Croatian Journal of Philosophy
Vol. XVI, No. 47, 2016

Coordinating with Language

JESSICA KEISER
Yale University, New Haven, USA

Linguistic meaning is determined by use. But given the fact that any given expression can be used in a variety of ways, this claim marks where metasemantic inquiry begins rather than where it ends. It sets an agenda for the metasemantic project: to distinguish in a principled and explanatory way those uses that determine linguistic meaning from those that do not. The prevailing view (along with its various refinements), which privileges assertion, suffers from being at once overly liberal and overly idealized. By parsing the most prominent aims we use language to achieve, noting their relations of dependence and the specific type of uses they involve, I arrive at a novel metasemantic account: facts of linguistic meaning are determined by locutionary action.

Keywords: Metasemantics, convention, illocutionary acts, speech acts, meaning.

> *If L is a possible language, and P is a population of agents, what relation must hold between L and P in order to make it the case that L is an actual language of P? The obvious answer is: the members of P must use the language L. But that answer is neither informative nor clear. They must use L in a certain way. If everyone in P used L by telling lies in L or by singing operas in L (without understanding the words), they would be using L but not in the right way; L would not be their language.* David Lewis (1969: 176–177)

> *I completely agree with Wittgenstein as far as he goes on that; the trouble is that he stops just short of the problem. I also think that the meaning [of an expression] is the use… I think this is one sufficiently broad characterization to cover the whole lot; the problem is, it's too broad… so it's a question of what use counts and what use is irrelevant and should be dismissed.... and there's where the problem begins, not where it ends.* W.V. Quine (2013)

Metasemantics is the study of what grounds facts of linguistic meaning. While it is uncontroversial that these facts are determined by use, given the myriad ways that any given expression can be used, this claim cannot be taken as ending systematic metasemantic inquiry; rather, it specifies its point of departure. One fundamental task of the metasemantic project is to provide a principled account of the sorts of uses that are meaning-determining, distinguishing them from uses that are not.

An initially plausible proposal, some variant of which has been either explicitly endorsed or tacitly assumed by the vast majority of philosophers and linguists for decades or more, is that *assertion* grounds facts of linguistic meaning. However, this view—along with its various refinements—suffers at once from being overly idealized and overly liberal. It fails both to account for standard and ubiquitous ways of using language that don't fit the idealized model of cooperative information exchange, and to sift out the kinds of use that are not relevant to determining linguistic meaning.

I offer a novel metasemantic account: facts of linguistic meaning are determined by *locutionary* action. My strategy is to parse the most prominent aims we use language to achieve, noting their relations of dependence and the specific type of uses they involve. This allows me to identify uses that meaning-determining and distinguish them in a principled and explanatory way from uses that are not. The resulting account provides a suitably robust—yet sufficiently fine-grained—account of linguistic meaning, while shedding light on the diverse roles that different kinds of uses will have in an overall theory of communication. Moreover, the account yields a new foundation for traditional topics in semantics and pragmatics which cannot be fully successfully theorized about independently of metasemantic considerations.

1. *Coordination Problems and Convention*

We use language as a tool for achieving a number of different aims. Many of these aims are situated in an order of dependence, in the sense that success in achieving one aim is a prerequisite for success in achieving another. We use language in different ways, corresponding to the different aims we are trying to achieve; this can lead to confusion in metsemantic theorizing, since not all of these uses are relevant to the question of what grounds facts about linguistic meaning. It is crucial, then, to get clear on the way these different ways of using language are related, and at what level each should appear in a full account of linguistic communication.

I will be making the simplifying assumption that language use is conventional, that a population speaks a certain language just in case there is a convention among them of performing a certain type of action.[1] This simplifies the task guiding metasemantic inquiry to that

[1] Most of my arguments (with the exception of one I make in 2.4) will not rely on this assumption—rather, they concern which type of action should appear in

of identifying the relevant action type, which—I will claim—is locutionary action. The notions of convention and coordination problems can be found in standard introductory economics textbook, and will be familiar to scholars working in any discipline that makes use of game theory. In the absence of uncontroversial yet rigorous definitions of these notions, I will be utilizing Lewis' account of both throughout the paper. Though there have been quibbles over the finer points of Lewis' definitions, they should nonetheless suffice here to provide a clear framework within which to talk about conventional uses of language.

A *coordination problem*, according to Lewis, is a situation in which two or more agents have a common interest in performing the same one of several alternative actions (or beliefs).[2] They must coordinate in order to achieve a mutually desired outcome, or equilibrium state. Coordination problems can theoretically be solved by pure luck; it is in principle possible that agents could inadvertently reach an equilibrium state by accidentally slipping on a banana peel or falling into a ditch. In most cases, however, coordination problems are solved by reasoning about another agent's mental states, predicting her actions, and then choosing one's own actions accordingly. In this way, agents are able to find a stable solution to a recurrent coordination problem by using the same strategy repeatedly on the expectation that others will do the same; in other words, a *convention*:

> A regularity R in behavior is a convention for a population P if it's common knowledge among the members of P that they follow R, that they expect one another to follow R, that R is a solution to the coordination problem that they face in S, and that there is another solution R′ that they could have conformed to instead of R. (Lewis 1969: 76)

In what follows I will tease apart several important coordination problems that we use language to solve, which happen to be nested inside of each other in the sense that solving one problem is a prerequisite for solving another. I think of these nested coordination problems like lay-

an account of linguistic meaning, regardless of whether the notion of convention is employed in such an account.

[2] In *Convention*, Lewis defines coordination problems in terms of action, going on to define conventions as regularities in action that are solutions to coordination problems. In "Language and Languages" he changes his account of convention so that the regularity is one of action or belief but does not make the relevant adjustment to his definition of coordination problem, which simply does not come up in that paper. This change in his account of convention was needed to accommodate the fact that in "Language and Languages" Lewis modifies his earlier account of language use as a convention of truthfulness (which is an action) to that of a convention of truthfulness *and trust* (which is belief). In this paper I will be using Lewis' amended account of language use and convention, which are given in terms of action *and belief* rather than merely action. Though he did not explicitly make the corresponding changes in his account of coordination problems, I do so here and consider this to be in keeping with his later views. For ease of exposition, I will often talk about coordination problems and conventions in terms of action only, but it should be implicitly understood that belief is also included (Lewis 1969: 24).

ers of an onion—peeling back one layer exposes another coordination problem underneath. I will suggest that up until this point, theorists of language have failed to peel the onion back far enough. The task is to peel back the layers of irrelevant use, stopping just at the point where the kind of use we have identified is robust enough to do the job of grounding linguistic communication. In what follows, I will consider roughly three layers of coordination problems—not because they are exhaustive, but because the first two correspond to the main metasemantic theories that have either been explicitly proposed or tacitly assumed in the literature up until now. The third, I will argue, is the one that corresponds to the *correct* metasemantic theory.

2. *Coordinating with Language*

Before teasing apart these different coordination problems and examining them more carefully, I'll first set up a bit of terminology. I will be using a broadly Austinian taxonomy of speech acts: The *locutionary* act is—roughly for now, but to be precisified later—to perform an utterance simply in order to transmit its conventionally encoded content (Austin 1975: 94–107).[3, 4] The locutionary act is successful when the audience entertains the content that the utterer has in mind. But merely transmitting a content rarely serves our goals—we often want to signal that we take some sort of attitude toward the content. Perhaps we want the audience to think that we believe a proposition, for instance, or to understand that we desire her to answer a question. So the locutionary act is generally in service of an *illocutionary* act; the performance of an utterance with some kind of force, such as assertion, command, query, etc. The illocutionary act is successful when the audience understands not only the content the speaker transmits, but the force that speaker attaches to that content. The illocutionary act, in turn, is usually performed in the service of the *perlocutionary* act; roughly, the act of doing something *by* uttering an expression, such as getting someone to close a door or to share one's beliefs, etc. The perlocutionary act is successful if the audience responds in the appropriate/desired way to the illocutionary act of the speaker, i.e., by producing the intended perlocutionary effect.

2.1. *Coordinating illocutionary act with perlocutionary effect*

Arguably, the aim of most ordinary language tends to be the production of perlocutionary effects—such as belief, action, and testimony. We navigate the world better when we are able to exchange information and

[3] For ease of exposition I will ignore context sensitivity here and make the simplifying assumption that the meanings of sentences are complete thoughts rather than functions.

[4] Some readers may find my interpretation of Austin to be controversial. While I believe I have provided a reasonable regimentation of Austin's taxonomy of speech acts, I don't have the space to defend that claim here. However, the fidelity of my interpretation of Austin should not be relevant to the arguments in this paper.

coordinate our behavior; it is hard to imagine how a community would have developed the use of language if not from pressure arising from this need.[5] This may explain early theorists' attraction to a metasemantic story in which perlocutionary effects play a primary role in linguistic convention. David Lewis' theory is a paradigm example of such an account: he worked out a rigorous metasemantic proposal, the general spirit of which is perhaps tacitly assumed by theorists who take the potential perlocutionary effect of an expression (e.g., change in context) to be indicative of its semantic value (e.g., context change potential). According to Lewis, parties in a conversation have a mutual interest in one member performing a certain perlocutionary response conditional on a certain state of affairs obtaining. They are able to achieve an equilibrium state if the utterer performs her illocutionary action only under certain circumstances (for instance, she will assert only true things, and command only actions that—if performed—will serve the interlocutors' relevant mutual interest) and the audience coordinates her perlocutionary response with the action of the utterer (by believing the content of the assertion, or obeying the command).[6] For instance, perhaps A and U ultimately have a mutual interest in both of them being able to unlock the house, though there is only one set of keys; then they will also share the more immediate interest for A to *believe* that the housekeys are hidden in the birdfeeder conditional on its being the case. Or for A to *hide* the housekeys in the birdfeeder conditional on its being the case that U will look for them there. Therefore, U will assert that the housekeys are hidden in the birdfeeder just in case they are, and she will command A to hide them there just in case that's where she is planning to look for them. In turn, A will respond to U's assertion that the housekeys are hidden in the birdfeeder by believing it, and will respond to her command to hide them there by doing it. Coordinating their action in this way serves their mutual interests: it allows them to navigate the world by exchanging information and acting together to achieve mutual goals.

While the motivation for this picture is clear, it does not provide a satisfactory basis for a theory of linguistic communication. As Austin

[5] I am uncommitted with respect to what was in fact the causal history of actual language use. Though, for instance, Noam Chomsky denies that language evolved as a response to a need for communication or information exchange, I take it to be uncontroversial that this is a primary function of language and strongly influences the development of individual languages over time. Moreover, in downplaying the conventionality of language, Chomsky is focused on syntax rather than semantics or pragmatics.

[6] Lewis uses a technical account of truthfulness and trust that is designed to describe the coordination between conversational participants in a way that generalizes across illocutionary act types. In each case one party is being truthful and one party is being trusting (in ways that will be cashed out differently depending on the type of illocutionary act). I will not go into the details of this account; what matters to the discussion here is that in each case the coordinating action on the part of the audience will be to perform a perlocutionary response (to believe the content of the assertion, fulfill the request, answer the question, etc.)

and Grice later started to point out, we can put language to many uses, not all of which determine conventional meaning or are even relevant to communication. What speaks decisively against the particular picture we are considering here is Austin's insight that communication has been successful when the hearer has identified the illocutionary action, regardless of whether she performs the corresponding perlocutionary response. If she knows what has been asserted, there has been successful communication regardless of whether she believes it. If she knows what has been commanded, there has been successful communication regardless of whether she does it. If she knows what has been asked, there has been successful communication regardless of whether she answers it, and so on. Furthermore, it seems that communication takes place even in situations in which we are actively and openly being *uncooperative* at the level which involves perlocutionary uptake.[7] So the lesson gleaned from Austin and Grice is that the layer of coordination involving perlocutionary action is irrelevant. The natural next step would be—and indeed *was*—to peel the onion back a layer and look instead to the problem of identifying the illocutionary act as a candidate for grounding language use. Something like this type of picture remains to be predominantly assumed in the literature today, though—as I will argue below—it is does not go far enough in stripping away irrelevant use.

2.2. *Coordination on identifying illocutionary act*

In order to coordinate her perlocutionary response with the speaker's illocutionary act, the audience must first identify which illocutionary act has been performed. Before forming a belief, for instance, she must understand what the speaker has asserted. Before answering she must understand the question and before obeying she must understand the command. This poses a separate coordination problem; the speaker and audience both want the audience to understand which illocutionary act the speaker is performing. They can solve the problem by relying on expectations and beliefs about the other's behavior: the speaker will generally encode her illocutionary act based on her expectations about how the audience will respond to her utterance, and the audience will generally form her beliefs about the speaker's illocutionary act based on her expectations of how the speaker will encode it. This is a coordination problem nested within a coordination problem, since it is only *after* the audience has successfully identified the speaker's illocutionary act that she can respond appropriately. But as we have seen, it seems as though successful communication occurs when the illocutionary act is identified, regardless of whether the audience responds in the appropriate or desired way. Indeed, agents need not even have an *interest* coordinating at the perlocutionary level in order to communicate suc-

[7] See Camp (forthcoming), McKinney (manuscript), and Pinker, Nowak, and Lee (2008).

cessfully. Such considerations make it appealing to look instead to the problem of identifying the illocutionary act to do the work of grounding linguistic meaning.

This general kind of picture which grounds meaning in illocutionary action is reflected in Grice and those following a similar framework, such as Stephen Schiffer and Brian Loar.[8] Differences in terminology can be misleading here: these philosophers took linguistic meaning to be grounded in acts of speaker meaning, but because they define speaker meaning in terms of a manifest intention to produce a propositional attitude (or other type of response) in the audience, it falls into the category of what I am calling *illocutionary action*, rather than what I am calling *speaker meaning*. Though Grice and his followers do not explicitly talk about the identification of the illocutionary action as a coordination problem, this way of framing things fits quite naturally with a Gricean account of communication. One of the key features of Grice's notion of speaker meaning, which sets it apart from natural meaning, is the involvement of coordination in choosing and interpreting the utterance. When a speaker *means p* she succeeds in communicating not simply by *causing* a belief, but by performing her utterance in a way that is meant to reveal her mental states, and to thereby provide the basis for the audience's reaction. In turn, the audience identifies the speaker's meaning not purely on the basis of external factors, but on her recognition of the speaker's intentions in producing that utterance, which serves as the basis for her response. The aspect of Grice's account of speaker meaning which makes it a fundamentally cooperative act is the requirement that the speaker intends for her communicative intention to be manifest to the audience—the recognition of which is meant to provide the audience a reason for identifying the speaker's illocutionary act one way rather than another. This kind of manifest, or reflexive intention functions as a mechanism for coordination. Though it is framed in a different way, the idea that communication fundamentally involves coordination by reasoning about each other's mental states and acting accordingly is as central to Grice's theory as it is to Lewis'. I take the general Gricean framework, then, to be a paradigmatic case of a metasemantic picture according to which language is grounded in conventions of illocutionary action.[9] Something in the vicinity also seems to be tacitly assumed by semanticists who posit illocutionary update potential as semantic values of expressions, as well as those who take illocutionary force to be relevant to locating the semantics/pragmatics boundary. As far as I can tell, this general type of picture seems to be widely taken for granted, but I will argue below that it is untenable.

[8] See Grice (1989), Loar (1981), Schiffer (1972).

[9] Grice had hesitations about appealing to convention in his account of linguistic meaning, and instead talked about members of a population having a certain procedure in their repertoires; however, because his account still grounds linguistic meaning in illocutionary action, the difficulties I outline here still apply (Grice 1989: 123–127).

This general approach runs into problems with something that is sometimes (somewhat disingenuously) called “non-standard speech”—specifically, cases in which language is used to illocute content that is non-identical to the standing meaning of the utterance, as well as cases in which language is used to do something else entirely from performing illocutions. Lewis was aware of these difficulties, which apply to his account as well (insofar as it grounds linguistic meaning *at least in part* in illocutionary action); he considers both cases in *Languages and Language*, though I will argue that his responses are inadequate. The two underlying problems are that linguistic meaning underdetermines illocutionary force, and that—like perlocutionary effect—illocutionary action does not seem to be necessary for communication or the use of language; these problems are intractable for a theory that seeks to ground linguistic meaning in conventions of illocutionary action.[10]

The first problem is that it seems as though there could be linguistic communities who speak non-literally more often than not; they primarily use sarcasm, metaphor, etc. According to the picture that grounds linguistic meaning in conventions of illocutionary action, however, such a community would not speak the language *L* that we’d naturally take them to be speaking; there would not be a convention among this population of using *s* to mean *L(s)*. Lewis responds by suggesting that the language we’d initially take to be the language of this population (call it *literal-L*) is only a simplified approximation to their actual language *L*, which may be obtained by specifying certain systematic departures from *literal-L*. (A simplified example: if *s* means *p* in *literal-L*, and is systematically used ironically by this population, then *s* means *not p* in *L* (Lewis 1975: 183)). The language of this community is determined by conventions of illocution, but they just happen to be speaking a different language than they initially appear to be.

But this response won’t do: the most serious problem is that Lewis’s solution assumes that the non-literal speech of this community will display enough systematicity to be able to derive *L* from *Literal-L*. However, this is at best extremely unlikely; it is certainly not necessary. Non-literal speech comes in many forms, and some of them—like metaphor—tend to be highly unsystematic; while we may be able to characterize the general features of different kinds of non-literal speech, there is no principled mapping from the meaning of a sentence *s* in Literal-English, for instance, to non-literal-English(*s*). The other issue is that the account rules out the possibility that for instance, Americans would share a language with the British in the case that the British started using irony to a sufficiently high degree—this seems like the wrong result.

[10] The objections can be modified to fit any theory that grounds linguistic meaning in illocutionary action, regardless of which/how many illocutionary acts are utilized in that theory. Furthermore, the objections will apply regardless of differences in the details of how these illocutionary acts are defined, as long as they meet the criteria outlined above in that they involve the speaker having a communicative intention over and above that of simply transmitting a content.

The second problem is that it seems that there could be linguistic communities who were not in the business of performing illocutionary actions at all; rather they may use a language L uniquely for the purpose of putting on plays, or telling jokes or stories.[11] If a population's language is determined by their conventions of illocution, this population does not count as speaking a language at all. Lewis' initial response here is analogous that given above; the language of this population L may be obtained by specifying their systematic departures from *literal-L* (which we would naturally identify as their language). In this case, however, Lewis recognizes that there may not be the requisite systematicity; he suggests an alternative solution, which is to restrict the account of the convention which governs language use to serious communication situations, which he defines as follows:

> We may say that a *serious communication situation* exists with respect to a sentence s of L whenever it is true, and common knowledge between a speaker and a hearer, that (a) the speaker does, and the hearer does not, know whether s is true in L; (b) the hearer wants to know; (c) the speaker wants the hearer to know; and (d) neither the speaker or the hearer has other (comparably strong) desires as to whether or not the speaker utters s. (Lewis 1975: 184)

The account would then be modified so that what determines which language a population speaks will be conventions of illocutionary action in *serious communication situations*. But again, this rules out plausible cases of shared language: we can imagine that a language was once used in a community to assert what were then taken to be religious truths, but which was later used to convey the *very same* stories, but this time as myths or parables. This situation does not reflect a change in language, but rather a change in religious belief. Perhaps more importantly, however, this restriction only serves to make salient the fact that serious communication situations are not the only contexts in which we use language. Information exchange is only one of many

[11] When joking around or telling stories, the speaker means something by her utterances—she intends her audience to attend to certain thoughts. Granted she is not *merely* intending to transmit these thoughts to the hearer; she will most likely have further intentions, perhaps including the intention to amuse her audience. However, this kind of further intention does not constitute an illocutionary act, since it is not a reflexive communicative intention. The speaker may intend that her audience be amused by the thought she has transmitted, but the basis for that amusement won't involve a recognition of the speaker's mental states (the fact that the hearer has these mental states will not typically be amusing); the source of amusement will be the transmitted thought itself. There need not be any coordination between the speaker and audience after the locutionary act has been identified. Grice touches on this point in *Utterer's Meaning and Intentions*, pointing out that while the audience's amusement may be partially *caused* by the recognition of the speaker's intentions (though I doubt that even this much is usually the case) it will not *be on the basis* of this recognition: "But though A's thought that U intended him to be amused might be a part-cause of his being amused, it could not be a part of his reason for being amused (one does not, indeed, have reasons for being amused)..." (Grice 1989: 92–3).

activities which we use language to facilitate. Our use of language is complicated and messy; we joke, we play, we say things we mean as well as things we don't mean—we often do these things unsystematically, and in the course of a single conversation. It is one thing to idealize away from complexities when constructing a model of language and language use, it is another thing to ignore or accept patently incorrect results for large swaths of perfectly ordinary use. Non-literal and "non-serious" speech are not peripheral cases; they are the norm, which any viable meta-semantic theory will need to account for.[12]

In the introduction I noted that metasemantic considerations will have significant implications concerning the viability of individual semantic theories. For instance, underlying any semantic theory which posits context change potential as semantic value seems to be the tacit assumption that meaning is in some way determined by illocutionary action or perlocutionary effect. This is because an update in the context, which can be modeled as the information state of the conversational participants, results from the recognition of the illocutionary action—and in some cases the subsequent production of the perlocutionary effect—in serious communication situations. For instance, assertion is standardly taken to effect a change in context when its intended perlocutionary effect (belief or acceptance of the asserted content by conversational participants) is achieved. In the case of questions and commands, the update in the context is not generally taken to require something quite strong as the production of the intended perlocutionary effect (an answer to the question or the performance the commanded action). However, it does involve recognition of the illocutionary action and a somewhat weaker response—for instance, the relevant conversationalists taking on the goal of answering the question or performing the commanded action. Given the underlying assumption that use determines meaning, semantic theories that posit context change potential as semantic values tacitly endorse the additional assumption that the *kind* of use that grounds linguistic meaning involves illocutionary action/perlocutionary effect. However, if the arguments above are correct and neither illocutionary action nor perlocutionary effect are necessary for language use, then it is hard to see what kind of metasemantics could support such a semantic theory.

2.3. *Coordination in identifying locutionary act*

So far we have considered two different coordination problems that language can be used to solve, one nested inside the other. The first was that of coordinating illocutionary action with perlocutionary response.

[12] In order to get around these problems one might try to incorporate mock illocutionary acts somehow (perhaps with a disjunctive account according to which meaning is grounded in conventions of illocutions or mock illocutions). I find little to recommend this approach. It is shamelessly *ad hoc*, and it generates the problem of giving a rigorous account of what mock illocutions amount to, exactly.

We saw that in order to reach the equilibrium state with this particular coordination problem, agents must first solve a different coordination problem: that of identifying the illocutionary act. But we can peel back yet another layer to find additional coordination problems nested within: that of identifying speaker meaning—and, where there is more than one level of meaning—identifying the speaker's *direct* meaning, which I will call the 'locutionary act'. I will argue that this is the coordination problem underlying language use. Though I am offering a novel metasemantic proposal, it is compatible (to varying degrees) with views about language and language use going back to philosophers like Locke, Austin and Strawson. Insofar as Locke considered language to signify mere contents of thought (devoid of illocutionary force), his view would naturally fit together with a metasemantic picture according to which language use is grounded in locutionary action, rather than illocutionary acts or perlocutionary effects. Austin thought that illocutionary acts were conventional, but that they were underdetermined by meaning; this seems to imply that the conventions of illocutionary action must be over and above those that determine meaning, at least in some cases. Strawson echoes this sentiment, agreeing with Austin that "we must refer. . . to linguistic conventions to determine what locutionary act has been performed in the making of an utterance, to determine what the meaning of the utterance is" but denying that illocutionary acts are always conventional (Strawson 1964: 442).

Given that it is possible for there to be language communities that employ language to different ends, it is important to identify what is common to all of these communities; I want to suggest that this is the act of speaker-meaning, which—to remind the reader—I define as follows:

> *Meaning*. By uttering *e*, U meant *m* iff for some audience A, U uttered *e* R-intending that A attend to *m* at least partially on the basis of her utterance.

The expression 'reflexive intention' is borrowed from Bach and Harnish (1979), and refers to intentions that are self-referential in the following way: a speaker has a reflexive communicative intention just in case part of her intention is that the audience recognize the full contents of that very intention. The speaker intends for her communicative intention to be manifest to the audience, the recognition of which is meant to provide the audience a reason for identifying the speaker's meaning one way rather than another. This definition of speaker meaning is adapted from Grice's account, but it differs crucially from his (as well as from others working within a broadly Gricean framework) in that it is less restrictive; a speaker means something merely by having the intention that her audience attend to a certain thought. The audience need not be intended to have any additional response, or to form any attitude toward the content over and above that of having it in mind.[13]

[13] Stephen Neale suggests a similar weakening in Neale (1992) p. 34–37. However, because he does not accept the implications that I take to follow from it—

If the speaker manages to get her audience to recognize her intentions and attend to the content that she has in mind, then this most basic of communicative acts—that of speaker meaning—has been successful. In contrast, more traditionally Gricean accounts of speaker meaning essentially involve some sort of additional communicative intention over and above what I am calling 'speaker meaning'—the intention that the audience do something with the content of the utterance after she has attended to it (such as to believe it, provide an answer to it, etc.) These additional communicative intentions constitute the illocutionary force of the utterance.

Of course, given that we have an interest in passing on information and coordinating action, speakers will often have communicative intentions that go above and beyond that of speaker-meaning, and instead constitute full illocutionary action. For instance, a speaker may utter a sentence not only with the intention that her audience *attend* to the proposition expressed by that sentence—but with the additional intention that the hearer recognize that the speaker believes the proposition, thereby coming to have reason to believe it herself. However, the performance and success of the act of speaker meaning are preconditions for the performance and success of an illocutionary act; so given that the agents have a shared interest in identifying the illocutionary act, they will have a shared interest in identifying the speaker's meaning in this more liberal sense. Here again we have a coordination problem nested within a coordination problem; in order to coordinate on solving the problem of identifying the illocutionary act, agents must coordinate on solving the problem of identifying what the speaker meant.

If we take the relevant coordination problem to be that of identifying speaker meaning, we get the desired results with respect to communities that do things with language other than perform illocutions. Paradigm cases in which a speaker means something by an utterance without performing an illocutionary act are conversational contexts in which the speaker is joking around or telling stories. In these contexts the speaker is not *merely* intending to transmit these thoughts to the hearer; she will most likely have further intentions, perhaps including the intention to amuse her audience. However, these further intentions do not constitute illocutionary action, since the source of amusement is typically not the recognition of the speaker's mental states, but rather the audience's impulsive reaction to her understanding of the thought being communicated. So while in this kind of situation there is no coordination needed to identify an illocutionary action, there is still coordination needed to identify the speaker's meaning. Since in non-serious communication situations speakers still mean what they say in the liberal sense that I've defined, by taking the language of a community to

for instance, that the literal content of indirect speech will be *meant* by the speaker this weakened sense—I assume he has in mind a more restricted account than the one I am proposing here.

be determined by conventions of speaker meaning, we will get the right results with respect to communities that use language solely in non-serious communication situations.

However, this does not provide the tools to deal with the other counterexample, which essentially involved indirection. In the example we considered, the indirection was with respect to the *illocutionary* acts being performed; that is, members of the hypothetical community used non-literal speech more often than not (in that, roughly, the content of their illocution tended to be non-identical to the semantic content of their utterance). But the example could easily be modified to create a problem for an account which takes the relevant language convention to be that of speaker-meaning. Indirection and non-seriousness are orthogonal issues; we could just as well consider a community of people who only joke around and tell stories, yet do all of this using indirect speech. In this kind of case, the speaker-meaning account will get the wrong results in regard to which language is the language of this population.

It is necessary to impose constraints on the account in order that the contents of implicature, ironical and metaphorical utterances and the like, do not get swept up by the theory—this can be done by restricting the account to *direct* meaning. Not all acts of speaker-meaning are on a par. With any given utterance, a speaker may perform multiple acts of speaker meaning; however she may use varying strategies to facilitate the hearer's recognition of the different things that she means. For a subset of the contents meant by the speaker, their communication will be linguistically mediated in a more direct sense than other contents that she meant in the course of that same utterance. That is, in the case that a speaker means multiple things with an utterance, there will be one content that she intends the speaker to entertain primarily on the basis of her utterance (together with shared background information), while there will be other contents that the hearer is intended to arrive at on the basis of other contents she recognizes to have been meant by the speaker. We can think of this subset as what is directly meant by the speaker. In the case that the speaker is part of a community with established linguistic conventions, what she directly means by her utterance will likely correspond with the content of what people have tended to think of as the locutionary act she performs with that utterance, in the sense that it will correspond to the conventional meaning of the expression she utters.[14, 15] In contrast, additional contents she

[14] This need not be the case. But the reason that it will often be the case is that usually the most efficient way for the utterer to get her audience to identify a certain content is to utter a sentence whose standing meaning is that very content.

[15] Again, for simplicity I am ignoring context sensitivity here, making the simplifying assumption that the meanings of sentences are complete thoughts rather than functions. In any case, the content of the locutionary act will generally bear some close relation to the conventional meaning of the utterance, though it may not be identity.

might have meant (for instance, those that she implicated with the utterance) neither correspond to the conventional meaning of the expression she used, nor were they directly meant. Rather, the hearer is intended to arrive at these additional meanings at least partially on the basis of other things meant by the speaker in that utterance. I will define locutionary action as direct meaning.

> *Indirect Meaning*: By uttering *e*, U indirectly meant *m* iff for some audience A, there is some content *k* (not identical to *m*) such that
>
> 1. By uttering *e*, U meant *k*
> 2. By uttering *e*, U meant *m*
> 3. U uttered *e* R-intending that A recognize (2) at least partially on the basis of (1).
>
> *Direct Meaning:* Whatever is meant: if it is not meant indirectly, is it meant directly.

My definition of locutionary action, then, differs from the traditional Austinian characterization in that the latter but not the former requires that there be some conventional meaning associated with speakers' utterance.

Again we find a coordination problem within a coordination problem: in the case of indirect speech, in order to identify the indirectly meant content *q*, the hearer must first identify that the speaker meant *p*, since it is on the basis of this fact that she is meant to identify the indirect content *q*. There can be many layers of meaning, but in each case the hearer must identify the content that is on the bottom-most level—the locuted content. Identification of the locutionary act, then, will be a part of any coordination problem of identifying speaker meaning. In the case that there is only one level of meaning, it will be the very same problem. In the case of identifying indirect speaker meaning, it will be a coordination problem within a coordination problem.

If we take the relevant coordination problem to be that of identifying the locutionary act of the speaker we get the right results not only with regard to the community that uses language only in non-serious communication situations, but also in regard to the community that uses indirect speech more often than not. But it is not simply a matter of making the right predicitions; the fact that this account manages to get the desired results in cases where the more traditional accounts failed is explanatory in that it is reflective of two broader facts about language use: (1) that locutionary action is common to all communication situations, and (2) that the locutionary act of the speaker is part of the mechanism by which indirect speech occurs—even when it not the "point" of the utterance, it is crucially involved in the communicative act. By grounding language use in locutionary action we peel back superfluous layers of use—following the lead of Austin and Grice, but continuing where they had stopped too soon—to find what is both common to all contexts of linguistic communication and robust enough to do the work of grounding linguistic meaning: locutionary action.

2.4. *Why not conventions of illocution as well as locution?*

As we have seen, identifying the speaker's illocutionary act involves coordination on the part of the speaker and her audience. Conversational participants use language to solve this coordination problem, in that the speaker chooses an utterance that she expects will provide the audience a way to identify her illocutionary act, and the audience will in turn identify the illocutionary act based on her expectations of the speaker. Surely, then, there can be conventions of illocutionary action; there is in principle no reason why a regularity could not develop within a community of using *s* to illocute L(*s*) and of taking others to do so. So why not take this convention—*in addition to* conventions of locutionary action—as relevant to metasemantic theorizing? We could get around the counterexample involving non-serious speech situations by introducing a disjunction into the definition: instead of the account I've offered—in which a population speaks a language *L* just in case they have a convention of using *s* to locute *L*(*s*)—we could say instead that a population speaks a language *L* just in case they have a convention of using *s* to locute *L*(*s*) and/or using *s* to (directly) illocute L(*s*). The counterexample involving indirection could be handled by modifying the account to impose a directness condition on the illocutionary action involved. While on the face of it, it might look like this picture would get the right results, this strategy should be resisted for the following reasons:

First, what looks like a convention may in fact just be correlation. According to an old picture of sentential force, different kinds of sentence types all have the same kind of content; the difference in perceived meaning does not amount to a difference in content, but rather a difference in force. According to this picture it would be somewhat arbitrary which kind of force could be paired with a particular content on an occasion of utterance, given that there is only one *kind* of content (traditionally supposed to be propositional) which is compatible with each type of force. On any occasion of utterance, there would be alternatives; the speaker can choose the content and the corresponding force based on her interests and goals on a given occasion. However, this traditional Fregean view—according to which expressions with different sentential force share the same type of content—has been shown to be untenable. Because these expressions can embed, the strategy of taking them to have a uniform kind of content leads to difficulties in providing a systematic compositional semantics, and has been largely abandoned as a result (Groenendijk and Stokhof 1997: 1061–1075). It has been more promising to posit different kinds of contents corresponding to different kinds of sentences. If this strategy is correct, and indicatives, queries, commands, etc. have different types of entities as their semantic contents, then the correlation we see between illocutionary action and sentence type might just be piggybacking on the convention of locution (which concerns merely content). For instance, we have conventions of locutions for indicative sentences that determine their contents to be propositions (or propositional functions); while indicatives may also correlate with the illocutionary act of assertion, it is not because there is a separate convention involved. Rather it is because of the nature of assertion; one can only assert contents that are propositions (and not, for instance, questions). *Mutatis mutandis* for queries with questions (semantic content of interrogatives), as well as requests with whatever kind of entity turns out to be the semantic content of imperatives. If my metasemantic proposal is correct, then in cases of direct illocutionary action we would expect to see the above correspondence. However, this correspondence would not indicate a convention because there would be no alternative; when the direct meaning is propositional, only assertion will be an option for the direct illocutionary act, when the direct meaning is a question, only a query will be an option for the direct illocutionary act, and so forth.

But let's suppose for the sake of argument that there could be conventions of direct illocutionary action. Even in this case I think that there are reasons—gleaned from the consideration of the counterexamples in *2.2*—to exclude these conventions from an account of what it is for a population to use a language. The first reason is that not all language communities need conventions of illocutionary action, though they all need conventions of locutionary action. Ideally, we'd want a simple theory that would generalize to all language speaking communities and all languages. The second is that there is an intuitive sense that two communities that share conventions of locution but not of illocution may share the same language. This seems to be evident from the hypothetical situation described above, in which Americans for the most part speak literally while the British primarily use non-literal speech. What they have in common is a convention of locution; they both use the same utterances to directly mean the same contents. They differ in their conventions of illocutionary force; whereas Americans directly assert (for instance) the locuted content, the British assert a different, indirectly communicated content. Yet there is the intuition in this case that even so, the two communities are speaking the same language. The same could be said if we were to compare a community that used language in serious speech situations in order to perform illocutionary actions, and one that did not—for instance, it seems as though *the very same language* that was used by one community to assert what they took to be religious truths (i.e., to perform illocutions), could later be used by another to convey the very same stories, but this time as myths or parables (i.e., to do something other than perform illocutions). These communities, again, would share the same conventions of locution but not the same conventions of illocution—and yet there is the intuition that they could share a language.

Acknowledgements

Many thanks to Zoltán Szabó, Jason Stanley, Liz Camp, Larry Horn, and the participants of the *Philosophy of Linguistics and Language Conference* (*Dubrovnik, 2015)* for helpful comments and feedback.

References

Austin, J. L. 1975. *How to Do Things with Words*. Oxford: Clarendon Press.

Camp, E. (forthcoming). "Insinuation, Indirection, and the Conversational Score." In D. Harris and M. Moss (eds.). *New Work in Speech Acts*. Oxford: Oxford University Press.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Groenendijk, J. and Stokhof, M. 1997. "Questions." In J. van Benthem and A. ter Meulen (eds.). *Handbook of Logic and Language*. Amsterdam: Elsevier: 1055–1124.

Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge: Harvard University Press.

Lewis, D. 1975. "Languages and language." In Keith Gunderson (ed.). *Minnesota Studies in the Philosophy of Science*, vol. VII. Minneapolis: University of Minnesota Press: 3–35.

Loar, B. 1981. *Mind and Meaning*. Cambridge: Cambridge University Press.

McKinney, R. Manuscript. Discursive Antagonism and the Value of Conversation.

Pinker, S., Nowak, M., and Lee, J. 2008. "The Logic of Indirect Speech." *PNAS* 105 (3): 833–838.

Schiffer, S. R. 1972. *Meaning*. Oxford: Clarendon Press.

Strawson, P. F. 1964. "Intention and convention in speech acts." *Philosophical Review* 73 (4): 439–460.

Quine, W.V. 2013. *In Conversation with W.V. Quine – The Boolos Panel.*
https://www.youtube.com/watch?v=2NmVVu7gF9c