exhaust the data available to actual interpreters" (1994: fn. 2).

[22] One might think that Eddy understands the *use* meaning of that sentence, inasmuch as Eddy is supposed to recognize when and only when he can also utter that sentence. But then, the distinction between the truth conditional theory and the use theory of meaning seems to collapse, which would be undesirable for proponents of the former.

[23] As outlined in footnote 15, the logic of the F-game is falsity preserving. Now assume that Eddy observes Greta's inference from (i) to (ii).

 (i) Schnee ist nicht weiss.

 (ii) Schnee ist nicht weiss *und p*.

 (where "*p*" is an arbitrary German sentence). The inference is correct with respect to the F-rules, in the sense that Greta's inference is falsity preserving. Also, as we've just seen, Eddy translates (i) into (iii).

snow is white. Despite his failure to identify the semantic meaning of the sentence Greta utters, Eddy's mental content ascription is appropriate.

Accordingly, the present version of the truth-conditional view is unsatisfactory given the Principle of Charity. If a radical interpreter does not know which set of rules is in operation, the semantic meaning of a sentence is not fixed by producing the T-sentence, which contradicts (II). Despite this failure of meaning specification, however, Eddy successfully attributes an appropriate belief content to Greta. So, the success of mental content ascription does not depend on the appropriateness of the meaning that the truth condition provides, as against (III).[24]

These consequences lead us to re-examine the modal status of the Principle of Charity. If the principle is taken to be an a priori constraint, it should not be characterized in terms of truth. The formulation of the principle usually maintains, directly or indirectly,[25] that the interpreter should trust that her subject is trying to generate truths. But, as we've just seen, Eddy fails to give the meaning of the sentence that Greta utters in the F-game, as long as he expects that Greta tries to utter only truths. So the Principle of Charity should not be characterized in terms of alethic notions. Other types of formulation employ the notions such as *sincerity* and *rationality*. Those notions seem much more appropriate, since we may assume that Greta and all other F-game players are sincere and rational. As to sincerity, recall Sect. 2.2; as to rationality, they are at least minimally rational in the sense that there is no

---

Footnote 23 (continued)

(iii) Snow is white.

Now Eddy observes that Greta and other German speakers, following the F-rules, always allow to draw a new sentence with "und". This performance coincides with our Disjunction Introduction. Thus, Eddy seems likely to conclude that "und" in German means "or" in English. Then, (ii) is translated into (iv).

(iv) Snow is white *or p*.

In a similar vein, he would mistake "oder" for "and". Thus, Eddy will mistake conjunction (disjunction) in English for disjunction (conjunction) in German.

[24] The discrepancy between the failure of specifying the sentence's meaning and the success of specifying speaker's mental content might suggest a weaker notion of 'interpretation'. In contrast to the standard notion, the weaker interpretation does not aim to provide the semantic meaning of an uttered sentence. It rests satisfied with identifying other's mental content, and it makes sense to think that, as long as Eddy knows what Greta *believes*, he successfully interprets this rational agent. It seems interesting to compare the distinction between two notions of interpretation with that between semantic and use meaning mentioned in footnote 22. But in this paper I only deal with the standard notion of interpretation.

[25] For example, *The Oxford Dictionary of Philosophy* (Blackburn 2008) characterizes the principle as the constraints on the interpreter 'to maximize the truth or rationality in the subject's sayings' (see the entry, "charity, principle of"). Notice that it mentions the truth at the level of "sayings", not at the level of doxastic attitudes. So, in general, if the formulation *forces* the interpreters to maximize the truths of utterances, then the F-game is a *direct* counterexample. Another type of formulation appeals to the truth in foreigners' beliefs. For example, Ludwig characterizes the principle as that "a speaker's *beliefs*, particularly those that are responses to his environment, are largely true" (Ludwig 2004: 353). Then, this formulation itself is not problematized by the F-game. There remains a problem, however, insofar as it is supposed to follow from the belief formulation that when a speaker holds true a sentence, by and large the sentence is true (ibid.). Hence, in general, if the formulation *leads* the interpreters to maximize the truths of utterances, then the F-game is an *indirect* counterexample.

inconsistency among the utterances they produce.[26] Given the playability of the F-game, we cannot take it as an a priori constraint that foreigners' utterances are by and large true.[27]

The F-game illustrates the possible discrepancy between giving the meaning to an uttered sentence and ascribing the belief content to the utterer in a context of radical interpretation. Holding the Principle of Charity as an a priori constraint for radical interpretation strategically amounts to excluding the possibility of the F-game; if the exclusion of the F-game is correct, Eddy will not experience the discrepancy between (II) and (III). However, only non-conceptual reasons can justify the claim that a given language community plays the T-game.

6 Conclusion

I have argued that the possibility of the F-game yields three significant consequences. First, since the T-game and the F-game are both playable, assertion and denial are conceptually on a par, and this repudiates the Frege-Geach account of denial. Second, the predicate "true" as it is used in the F-game is a device for making denials (though its generalizing role is unaffected). So illocutionary deflationism is false. Third, the notion of truth conditions fails to play the role of meaning-giver when one tries to interpret foreign linguistic practice governed by the F-rules, given the Principle of Charity as an a priori constraint. These three doctrines take for granted the problematic assumption that the T-game has conceptual priority. Certainly, it is a cross-linguistic fact that any natural language practice satisfies the T-rules, not the F-rules. But it is a *fact*, not a conceptual necessity. Given the conceptual possibility of the F-game, the T-rules—among them, the maxim of quality—actually obtain for non-conceptual reasons.

References

Ayer, A. J. (1946). *Language, truth, and logic* (2nd ed.). New York: Dover.

Baron, D., & Simmons, K. (2007). The use of force against deflationism: Truth and assertion. In D. Greimann & G. Siegwart (Eds.), *Truth and speech acts: Studies in the philosophy of language* (pp. 61–89). New York, NY: Routledge.

Blackburn, S. (2008). *The Oxford Dictionary of Philosophy. Second* (revised). Oxford: Oxford UP.

Davidson, D. (1974). Belief and the basis of meaning. *Inquiries into truth and interpretation* (2nd ed., pp. 141–154). Oxford: Oxford University Press.

Davidson, D. (1994). Radical interpretation interpreted. *Philosophical Perspectives, 8,* 121–128.

Dickie, I. (2010). Negation, anti-realism, and the denial defence. *Philosophical Studies, 150,* 161–185.

---

[26] Here I take consistency in utterances to be a hallmark of (minimal) rationality. The F-game players keep producing false sentences, but such a practice itself does not generate any inconsistency in what they say, as illustrated in footnote 15.

[27] As to more detailed discussions about the epistemological and metaphysical status of the principle, see (Glüer 2006; Pagin 2006), in which the authors argue, against Davidson, that the principle is not an a priori but an a posteriori necessity.

Dummett, M. (1959). Truth. In *Truth and other enigmas* (pp. 1–24). Cambridge, MA: Harvard University Press.

Frege, G. (1879–81). Logic. In H. Hermes, F. Kambartel, & F. Kaulbach (Eds.), *Posthumous writings* (pp. 1–8). Oxford: Basil Blackwell.

Frege, G. (1892). On Concept and Object. (In M. Beaney (Ed.), *The Frege Reader* (pp.181–93). Oxford: Blackwell Publishing, 1997)

Frege, G. (1919). Negation. In P. T. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege* (pp. 117–136). Oxford: Oxford University Press.

Geach, P. T. (1965). Assertion. *Philosophical Review, 74,* 449–465.

Glüer, K. (2006). The status of charity I: Conceptual truth or a posteriori necessity? *International Journal of Philosophical Studies, 14,* 337–359.

Green, M., & Williams, J. (2010). Moore's paradox, truth and accuracy. *Acta Analytica, 26,* 243–255.

Grice, P. (1989). *Studies in the ways of words*. Cambridge, MA: Harvard UP.

Goldberg, S. (2015). *Assertion: On the philosophical significance of assertoric speech*. Oxford: Oxford UP.

Hanks, P. (2015). *Propositional Content*. Oxford: Oxford UP.

Horn, L. (2001). *A natural history of negation*. Stanford, CA: CSLI Publications.

Incurvati, L., & Schlöder, J. J. (2017). Weak rejection. *Australasian Journal of Philosophy, 95*(4), 741–760.

Lewis, D. (1975). Languages and language. In *Philosophical papers* (Vol. 1, pp.163–88). Oxford: Oxford UP.

Lewis, D. (1980). Index, context, and content. In S. Kanger & S. Öhman (Eds.), *Philosophy and grammar* (pp. 79–100). Holland: D. Reidel.

Lowler, K., & Perry, J. (2013). Moore's paradox. *Australasian Journal of Philosophy, 86,* 421–427.

Ludwig, K. (2004). Rationality, language, and the principle of charity. In A. R. Mele & P. Rawling (Eds.), *The oxford handbook of rationality*. Oxford: Oxford UP.

MacFarlane, J. (2011). What is assertion? In J. Brown & H. Cappelen (Eds.), *Assertion: New philosophical essays* (pp. 79–96). Oxford: Oxford UP.

Pagin, P. (2006). The status of charity II: Charity, probability, and simplicity. *International Journal of Philosophical Studies, 14,* 361–383.

Price, H. (1990). Why 'not'? *Mind, 99*(394), 221–238.

Price, H. (2011). Three norms of assertibility, or how the moa became extinct. *Philosophical Perspectives, 12,* 241–254.

Quine, W. V. O. (1970). *Philosophy of logic*. Englewood, NJ: Prentice Hall.

Restall, G. (2005). Multiple conclusions. In P. Hájek, L. Valdés-Villanueva, & D. Westerståhl (Eds.), *Logic, methodology and philosophy of science: Proceedings of the twelfth international congress* (pp. 189–205). London: King's College Publications.

Ripley, D. (2011). Negation, denial, and rejection. *Philosophy Compass, 6,* 622–629.

Ripley, D. (2020). Denial. In V. Déprez & M. T. Espinal (Eds.), *Oxford handbook of negation* (pp. 45–57). Oxford: Oxford UP.

Rumfitt, I. (2000). Yes" and "No. *Mind, 109,* 781–823.

Scharp, K. (2013). *Replacing truth*. Oxford: Oxford UP.

Smiley, T. (1996). Rejection. *Analysis 56*(1), 1–9.

Strawson, P. F. (1949). Truth. *Analysis, 9,* 83–97.

Weiner, M. (2005). Must we know what we say? *The Philosophical Review, 114,* 227–251.

Whiting, D. (2013). Stick to the facts: On the norms of assertion. *Erkenntnis, 78,* 847–867.

Williams, M. (1999). Meaning and deflationary truth. *The Journal of Philosophy, 96,* 545–564.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford UP.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. (D. F. Pears and B. F. McGuinness (trans.), London: Routledge.

Wright, C. (1992). *Truth and objectivity*. Cambridge, MA: Harvard UP.