Philosophia (2018) 46:39–55
DOI 10.1007/s11406-017-9909-x

Indexicals in Remote Utterances

Adrian Briciu[1]

Received: 30 March 2017 / Revised: 12 July 2017 / Accepted: 12 September 2017 /
Published online: 10 October 2017

Abstract Recording devices are generally taken to present problems for the standard Kaplanian semantics for indexicals. In this paper, I argue that the remote utterance view offers the best way for the Kaplanian semantics to handle the recalcitrant data that comes from the use of recording devices. Following Sidelle (1991) I argue that recording devices allow agents to perform utterances at a distance. Using the essential, but widely ignored, distinction between tokens and utterances, I develop the view beyond the initial sketch given by Sidelle, and I answer the main objections raised against the view. The paper is structured as follows. Section 1 gives a succinct presentation of Kaplanian semantics and of the problem raised by the use recording devices, Section 2 presents the remote utterance view and Section 3 answers the objections put forward against the view and further develops it. I conclude that the remote utterance view can handle the data that comes from the use of recording devices with only modest modifications of the Kaplanian semantics.

Keywords Indexicals · Kaplanian semantics · Reference · Speech acts · Recording devices, Intentions

1 Introduction: Standard Kaplanian Semantics and the Answering Machine Puzzle

Formal semantic theories use formal languages as explanatory models of natural ones, whereby these models are supposed to represent the meaning properties and meaning relations of the natural language. Before any semantic work can begin, the theoretician must explain how the formal model makes contact with the linguistic reality it aims to capture. The primary data that semantic theories aim to predict are competent speakers' intuitions about the correctness of utterances of natural language sentences. Utterances

✉ Adrian Briciu
adrian.briciu@gmail.com

[1] Department of Logic, History and Philosophy of Science, LOGOS Research Group, University of Barcelona, C. Montalegre, 6-8, 4a Planta, 08001 Barcelona, Spain

that speakers have intuitions about are concrete events of language use. Each utterance event is represented in the formal model as a pair consisting of a sentence and a context. Contexts, as objects of the model, are supposed to represent the situations in which utterance events occur. At minimum they can represent the fact that utterances are performed by an *agent*, at a *time* and *place* in a *possible world*. Then, contexts are modeled as a sequence of particulars that include an agent, a location, a time and a world, which, following established use, I will call *contextual parameters*. Kaplan (1989: 509) proposes that the contexts available for semantic evaluation are restricted to *proper contexts*: contexts in which the agent of an utterance is at the place and time of the utterance in the world of the utterance.

Matching a particular utterance event with the appropriate sentence-context pair is no trivial matter. On one hand, the semantic machinery makes claims about syntactically and semantically disambiguated sentences, but, on the other hand, natural languages exhibit both lexical and syntactic ambiguity. Therefore, matching the sentence uttered in a particular event with the corresponding sentence in the semantic model requires syntactic and semantic disambiguation [1]. Likewise, matching a particular, concrete situation in which a sentence is uttered with a sequence of contextual parameters in the formal model is not an easy matter either, as it will become clear in the next sections[2].

It is customary, although not uncontroversial, to divide a theory that gives a description of the meaning properties of a natural language into two-parts: a theory that compositionally pairs each sentence of the language with its satisfaction conditions and a theory of the illocutionary forces with which those sentences are used. When it comes to compositionally pairing sentences with their satisfaction conditions, Kaplan (1989) advises us to distinguish two types of meaning: what an expression means independently of any context of utterance, and what an expression means relative to a context of utterance. He calls the first the *character* of an expression, and the second the *propositional content* of an expression. The character of an expression is the convention associated with that expression and something like a rule of use: it tells what an expression can say when used in any arbitrary context. For example, the character of the first person pronoun "I" can be thought of as a rule that states that "I" when used in an arbitrary context refers to the speaker of that context. Characters can be modeled as functions from possible contexts of utterance to propositional contents. For some expressions their character is a constant function that returns the same content relative to any context of utterance (i.e. numerals, proper names), but for indexicals ("I", "today", "here", "now") their character is a non-constant function that returns different contents for different contexts. The *propositional content* of an expression can be thought of as the information that determines the extension of that expression at possible states of affairs. Content can be modeled, then, as a function from possible worlds (which here are taken as circumstances of evaluation) to extension: individuals for singular terms, sets for predicates, truth-values for declarative sentences. The

[1] For example, the English sentence with the audible, or surface, form "Mary saw the boy with binoculars" will be matched with two sentences in the model:

(s_1) [Mary_N [saw_V[the_{DET}[boy_{CN}[with binoculars$_{AdjP}$]$_{NP}$]$_{NP}$]$_{VP}$]$_S$

(s_2) [Mary_N [[[saw_V[the_{DET} [boy_{CN}]$_{NP}$]$_{VP}$[with binoculars]$_{AdvP}$]$_S$.

[2] Predelli (2005: 23-34) and Kölbel 2009: 376-382) offer an extensive discussion of this. See also Kaplan (1989: 522-523) for a discussion.

content of some expressions is a non-constant function that returns different extensions at different circumstances of evaluation, but for indexicals their content is a constant function that returns the same extension for any circumstance of evaluation[3]. Since the content of indexical expressions is a constant function we can say that, in a sense, their character directly determines their reference relative to a context, and just for matter of expedience I will talk that way.

The fact that expressions such as "I", "here", "now" can have different referents on different occasion of use is captured in standard Kaplanian semantics in the following way. The character of "I" is modeled as a function defined on the set of possible contexts of utterance and which returns the agent of the context, the character of "now" as a function defined likewise and which returns the time of the context, and the character of "here" as a function defined likewise and which returns the location of the context.

Now, given that

(a) The referent of indexical expressions, relative to a context of utterance, is determined by evaluating its character to that context of utterance,
(b) The character of "I" is a function that yields the agent of the context of utterance and the character of "here" is a function that yields the location of the context of utterance and the character of "now" as a function yields the time of the context of utterance,
(c) The agent is always at the location of the context of utterance at the time of the context of utterance.

it follows that every utterance of (1) is true; in other words, (1) is true at every context of utterance.

(1) "I am here now"

Then, given standard semantics for negation, it follows that every utterance of (2) is false: there is no context of utterance at which (2) is true.

(2) "I am not here now"

But the pervasive use of recorded messages seems to indicate that there are true uses of (2). For example if we find a scrap of paper, or a post note, on someone's office door with (2) written on it, we judge (2) to be true if that person is not in her office at that time. Or if we call someone's house and (2) is played back on the answering machine, we judge it to be true if that person is not home at the time of the call. So, as opposed to what the standard semantics predicts, there are true uses of (2). This has been called by Sidelle (1991) the answering machine paradox for indexicals, although it is more of a puzzle than a paradox.

[3] Kaplan was moved to make this claim by the observation that when embedded under modal, temporal or locational operators, indexicals do not shift their reference. This can be seen in sentences like: "It is possible that in Pakistan, in five years, only those who are actually here now will be envied" where the world, location and time referred to by "actually", "here" and "now" are the world, time and location of the context of utterance, and not those determined by the modal, temporal and locational operators (Kaplan 1989: 499).

The task, then, is to extend or modify the standard Kaplanian semantics in a way that can handle this puzzle. Responses to the puzzle can be classified depending on whether they reject (a), (b) or (c). The account that I'll defend rejects (c)[4]. This account is a version of the remote utterance view, a view first proposed by Sidelle (1991). I'll develop and modify Sidelle's version of the view in ways that can handle criticism leveled against it[5].

2 The Remote Utterance View

2.1 Communicating with the help of recording devices

It is obvious that not all communication takes place face-to-face. Some communication is indirect in that it takes place through recording devices. There are at least two types of situations in which recording devices come handy. One is when the agent wants to communicate something about her time and place to an absent audience. This is the case when, for example, during a vacation in Paris one video-records herself in front of the Eiffel tower and says "I'm here now. I'm really impressed!" Intuitively, in this case one tries to communicate to an absent audience the content that the uttered sentence has at the time and place of the

[4] All these responses accept the data (i.e. the intuitions concerning the truth-values of (2) when used in recorded messages) as semantically relevant. Another type of response denies that the data is semantically relevant and tries to explain away the intuitions by appeal to pragmatic effects. According to this type of answer, the intuition that (2) is true when the message is played back is due to what is conveyed by that use of (2) and not because its propositional content is true at that context of use. On the contrary, when the message is played back, (2) is false *even though* the agent is not present at the time and place of the play back. It must be, then, that when played back (2) doesn't expresses the content that the agent is present at the time and place of the playback. Rather, the content expressed is the false content that agent is not present at the time and place of recording. As Cohen and Michaelson 2013: 581-582) convincingly argue, the pragmatic answer fails for several reasons: (i) unlike conversational implicatures, answering machine recordings of (2) cannot be canceled as shown by the infelicity of "I am not here now, but I might be when you are listening to this."; (ii) When (2) is played back on an answering machine, "understanding that the 'here' and 'now' refer to the place and time of playback is entirely non-optional, in that if one fails to interpret these indexicals in this way, one fails" to understand (2) Michaelson (2014: 535-536). There is a further reason to set aside the pragmatic response. The response assumes that when one records (2) on the answering machine, one performs an utterance. I'll argue in this paper that this is not the case.

[5] This paper has a rather limited scope: that of showing that Kaplanian semantics can handle the answering machine puzzle. Ever since the publication of Demonstratives many other objections, or proposed revisions, have been put forward for Kaplanian semantics. It is beyond the scope of this paper to deal with all of them. One other phenomenon that some authors take to require a revision of Kaplanian semantics, and is often discussed in conjuction with the answering machine puzzle, is the historical present: the use of "now" when narrating past events in order to refer to a moment in the past (e.g. as when during a history class about Napoleonic wars the professor says "Now, Napoleon prepares to cross the Niemen river into Russia", "now" doesn't refer to the time of the utterance, but it refers to 1812.) As Predelli (2005: 53) points out, the historical present is distinct from the challenge posed by recorded messages, but both are taken to motivate the same solution. This is debatable, though. Cohen (2013) and Corazza (2004) convincingly argue, in my opinion that occurrences of "now" in the historical present are not indexicals at all but bound variables. In the above sentence it is plausible that "now" is bound to the time of Napoleon's crossing of Niemen by an immediately preceding discourse fragment. Kaplan (1989:489-490) was well aware that words like "he", "now", "here", etc. have both an indexical use and a non-indexical use, and he explicitly limited his semantics to their indexical use. This fits well with the data presented in Nunberg (1993) that expressions like "that", "he", "here", "now" can be used indexically, anaphorically and as bound variable.

recording. We can say that, in this case, one records her utterance so that an absent audience can entertain it.

Another situation is when one wants to communicate something about a time and place at which she is absent, to an audience present at that time and/or place. This is the case when, for example, one inscribes (2) on a post note and places it to her office door before leaving the office. This is also the case when one records (2) on an answering machine at her home. Intuitively, in this case one tries to communicate to an audience the content that (2) has when uttered at the time and place of its decoding. An intuitive way to look at this situation is to say that one uses the recording device to perform utterances at times and place she is absent. This is, in fact, the gist of Sidelle's approach to the puzzle. Recording devices, be them post notes, letters, answering machines, videos, or of other type, give us the ability to act at a distance. More precisely, they give us the ability to perform utterances at space-time locations at which we are absent. As other types of actions show, it is not the case that one is always located at the place at which one performs an action, at the time she performs it[6]. Recording devices allow us to perform utterances at a distance, just as various other kinds of devices allow us to perform various other actions at a distance.

In brief, one can use recording devices either to record their utterances so absent audiences can entertain them, or one can use recording devices in order to perform utterances by their means.

2.2 The Remote Utterance view: The basic idea

Concentrating only on the second way of using recording devices, Sidelle (1991) argues that when someone records (2) on an answering machine, or writes it on a note, she is arranging to make an utterance at a later time, namely the time when the answering machine is played, or the note is read. In other words, when one records (2) on the answering machine she is not, at that time and place, making an utterance but merely makes preparations to perform an utterance at a future time. By means of the recording device, the agent genuinely utters (2) at the time when someone calls and the message is played back on the answering machine, even though she is absent at that time and place. If a genuine utterance is performed at the time of the playback, then "here" refers to the location of the utterance, and "now" refers to the time of the utterance. Which means that (2) is true at the time of the call, if the agent happens not to be at home at that time. This captures the intuition that there are true uses of (2).

When we record a message we do it with the intention, or at least the expectation, that the message will be triggered (i.e. played back, read, etc.) at a specific time, or within a certain time period, and at a specific place. As Sidelle puts it "there are parameters we set when making deferring utterances upon the situations in which the deferred utterances may occur" (Sidelle 1991: 527). Consider, for illustration, the case

[6] Sidelle (1991:535) gives this as an example of an action performed at a distance: one plants a bomb which much later on, when the agent is thousands of miles away, is detonated. The detonation is an action she performs at a distance.

of someone who shortly leaves her office for the nearby store and inscribes (3) on a post-note which she pins on her office door.

(3) I am at the store

Intuitively, her intentions in writing the note are such that an utterance of (3) is made only if the inscribed sentence is read at a certain time period (i.e. the time immediately after her departure) and place (i.e. her office). In other words, an utterance by means of a recording device occurs only at the time and place intended by the agent. To buttress this claim Sidelle asks us to consider what would happen if the note gets lost and is discovered only years later? His intuition is that no utterance is being made when the note is discovered and read. This seems right since making an utterance is the result of acting on one's communicative intentions, and the agent did not intend to utter (3) years after its inscription. Sidelle also claims that an utterance *is* made at the time intended by the speaker, despite the note not being read at the intended time (Sidelle 1991: 527). I believe that he is right to claim that no utterance is performed when one discovers the note years later, but it is open to debate whether an utterance is performed at the time intended by the agent, if the note is not read. After all, nothing relevant whatsoever occurs at the time and place intended by her. The reasons for why I believe that Sidelle is right about the first claim but why his second claim is open for debate will have to wait for the next section, where I introduce an important distinction and further develop the account.

The remote utterance view requires that we give up the claim that the agent needs to be at the location of the utterance at the time of the utterance. The positive part of the view is that it preserves intact much of the standard semantics of indexicals. More precisely, it preserves the standard mechanism of semantic evaluation: the referent of an indexical, relative to a context of utterance, is the result of applying the character of that expression to *that* context of utterance. Sentences are evaluated at contexts of utterance and not at contexts intended as relevant by the speaker and which are distinct from the context of utterance, as in Predelli (1998a, b, 2011) and Recanati (2005). Secondly, it allows us to give rather simple lexical entries for indexical expressions: the character for "I", "here" and "now" can be the standard kaplanian ones, as opposed to competing accounts that needlessly introduce multiple characters associated with indexical expressions (e.g. Smith, 1989; Michaelson 2014).

It should be noted that the remote utterance view is a theory about language and its use. It claims that Kaplanian semantics can handle the *prima facie* recalcitrant data that comes from using recording devices to communicate. It is built around the basic insight that agents can perform utterances at times and places they are absent, by using such devices. Its central claim (that an utterance is performed by means of a recording device at the time and place where a speech act is performed) is defended by appeal to speech act theory. The remote utterance view is not a theory about recording devices themselves, or more precisely, it is not a theory that makes predictions about the use of recording devices in non-linguistic actions. For example, the theory doesn't make any predictions on whether a singer using a playback in a performance is actually singing or not. Being a theory about speech acts it remains completely silent about non-linguistic actions where recording devices might be employed.

2.3 Tokens vs. Utterances

It is a trivial observation that in order to communicate with sentences we must utter them, and uttering a sentence involves using an acoustic, written or manual token of the sentence. Nevertheless, it is important to distinguish *tokens* from *utterances*: tokens are physical realizations of expression types; they are *concrete particulars* like noises, inscriptions, gestures and so on. Utterances are *intentional actions* performed by agents at space-time locations in order to communicate. That the two should be distinguished can be seen from the fact that one token can be used to make multiple utterances, as when at a protest a sign inscribed with (4) is passed from one protester to another.

(4) I support electoral reform

There, each protester utters (4) when they hold the protest sign, and they all do it with the very same token[7]. Conversely, an utterance can involve many tokens as when, for example, the queen of Denmark publishes her oath in multiple newspapers.

The distinction between the production of mere physical tokens and the performance of utterances holds the key to a full articulation of the remote utterance view. As Searle notes "people perform illocutionary acts: they make statements, give orders, ask questions, etc. [and that] in doing so they make noises, or marks on paper; they draw pictures, or wave their arms about etc." But "what must be added to these noises, marks, etc., in order that they should be statements, orders, etc.? What, so to speak, must be added to the physics to get the semantics? " (Searle 1986:209)

Utterances embody and manifest communicative intentions, while mere tokens don't. One can produce a token with no communicative intentions behind her action, as when one produces a token as a result of her interest in the phonetic properties of the token (for noises), or in its calligraphic and orthographic properties (for inscriptions). In other words, what differentiates utterances from mere tokens are the agent's intentions. Since utterances execute communicative intentions, for the use of a token to count as an utterance, the agent who uses the token must intend to communicate something by *that* use of the token. More precisely, the use of a token *t* of an expression type *e* is an utterance of *e*, if and only if the agent intends to express, by *that use* of *t*, a propositional content *p* with an illocutionary force *f*. The content expressed can, but need not, be determined solely by the linguistic conventions associated with the expression type *e*. [8]

The distinction between tokens and utterances can be exemplified in various ways. Consider Londoner Jill, who in order to improve her American English pronunciation decides to do some pronunciation exercises, and as a result she tokens (1). Her token of (1) is produced merely as a way of practicing American English pronunciation, and as such it

[7] Or, to borrow an example from Searle (1978:209), consider the case of someone who holds up the same sign inscribed with "Stop here" on several occasions. Each time she holds up the sign she utters the sentence "Stop here" and she does it with the very same token.

[8] The distinction between tokens as concrete particulars composed of ink, pixels, sound waves, hand movements, etc. and utterances as actions executing communicative intentions is found in other places in the philosophical and linguistic literature: Searle (1978:209), Bromberger (Bromberger 1992;191-193), Millikan (2012: 221), Perry 2001: 38-39), O'Madagain (2014: 75).

doesn't count as an utterance of (1). What is missing when she vocalizes (1) in practicing American English pronunciation is her intention to communicate something by way of her action. Contrast this with the situation in which Jill produces a token of (1) the moment she enters her office with the intention to inform her co-workers that she has arrived. In this case, her tokening of (1) counts as an utterance, for she intended to perform a speech act by means of that tokening.

The use of pre-recorded messages in public announcements lends support to this distinction. Imagine, for example, a voice actor who records the following line which is to be used in all train stations across the country:

> (5) "The train now approaching does not stop here. Please stay well clear of the edge of the platform".

Obviously, she doesn't intend to assert anything about a train approaching the studio at the time and place of the recording. Or imagine a factory which produces various signs used in public communication. One of the signs reads

> (6) "Do not go beyond this line".

Intuitively, at the time of the production, the worker who produces the sign is not asking anybody not to go beyond a line at the place of the production. Rather, she is crafting a token which can be used, later on by the appropriate authority to make requests or give commands.

In order to perform an utterance at a given time and place, an agent must intend to assert, or command, or request something, in short she must intend to perform an illocutionary act at that time and place. But neither the voice actor intends to assert something about a non-existent track ahead of her, nor does the factory worker intends to make a request about any given line located in her vicinity. They, instead, merely craft tokens that can be used to make assertions and requests, respectively. In other words, neither of them produces an illocutionary act at the time and place of the recording and, therefore, neither of them performs an utterance at their respective time and places.

The support for the token-utterance distinction, given by the use of pre-recorded messages in public announcements, doesn't essentially rely on the fact that an agent produces a token which is later on used by another one. Consider the following story: in order to ease her job, the station master of a one-person-operated railway station decides to record the public announcements made in her station. In a recording studio, far away from the railway station itself, she records numerous sentences, including (5). Later on, back at the railway station she uploads the recorded sentences on the station's automated public announcement system and programs it to play the sentences according to the station's schedule. She programs the system to play (5) every time a passing train approaches the station.

What happens when the station master records (5) in the studio? Obviously, she doesn't intend to assert that a train is approaching at that moment her current location (i.e. the studio) and that the train does not stop at that location. And if no train approaches the studio at that time we are not inclined to say that she lied when she records the message, Observe that if she is accused of lying or of saying something false, it is all too natural for her to reply that she did not, for she did not intend to communicate anything at that time and place. Given that communicative intentions are a necessary condition for utterances to occur, and that the

station master doesn't intend to communicate anything at the time and place of the recording, she doesn't perform an utterance at that time and place, but merely crafts a token of (5).

When the message is played back on the public announcement system, we intuitively take the station master to *assert* that a train is approaching the station and *request* passengers to stay clear of the platform. This is backed by the further intuitions. Imagine that the train is late, and it doesn't approach the station at the scheduled time. Even though she knows full well that no train is approaching the station, the station master decides to let the PA system do its job and play (5). There is a strong intuition that, in this case, the station master *lied* or, at least, *misinformed*, passengers. But if a speech act is performed when the train doesn't approach the station, then a speech act (that of *asserting* or *informing*) is also performed when the train approaches the station. There is another intuition that suggests that a speech act is performed at the time and place of the playback: if a passenger doesn't stay clear of the edge of the platform, we take her not to fulfill the station master's *request*.

These intuitions suggest that in the initial scenario two speech acts are being performed by the station master at the time of the playback: the assertion that a train is approaching the station, and the request that passengers stay clear of the line. And if the station master performs, by means of her recording, these speech acts at the time and place of the playback, and is causally responsible for that tokening of (5) then she utters (5) at that time and place.

Now, I'll spell out how the remote utterance view handles the paradigmatic cases put forward against Kaplanian semantics. Consider the case of answering machines. According to the remote utterance view, no utterance of (2) is performed when Jill merely records (2) on the answering machine. When she records (2) on the answering machine Jill merely crafts a token and does not utter (2) at that time and place. More precisely, since she doesn't intend to assert the content of (2), nor any other content, at the time and place of the recording, her production of the token doesn't amount to an utterance (2) at that time and place. However, when someone calls and (2) is played back on the answering machine Jill utters (2), because she made arrangements that a token of (2) is produced at that time and she intended to use *that* token to inform the audience that she is not present at her home at that time.

Suppose that Jill happens to be home when someone calls, but decides to let the machine go on and play (2). Did she lie to the caller? There is a robust intuition that, indeed, this is a case of lying. The fact that Jill lies when she is home and decides to let the machine play (2), shows that she is asserting (2) when she is not home and the machine plays back (2) [9]. But if she performs, by means of the machine, the act of lying when she is at home, and she performs, also by means of the machine, the act of informing when she is not home then in both situations she utters (2) at the time of the playback.

There are certain regularities in our use of answering machines, namely we use them with the intention that whenever someone calls, and the machine plays back the message, the caller is informed that the owner of the answering machine is not home. So is the case with Jill: she records (2) with the intention that whenever someone calls, and the machine plays the message, the caller is informed that she is not home at the time of the call. Then, as long as the machine is used for its initial purpose (i.e. it

[9] As Sidelle (1991: 553) puts it "If you lie when you are home, you must be telling the truth when you are not. It can't be because one is at home at the time of the machine; one can use the machine to utter "I am not here now". One being home at the time of the call simply provides one with the opportunity to use the sentence to lie".

doesn't change owner, it doesn't became part of an art installation, etc.) every time someone calls and the machine plays back (2), Jill utters (2) at the time of the playback.

What about when things don't go as planned? Consider, again, the situation when Jill leaves her office for the nearby store and inscribes (3) on a post-note which she pins on her office door. A strong wind, though, blows the note off before anyone can read it. The note is lost and discovered only years later.

(3) I am at the store

The remote utterance view should agree with Sidelle that no utterance of (3) is being made when the note is discovered and read. Jill inscribes (3) on the post-note with the intention that anyone who drops by her office, while she is at the store, and reads the note will be informed about her whereabouts at that time. Since agent's communicative intentions are necessary for a tokening to be an utterance, an utterance occurs only at times and places intended by its agent[10]. Being the performer of an utterance, Jill gets to choose what sentence to use in order to communicate what she intends to communicate, and gets to choose when and where to perform the utterance. In other words, no utterance of (3) by Jill takes place when the note is read outside the time and place at which she intended to express something by means of that note. In this situation as well, Jill is not made a liar when the note is found years later if she happens not to be at the store at that time, for her intention was to express something by means of that note only in the short time period after she left her office. And of course, no utterance is performed by Jill at the time and place she intended because her note is blown away by the wind and, as a consequence, there is no tokening of (3) at the time and place she intended.

Enriched with these considerations about how recording devices are used, and the distinction between tokens and utterances, the remote utterance view can easily answer the objections raised against version initially proposed by Sidelle (1991).

3 Objections against the remote utterance view

3.1 Decoding recorded messages

Several authors have argued that although the remote utterance view handles the answering machine case quite well, it faces challenges from other kinds of recorded messages, such as postcards or even letters[11]. Consider, for example, a postcard on which is inscribed only:

(5) It is beautiful here now.

According to them, the remote utterance view predicts, counterintuitively, that this use of (5) is true iff it is beautiful at the time and place where the postcard is read, not where it is written.

[10] What counts as a tokening (i.e. use of the token) is less clear when we deal with written messages (like post-notes) than it is when we deal with answering machine. There seem to be a consensus, though, that each instance of the post-note being read constitutes a tokening (i.e. use of the token). See Michaelson (2014: 251-253), Cohen (2014: 24)

[11] See various joint and individual papers by Cohen and Michaelson (Cohen and Michaelson 2013: 583, Cohen 2013: 24 fn. 23. Michaelson 2014:331-332), Corazza et al. (2004: 4) and Bianchi (2008: 312).

I believe that these authors are wrong about the predictions made by the remote utterance view. The remote utterance view is not committed to the claim that recording devices can be used solely to *perform* remote utterances. As I already mentioned there are at least two types of indirect communication for which recording devices come handy: (a) to record our utterances for audiences not present at that time and place and (b) to perform utterances at time and places we are not present. In the first case the agent performs an utterance at the time and place of the recording, and uses recording devices so that absent audiences can entertain her utterance. This is what's going on when one writes (5) on a postcard. The agent intends to assert that it is beautiful at the time and place of the inscription. In writing (5) on the postcard she performs an utterance that asserts exactly that. She uses the postcard so that her utterance reaches an audience that is absent at that time and place.

It is a legitimate question to ask: how does an audience know whether the recorded message is a recorded utterance (and thus the sentence needs to be evaluated at the time and place of the recording), or whether an utterance is being performed at the time the recording device is activated (and thus the sentence needs to be evaluated at the time and place of the decoding)[12]?

The audience's ability to identify what utterance was produced is probably explained mostly by appeal to concepts and abilities that are not essentially linguistic. Other than recognizing some marks as being a token of (5) her linguistic abilities play little role in her identifying what action (i.e. what utterance) was produced. Rather, the audience uses various contextual cues, her knowledge of practices that involve recording devices, and some form of plan recognition to correctly identify the utterance performed. There is nothing inherently linguistic about this[13]. This is the same ability by which an agent recognizes what particular action another agent performs.

[12] Michaelson (2014: 531) does raise this question for the remote utterance view.

[13] Corazza et al. (2004) argue that there are conventions associated with each type of recording device which determine whether (5) is to be interpreted at the context of encoding or at the context of decoding (i.e. context of reading, hearing, etc): if (5) is heard on an answering machine, then the relevant time and place for its interpretation are the time and place of the playback; if it is read on a postcard or letter the relevant time and place are those when the postcard or letter was written. But this is doubtful. At best certain usages are more common than others. And certainly the remote utterance view need not adhere to it. For sure, both answering machines and postcards/letters can be used in ways that go against the conventions stipulated by Corazza et all, and these uses are neither linguistically nor socially deviant. Imagine that one records the following sentences on an answering machine: "I decided to give up everything. I'm signing the donation agreements right now. From now on this house no longer belongs to me". The time relevant for interpreting these sentences is the time of the recording and not the time of decoding. As far as letters are concerned, imagine that one writes the following in a letter: "You are about to read my side of the story. Today, you'll finally understand why I decided to leave". The time relevant for interpreting these sentences is the time of the decoding/reading and not the time of the encoding/writing. Moreover, people can successfully communicate with the help of *novel* recording devices which are not governed by social conventions, as when one video records their will. To borrow an example from Sherman (2015), one can video record either "Today I met with my lawyer to go over all the details before making this videotape" and "Today you all received a phone call telling you to come to my lawyer's office." The time relevant for interpreting the first sentence is the time of the recording, while the time relevant for interpreting the second sentence is the time of the decoding. The remote utterance view can easily account for this and it doesn't need to appeal to conventions. Both sentences recorded on the video will are evaluated at their respective contexts of utterance. Video wills, like all recording devices, can be used both to record and to perform utterances. When one video records "Today I met with my lawyer to go over all the details before making this videotape" on her will she records her utterance for a future audience. When one records "Today you all received a phone call telling you to come to my lawyer's office" on her video will, she makes preparations so to utter the sentence when the video is being played to a future audience.

Importantly, the remote utterance view is not committed to the claim that recording devices can be employed only to perform utterances at a distance. And, therefore, it is not committed to the view that a recorded sentence must always be interpreted at the time and place of the decoding. For it is obvious that recording devices can also be employed to record one's utterances so that they can reach audiences that are absent at the time and place of recording, as when one video-records (5) in front of the Eiffel tower and sends the recording to her family; or when the queen of Denmark records her oath in print. In other words, whether (5) is interpreted at the time and place of the recording, or at the time and place of the decoding depends on whether (5) is uttered at the time and place of the recording or whether it is uttered the time and place of the decoding. The audience's ability to discern the right interpretation partly depends on her (non-linguistic) ability to discern the agent's action: when and where (5) was uttered.

3.2 Misfiring intentions

The main objection against the remote utterance view comes from Predelli. I should point out that Predelli doesn't deny the possibility of remote utterances (see Predelli 2005: 61); he denies that the view can account for all our intuitions about recorded messages. He imagines the following scenario. John, before leaving home at 8AM, writes a note to his companion Jill who, usually, is back home at 5PM:

(6) As you can see I am not here now. Meet me in two hours at Cabo Real.

Predelli rightly points out that when read, the note does not convey the false content that John is not at home at 8AM (the time the note was written), nor does it request Jill to be at Cabo Real at 10 AM (two hours from the time of inscription). Given what is common knowledge among them (i.e. that Jill is back home at 5PM) John inscribes (6) with the intention to inform Jill that he is not home at 5PM, and that he wants to meet her at Cabo Real at 7PM.

As it happens Jill is late, she arrives home at 8PM and reads the note only then. Intuitively, (6) must be evaluated with respect to the expected time of decoding, namely 5PM since John intends to inform Jill that she should meet him at 7PM, and Jill would misinterpret the sentences written on the note were she to conclude that John will be waiting for her at Cabo Real at 10 PM. According to Predelli (1998b:405) the deferred utterance view is unable to give this "pre-theoretically desired outcome", because it fails to provide a context of utterance that has 5PM as its temporal parameter. More exactly, Predelli's objection is that the remote utterance view cannot account for our intuition that "here" and "now" refer to John's house and 5PM, respectively (Predelli 1998a: 118). According to him, no utterance takes place in John's house at 5PM, since no relevant event occurs at that time at that place. Then, there is no *context of utterance* at which to evaluate (6) and, therefore, (6) remains semantically incomplete. Then, the indexicals inscribed on John's note should lack reference. This, though, doesn't bode well with the intuition that they do have a referent, namely the time and place that John intended his note to be read: 5PM and his house.

The intuitions raised by Predelli's scenario are robust and must be accounted for. But, *pace* Predelli, I believe that the remote utterance view can give predictions that accord with these intuitions[14].

There are two ways in which the remote utterance view can answer Predelli's objection. One way is to reject his claim that no utterance of (6) takes place at the time and place intended by John, namely at 5PM and his home. It can be argued that John does utter (6) at 5PM at his house because he tokens (6) at that time and place, since he caused that token to be at that time and place, and he has the intention to use *that* token to utter (6) precisely at that time and place. The fact that the audience is absent doesn't mean that John didn't make an utterance of (6) at that time and place. This seems to be the position favored by Sidelle (1991) and more recently by O'Madagain (2014). If this is correct then Predelli's argument doesn't even take of the ground, since an utterance of (6) is, in fact, produced at 5PM in John's house. Although this is an interesting approach, I won't pursue it here, for I believe that it runs into trouble in explaining how an utterance of (6) occurs at that time and place. Moreover, Predelli's claim that no utterance of (6) by John takes place at 5PM has strong intuitive pull, and most authors accept it[15].

A second way to answer Predelli's objection is available to the remote utterance view, and I find it more plausible. The remote utterance view can accept that no utterance of (6) occurs at the time and place intended by John and, *pace* Predelli, can still account for the intuitions raised by his scenario. In fact, the remote utterance view can explain our intuitions even if it agrees that no utterance whatsoever is performed by John in the imagined scenario: no utterance at the time and place intended by him, since no event whatsoever occurs at that time and place, and no utterance at the time Jill reads the note, since John did not intend to perform an utterance at that time.

John's situation can be described as one in which his intention misfires. The speech acts that he intended to perform by means of his note are not achieved because of a hitch in the execution of the acts: no relevant event takes place at the intended time and place. For the sake of clarity let's keep separate the two sentences inscribed on John's note, since he intends to use each of them with a different illocutionary force: (6a) to make an assertion and (6b) to make a request.

(6a) As you can see I am not here now.
(6b) Meet me in two hours at Cabo Real.

My disagreement with Predelli lies in the way in which he describes the datum that the theory needs to account for. According to him, the datum is "the insight that "here" and "now" in John's note refer to" John's house and to 5PM, respectively (Predelli 1998a: 111). This, unfortunately, is a rather loose way of presenting the datum. Once we are methodologically more rigorous it will be easy to see that the remote utterance view can account for the datum. I will argue that the intuitions elicited by Predelli's

[14] I should point out that not all authors accept Predelli's intuitions. For example, Cohen (2013) whose position is that recorded expressions are semantically interpreted always at the context of decoding (i.e. context of reading, hearing, etc) claims that the indexical "now" in the note refers to 8 P.M. and not to 5 P.M. I take this to be a serious drawback of Cohen's position.
[15] See Cohen and Michaelson (2014) for an overview.

scenario are intuitions about the reference of "here" and "now" in a possible but never actualized utterance of (6a) that John intended to perform by means of his note.

One uncontroversial type data that semantic theories aim to predict are intuitions about correctness of use[16]. Competent speakers are able to respond to uses of sentences and say whether these uses are correct or not – and if they are not correct, competent speakers are able to say under what conditions would they be correct. There are different dimensions along which the use of a sentence can be said to be correct or not. To illustrate with the help of declarative sentences: one dimension of correctness concerns the fact that one can use a declarative sentence to make an assertion only if certain conditions are met (e.g. she is sincere, and so on) and competent speakers have intuitions on whether a particular use of a sentence succeeds or fails to be an assertion. A second dimension of correctness concerns the fact that an utterance of a declarative sentence is true only under certain conditions, and speakers have intuitions about the conditions under which an utterance of the sentence is true. Of course, competent speakers' intuitions are not limited to declarative sentences. There are, then, at least two types of intuitions which semantic theories seek to predict: intuitions on whether an utterance of a sentence succeeds or fails to be a certain type of speech act, and intuitions on whether an utterance of a sentence satisfies the sentence's satisfaction conditions. Importantly, for reasons having to do with the productivity of natural languages, semantic theories aim to predict not only intuitions about actual utterances but intuitions about possible utterances as well. My claim is, then, that the data in Predelli's scenario are, in fact, intuitions about possible utterances of (6a) and (6b) that John intended to perform but that failed to actualize due a glitch in execution. Intuitions about the conditions under which a particular propositional content can be asserted using a given sentence are at play in Predelli's scenario.

In describing the scenario Predelli makes us privy both to John's plans and intentions (that of *informing* Jill that he is not home at 5PM and *asking* her to meet him at 7PM) and to what is common knowledge between John and Jill (that she is usually home at 5PM). Given what we know about John's intentions and expectations we have the intuition that he can assert or inform, by way of (6a), that he is not home at 5PM if and only if he utters (6a) at 5PM. More precisely, asserting that John is not home at 5PM, by means of (6a) can only be achieved if (6a) is uttered by John at his home at 5PM. Likewise, given what we know about John's intentions and expectations we have the intuition that in order for him to request Jill, by way of (6b), to meet him at 7PM John must utter (6b) at 5PM. In other words, the request that Jill meets John at 7PM can only be achieved by means of (6b) only if (6b) is uttered by John at 5PM. The intuitions raised by Predelli's scenario about the referents of "here" and "now" inscribed on the note are best described as intuitions about a possible but never actualized utterance of (6).

Predelli's objection is, basically, that since no utterance of (6) occurs at 5PM in John's house, the remote utterance view cannot account for our intuitions. He seems to presuppose that the remote utterance view can account only for intuitions about actual utterances. But why should it be so? Speakers have intuitions both about actual and possible utterances, and Kaplanian semantics (like any semantic theory) can give

[16] For the role of intuitions about use as primary data in semantic theorizing see Kölbel (2009) and Martí (2012)

predictions about both actual and possible utterances of English sentences. (Remember, the remote utterance view is just Kaplanian semantics minus the claim that all semantically relevant contexts are proper ones.) Kaplanian semantics predicts that if uttered at a context that has 5PM as its temporal parameter "now" will refer to 5PM; and if uttered at a context that has John's house as its location-parameter "here" will refer to John's house, and if uttered at a context that has John as its agent "I" will refer to John. Given the syntactic structure of (2) and standard semantic composition, if (2) is evaluated relative a context consisting of the above parameters, (2) expresses the propositional content that John is not at his home at 5PM. This correctly predicts the intuitions raised by Predelli's scenario[17].

I would like to end with a few remarks about the role of speaker's intentions in determining the reference of indexical expressions. There is a vast debate in semantics over what determines the reference of an indexical expression on a given occasion of use. According to *conventionalism*, the reference is determined solely by the linguistic convention associated with the expression and the context of utterance. According to *intentionalism*, the reference is determined partly by speakers' intentions. Where does the remote utterance view fit within this debate? The way I see it, according to the remote utterance view, speakers' intentions play no role in the semantic evaluation of indexical sentences. They do play a role, though, in determining the utterance produced, namely they determine what sentence is to be uttered, and determine when and where to utter it. And this is what we should expect since utterances are purposeful, intentional acts. But intentions play no role in the semantics of indexicals: they do not determine, from one occasion of use to another, the character that governs that particular use as in Smith (1989), nor do they determine the context at which indexicals are to be semantically evaluated as in Predelli (1998a, b, 2011) and Recanati (2005). Moreover, the remote utterance view allows us to keep a rather simple form of conventionalism: one according to which expressions like "I", "here" and "now" have each a single convention associated with them (i.e. "I" the convention that it refers to the agent who utters it, "here" the convention that it refers to the place where it is uttered, and "now" the convention that it refers to the time when it is uttered). This stands in contrast with conventionalist accounts that posit multiple linguistic conventions for every indexical expressions, where each convention governs the reference of the indexical depending on the communicative channel (i.e. type of recording device) used to deliver the message: one convention governs their reference in face-to-face contexts, a different convention governs when they are played back on answering machines, a yet different one when they are or written on postcards or letters, and so on (see Michaelson 2014). This is not the place to go over the problems faced by the multiple-character type of conventionalism, but suffices to point out that a standard conventionalist position (one which associates expressions with only one character) that accounts equally well for the data should be preferred because it is, if nothing else, more economical.

[17] To make his case Predelli must show that it is *metaphysically impossible* for John to utter (6) at his home at 5PM by means of the note, and that yet we have intuitions about the truth-value of such an impossible utterance. But surely there are *possible worlds* in which John does utter (6) at 5PM at his home by means of the note: those in which someone is home at 5PM and reads the note.

4 Conclusions

In summary, I have argued that the remote utterance view offers an intuitive and elegant solution to the answering machine puzzle. The view can easily answer the objections raised against it once it takes on board two observations: (a) that recording devices can be used both to record utterances and to perform utterances, and (b) that utterances are distinct from tokens. The view requires only a modest modification of Kaplanian semantics, namely that we give up the requirement that proper contexts are the only semantically relevant contexts of utterance. This comes only natural if we take contexts to model real speech situations and we look beyond speech that occurs in face-to-face communication. Of course, future work should spell out in detail a Kaplanian semantics that gives up the commitment to proper contexts[18]. Moreover, the remote utterance view nicely spells out the insight that the purpose of recording devices is to allow agents to assert, request, order, and so on, in situations in which face-to-face communication is not possible.

References

Bianchi, C. (2008). Indexicals, Speech Acts and Pornography. *Analysis, 68*(4), 310–316.

Bromberger, S. (1992). *On What We Know We Don't Know: Explanation, Theory, Linguistics, and How Questions Shape Them*. Chicago: University of Chicago Press.

Cohen, J. (2013). Indexicality and the Puzzle of the Answering Machine. *Journal of Philosophy, 110*(1), 5–32.

Cohen, J., & Michaelson, E. (2013). Indexicality and the Answering Machine Paradox. *Philosophy Compass, 8*(6), 580–592.

Corazza, E. (2004). On the alleged ambiguity of 'now' and 'here'. *Synthese, 138*(2), 289–313.

Corazza, E., Fish, W., & Gorvett, J. (2004). Who Is I? *Philosophical Studies, 107*(1), 1–21.

Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 565–614). Oxford: Oxford University Press.

Kölbel, M. (2009). The evidence for relativism. *Synthese, 166*(2), 375–395.

Martí, G. (2012). Empirical Data and the Theory of Reference. In P. Kabasenche, M. O'Rourke, & M. Slater (Eds.), *Reference and Referring. Topics in Contemporary Philosophy* (pp. 63–82). Cambridge: MIT Press.

Michaelson, E. (2014). Shifty characters. *Philosophical Studies, 167*(3), 519–540.

Millikan, R. (2012). Are there mental indexicals and demonstratives? *Philosophical Perspectives, 26*(1), 217–234.

Nunberg, G. (1993). Indexicality and deixis. *Linguistics and Philosophy, 16*(1), 1–43.

O'Madagain, C. (2014). Indexicals and the Metaphysics of Semantic Tokens: When Shapes and Sounds become Utterances. *Thought: A Journal of Philosophy, 3*(1), 71–79.

Perry, J. (2001). *Reference and Reflexivity*. Stanford: CSLI Publications.

Predelli, S. (1998a). I am not here now. *Analysis, 58*(2), 107–115.

Predelli, S. (1998b). Utterance interpretation and the logic of indexicals. *Mind and Language, 13*(3), 400–414.

Predelli, S. (2005). *Contexts: Meaning, Truth, and the Use of Language*. Oxford: Oxford University Press.

Predelli, S. (2011). I am still not here now. *Erkenntnis, 74*(3), 289–303.

Recanati, F. (2005). Indexicality, Context, and Pretense: A Speech-act theoretic Account. In N. Burton-Roberts (Ed.), *Advances in Pragmatics* (pp. 213–229). Basingstoke: Palgrave-Macmillan.

Searle, J. (1978). Literal meaning. *Erkenntnis, 13*(1), 207–224.

[18] One immediate consequence of giving up the restriction of semantic evaluation to proper context is that the logic of indexicals developed by Kaplan (1989) must be slightly rewritten. But Stojanovic (2011) argues convincingly there are independent reasons for a revision of Kaplanian logic of indexicals, one in which (2) doesn't come true at every context of utterance.

Searle, J. (1986). Meaning, communication, and representation. In R. Grandy & R. Warner (Eds.), *Philosophical Grounds of Rationality: Intentions, Categories, Ends* (pp. 209–226). Oxford: Oxford University Press.

Sherman, B. (2015). Constructing Contexts. *Ergo, an Open Access Journal of Philosophy, 2*(23), 581–605.

Sidelle, A. (1991). The Answering Machine Paradox. *Canadian Journal of Philosophy, 21*(4), 525–539.

Smith, Q. (1989). The Multiple Uses of Indexicals. *Synthese, 78*, 167–191.

Stojanovic, I. (2011). Context, Meaning and Logical Truth. In A. Gupta & J. van Benthem (Eds.), *Logic and Philosophy Today*, *Studies in Logic Series, vol. 30* (pp. 11–24). London: College Publications.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com