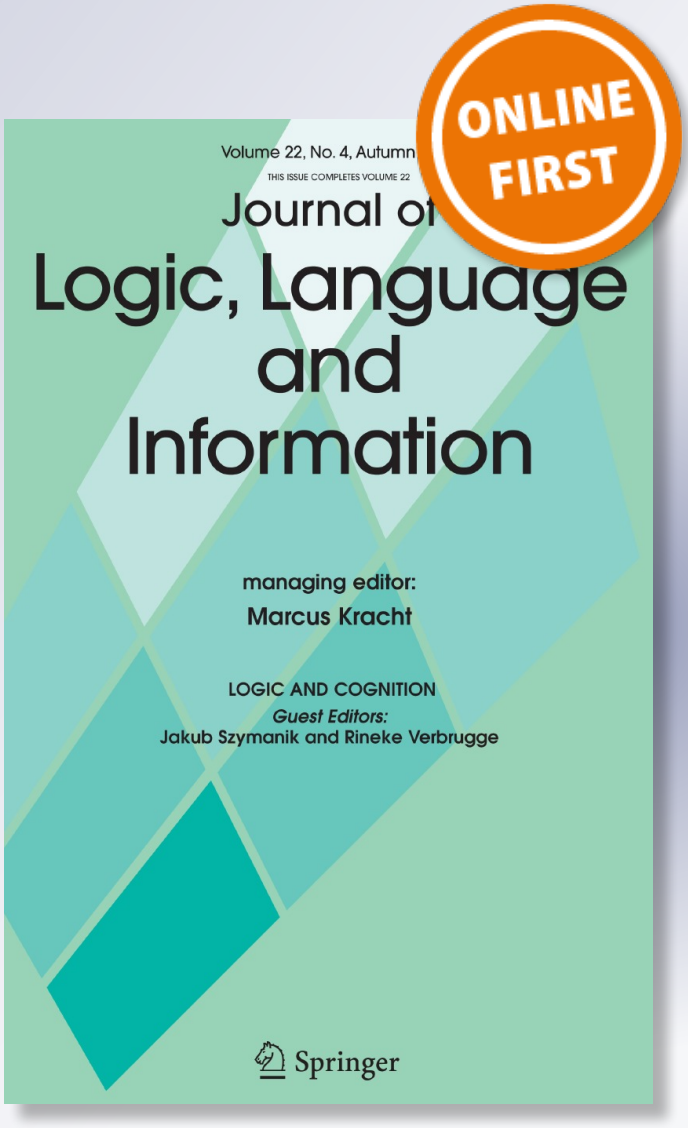
# *Natural Language and Logic of Agency*

## Johan van Benthem

**Journal of Logic, Language and Information**

ISSN 0925-8531

J of Log Lang and Inf
DOI 10.1007/s10849-014-9188-x




Springer

J Log Lang Inf
DOI 10.1007/s10849-014-9188-x

# Natural Language and Logic of Agency

**Johan van Benthem**



**Abstract** This light piece reflects on analogies between two often disjoint streams of research: the logical semantics and pragmatics of natural language and dynamic logics of general information-driven agency. The two areas show significant overlap in themes and tools, and yet, the focus seems subtly different in each, defying a simple comparison. We discuss some unusual questions that emerge when the two are put side by side, without any pretense at covering the whole literature or at reaching definitive conclusions.

**Keywords** Natural language · General agency · Information · Evaluation · Belief revision · Dynamic–epistemic logic

## 1 Semantics and Pragmatics of Natural Language

To reach our main theme, we need to set a scene with two actors. We start with logical studies of natural language. Inspired by the methods of modern logic, the classical semantics of natural language focuses on truth conditions. These describe recursively, for any sentence $\varphi$, its content: what the world must be like in order for $\varphi$ to be true. This is a *descriptive* focus on language as describing the world through combinations of atomic and complex facts. This view is close to ontology,

I thank the editor and two referees for their helpful critical comments. I have also profited from feedback by colleagues at presentations of these thoughts in Beijing and Tilburg.

J. van Benthem (✉)
Institute for Logic, Language and Computation ILLC, Amsterdam, The Netherlands
e-mail: johan.vanbenthem@uva.nl

J. van Benthem
Department of Philosophy, Stanford, CA, USA

Published online: 02 March 2014

and fits with logical consequence as an eternal relationship between propositions stating properties of worlds. The descriptive truth-conditional perspective has proven highly fruitful, in the wake of classics such as Montague (1976) and other pioneers in his generation. Notice that no human agents are involved here. Language could truthfully reflect an empty world, like a signpost in a desert. Still, language is also a natural social phenomenon that emerged for a reason. And that reason is surely the behavior of human agents who not only describe the world, but interact with it, and with one another, in a wide variety of ways. With this *functional* focus, the primary issues in understanding what makes a language tick are what agents do with it, in particular, communication. This second focus moves closer to an epistemological rather than an ontological view of language—though, of course, the two stances are not incompatible.[1] In logical studies of natural language, this second perspective goes back to classics in the pragmatics of language such as Lewis (1969, 1979), Stalnaker (1970, 2002)—and many others.[2,3]

Descriptive and functional perspectives have interacted continuously in the study of language, resulting in a fluid borderline between logical semantics and pragmatics. In particular, innovative ideas on information and agency in language occur in discourse representation theory (Kamp 1981), situation semantics (Barwise and Perry 1983), dynamic semantics (Groenendijk and Stokhof 1991; Veltman 1996), inquisitive semantics (Ciardelli et al. 2012), and varieties of game-theoretic semantics (Hintikka and Sandu 1997; Jäger and van Rooij 2007; Gärdenfors and Warglien 2006—with surveys in Clark 2012; van Benthem 2008). Truth conditions for linguistic expressions now come to emphasize how agents dynamically represent and modify information, singly, or interactively. Here, truth about the world does not drop out altogether, but it becomes one dimension in a richer view of interpretation and meaning. Much current research on language shows this richer semantic-pragmatic mixture. Two examples, out of a thriving literature, where semantics meets philosophical logic are epistemic modals (Yalcin 2007), and vagueness (Cobreros et al. 2012).

What is the overall agenda in contemporary semantics and pragmatics? There is a lively discussion of this issue in the literature, which does not seem to have reached a consensus yet (cf. Stojanovich 2008; Stokhof 2007 for interesting views). For our present purposes, we just note the following. Formal semantics of natural language started with a basic repertoire borrowed from philosophical logic. Montague gave truth conditions for the logical quantifiers, plus a few modalities for time and necessity, and the rest of the system was general categories whose expressions were left unanalyzed, though a few verbs came with meaning postulates constraining their interpretation. There was no discussion to which extent this particular focus was characteristic of

[1] There have been related major debates on where to place the focus in understanding natural language, witness the famous Chomsky–Piaget clash recorded in Piattelli-Palmarini (1980).

[2] Here and throughout this discussion piece, we will only give a few non-exhaustive references.

[3] Dag Westerståhl has emphasized that one can view the functional stance as descriptive, too, but then of linguistic practices—with the logical systems discussed later describing valid laws of such practices. This is right, though I do not think this undermines the intuitive contrast.

how language actually functions.[4] Over time, however, research in semantics and pragmatics has gone for major 'systems' in natural language such as anaphora creating coherence in discourse, temporal perspective making narrative flow, the mechanics of categorial combination, natural subsystems of reasoning in language use, comparatives and other ordering devices, questions, imperatives and other speech acts, or the rich varieties of count and mass quantification. I am in no position to paint a credible overall picture of language and its users emerging from all of these strands, but this general perspective is the appropriate level for the discussion to follow. I will explore whether dynamic logics of agency have something to say here—so let us turn now to such logics.

## 2 Dynamic Logics of Agency

Here is a second tradition, coming with a different set of motivations. In recent decades, information-driven agency has become a major topic of study by logicians working at interfaces of philosophy, artificial intelligence, computer science, game theory, and the social sciences. This area is even more vast than the semantics and pragmatics of natural language, and I cannot begin to summarize it here.[5] Let me just mention one motivation that will be the basis for my discussion later on. The above descriptive-functional interplay also occurs inside logic itself. In addition to truth-conditional world description, historically, there has always been a 'functional' countercurrent with a focus on people arguing, drawing conclusions, and exchanging information.[6] My own work on 'logical dynamics' (van Benthem 1996, 2011, 2014) has been driven by the view that logic is about all information-driven activity, involving inferences, but on a par with these, observations or questions to others. Systems of 'dynamic–epistemic logic' satisfying traditional standards of mathematical rigor then help us see the structure of this richer practice, as well as its fundamental laws. Without going into details about such technical systems, what is important to my intended comparison is the *agenda* of this sort of research, whose picture of agency I will now explain.

### 2.1 Basic Events and Attitudes on a Par

The starting point for our analysis of information-driven agency is dealing with two things on a par: the *actions* or events that produce information, and the *attitudes* that agents have, triggered and modified by these events. One cannot fully understand the one without understanding the other. Philosophical logic has tended to emphasize attitudes, such as knowledge or belief, but now we also want to bring the events that modify knowledge or beliefs into the scope of the logic, as first-class citizens in the syntax and semantics that obey their own logical laws.

[4] However, an explicit discussion of the basic mechanics of natural language is found in Barwise and Perry (1983), as a prelude to their proposed paradigm shift to situation semantics.

[5] See Fagin et al. (1995), Shoham and Leyton-Brown (2008), Wooldridge (2001).

[6] One can often see both motivations living together even in basic textbooks on logic.

### 2.2 Pilot System: Public Announcement Logic

That this can be done is shown by a simple system that 'dynamifies' epistemic logic with explicit actions of updating information. Our base language is a standard one with static modalities $K_i\varphi$ saying that agent $i$ knows that $\varphi$ is the case. Now knowledge can be gained or lost through informational events. The simplest such event is a truthful public announcement $!\varphi$[7] which updates a current epistemic model $\boldsymbol{M}, s$ with actual world $s$ to the sub-model $\boldsymbol{M}|\varphi, s$ consisting of only those worlds where $\varphi$ is true. This intuitive model transformation occurs in many areas, from philosophy to the sciences. Technically, it can be described by modalities $[!\varphi]\psi$ saying that $\psi$ holds after the update with $\varphi$ has taken place. The special case $[!\varphi]K_i\psi$ describes what an agent $i$ knows after learning that $\varphi$ is true. The resulting system of *public announcement logic* has a complete axiomatization that extends the epistemic base logic with various intuitive principles, including a fundamental law for knowledge after update with new information: $[!\varphi]K_i\psi \leftrightarrow (\varphi \rightarrow K_i[!\varphi]\psi)$. This may be considered the driving 'recursion equation' for knowledge under update with totally reliable information, and it is worthwhile pondering the interplay of dynamic update actions and knowledge states in its formulation, a quite different concern from the usual issues about epistemic closure or introspection. There is no need for us to probe public announcement logic in more technical depth here (there is more to it than meets the eye), but note how it satisfies classical standards of rigorous design.

For what follows, note three features of the methodology at work here: one identifies the major update actions that are relevant, one writes them in a suitable logical syntax, and one determines their valid principles. The resulting two-tier dynamic logics containing both formulas and action expressions tend to be conservative extensions of the static base logic one started with. That is, they do not change the valid laws of the base language, but only add principles of reasoning for the new operators. Such dynamic logical systems can be written for all scenarios to be discussed in what follows, but no such detail is needed for the points to be made in this paper.

### 2.3 Toward an Architecture of Agency

However, this modus operandi is only a beginning. We now list a number of further essential features that feed into dynamic logics of agency, all with counterparts for natural language in what follows. (a) First, rational behavior is not just driven by knowledge and perfect information, perhaps more importantly, it involves *beliefs* that can turn out to be wrong, and acts of belief revision and learning from mistakes. As is stressed in learning theory, not eternal correctness, but a talent for *correction* seems the hallmark of rational behavior.[8] (b) But crucial features of agency still remain under the radar in

[7] This common terminology may be suboptimal, since we are really modeling the update for any totally reliable information: coming from communication, observation, or yet other sources.

[8] The same logical methodology applies. One identifies key events of 'hard' and 'soft information' that transform a 'plausibility order' of worlds inside the current epistemic range. The logic makes these events explicit, and has the complete laws governing the induced belief changes. Similar logics govern changes in agents' syntactic inferential information, or their preferences.

the usual models for knowledge update or belief revision. The ubiquitous phenomenon of *acts of inference* enriches agents' information, but not in its usual semantic sense of the epistemic options. After all, by definition, a valid consequence does not change such a range. Dynamic logics of inference highlight a variety of notions of information that drive agency, from semantic models with a lot of information reduction to fine-grained symbolic or representational structures, allowing for 'awareness' or 'access' to what the world is like.[9] (c) Next, in addition to information, of whatever kind, rational behavior crucially involves agents' evaluation of states of the world or outcomes of a process, often encoded in agents' *preferences*. The balance of the two: informational attitudes and preference, is essential here. For instance, decision theory and game theory typically involve entangled notions such as choosing a best action among one's options given what one believes about their outcomes. Again, in tandem with these attitudes, we now see a richer dynamics of acts of preference change, triggered by pure evaluation changes or entangled changes in information. (d) Next comes a fourth important step. All that we described so far were single local acts and events in agency. But these normally take place inside a longer-term process of conversation or inquiry. This *temporal process structure* can be modeled, too, making it clear what an individual step 'means' in the setting of some current process known, partially or wholly, to the agents.[10] (e) A final natural move in dynamic logics of agency takes us from individuals to *social interaction*. Many intelligent tasks involve the interplay of two or more agents, with communication or games as prime examples. Just as the physical world is built from many-body interactions between particles, human agency revolves around many-mind problems. Moreover, in this setting, groups of agents themselves may acquire an independent status as new objects of their own in social behavior. Technically, this links dynamic and temporal logics of agency to the logics of interactive strategic behavior for individuals and groups found in game theory and social choice theory.[11]

Thus, a serious picture of agency has a rich architecture: it involves local triggers for changes in several entangled attitudes, as well as long-term process structure, and all of this takes place in an interactive social setting. What ties all this together is the availability of logical methods for describing major structures and their laws at these various levels. This is the picture we will have in mind when discussing natural language in what follows.

## 3 Basic Attitudes and Core Actions in Language Use

Our first question emanating from the preceding discussion concerns the basic dynamic steps of the process. What basic attitudes and actions, considered in tandem, make

[9] In this connection, while the mainstream of dynamic logics of agency is closer to standard semantics for natural language, having different levels of information might come closer to the richer structures used in Discourse Representation Theory (Kamp and Reyle 1993).

[10] This is reflected in the emergence of 'protocols' encoding procedural information (Hoshi 2009). Typical systems handling protocols are epistemic-temporal logics, or extensions thereof.

[11] For instance, a game is a typical temporal protocol for many interacting agents, whose step by step dynamics can be studied using dynamic–epistemic update methods (van Benthem 2014).

language users tick when one considers them as general agents engaged in a cognitive social practice?

### 3.1 Knowledge and Information Dynamics

It is very tempting to move from a classical descriptive perspective on truth to what language users *know* about the world, and how successive rounds of communication modify this knowledge. Indeed, this seems to be the major emphasis in current dynamic views of language use. If there is an action behind the standard truth-oriented paradigm for language, it seems to be information transfer via indicative discourse. Now it is an empirical question how much of ordinary discourse is really of this character. I myself have doubts about its centrality, since so much of daily talk is about informing people only to the extent of influencing them. Likewise, the well-known conversational postulates of Grice start from the idea that people are normally helpful sources of information. Again, it seems an empirical issue to which extent this is true, while there is also the methodological issue whether this special type of discourse is going to lead us to the heart of the functioning of language.[12] Please note that I am only raising questions here, trying to shake the tree without breaking it. But it is of interest to look at some alternative views.

### 3.2 Belief and Correction

As with agency, which is largely belief-driven, *belief* might be the more central attitude. Most language use cannot guarantee reliable information (which source would you utterly trust?), and hence our theory should not: belief based on plausibility or probability then seems more congenial to how communication actually functions.[13] What would it mean to make belief the basic ingredient of analyzing language? One very radical view might be that we need to change the semantics for any kind of expression from truth conditions to belief-oriented 'plausibility conditions'.[14] However, this radical move is just a speculation so far—and it might lead to making semantics and pragmatics probabilistic. And even then, there may be serious obstacles on this road, yielding a favored status for truth-conditional semantics after all.[15,16]

Closer to our earlier picture of agency is the issue of the basic *dynamic acts* of belief change that matter to natural language. Dynamic logics for belief revision contain

[12] It has been suggested that this emphasis is mainly a result of the philosopher's preoccupation with written rather than spoken language, an emphasis that can be very misleading. See Wenzel (2010) on the unfortunate effects of making Chinese written characters our exclusive yardstick for understanding Chinese culture, rather than also the structure of spoken Chinese.

[13] Goodman and Lassiter (2013) analyze brain processing of language as a belief revision machine.

[14] For some initial work in this direction, see Baltag et al. (2012).

[15] E.g., truth decomposes recursively along major logical constructions, but does plausibility?

[16] One power may be this: even with a much richer picture of agency, the principles governing that picture are true or false, and hence truth-conditional semantics seems appropriate to *reports on* a wide range of informational and evaluative attitudes by linguistic agents.

many such changes, more radical or more conservative, reflecting the strength with which an agent takes the new information. However, let me just pose one question here, reflecting an earlier point about cognitive abilities. Ask yourself what you find remarkable episodes that are paradigmatic for *mastery of a language*. I myself find indicative discourse one important strand, but many other speech acts come to mind, such as questions and commands. And one can go further still, in line with the above view of agency. A major focus in the linguistic literature is correct language use and successful communication, including safeguards for this to happen. But I myself am much more struck by episodes of *misunderstanding*, where we detect that something is going wrong, and repair the situation by further communication. The dynamics of correction seems more important to competent language use than absolute guarantees of correctness and mutual understanding.[17] Language can tolerate a huge amount of ambiguity, mistaken beliefs, and mutual misunderstanding precisely because we have correction mechanisms—and we still lack a good understanding of why this works better than complex systems that always guarantee correctness.[18] It would be good to develop a more systematic view of a much richer set of correction scenarios for belief, including the crucial role of disagreement in discourse as a force toward clarification and repair.

### 3.3 Evaluation and Preference

A major feature of rational agency is the lack of a natural separation between the information agents have and the ways they evaluate states of the world or outcomes of their actions. This entanglement has been there ever since the birth of probability theory in the seventieth century with its crucial notion of expected value. But the same seems true for language. It is hard to imagine any purely informational communication, since we are always broadcasting evaluative aspects, such as wanting to be in this conversation, desiring some goal, or how we feel about things in the world. It has even been claimed that pure information transfer without evaluative 'resonance' between agents cannot succeed at all. But if all that is so, what is the basic preference dynamics underlying natural language use? Again, the literature on dynamic–epistemic logic has many candidates (cf. Liu 2012), but I will make no concrete proposal here. My only point is highlighting the central importance of evaluation and its dynamics for linguistic communication, rather than as a mere side effect of conveying information.

## 4 The Language of Agency

Our discussion so far may have emphasized differences between the semantics and pragmatics of natural language as an unalterable given phenomenon, and free design of logics for agency of whatever sort. But many areas of research have features of

[17] The emphasis on correctness in language may be a remnant of old attitudes in the foundations of mathematics, with its aim of proving correctness of mathematical reasoning once and for all.

[18] Indeed, special cases of correction have been studied in semantics and pragmatics, witness a phenomenon like accommodation in presuppositions (Stalnaker 1973; Schlenker 2008).

both. After all (and this is a theme to which we will return), natural language comes with a special sort of agency, but at the same time it is a medium that accompanies any activity whatsoever. In this section, I discuss a few practical ways in which the two sides can meet, postponing a few more theoretical considerations until the end of this paper.

### 4.1 Analogies

For a start, whatever their origin, many papers can be read as contributions to both areas. For instance, current work on epistemic modals (Yalcin 2007) brings to light many subtleties in our use of expressions such as "may", "must", "certainly", "probably", that are the linguistic traces of general agent attitudes.[19] Surely, our rich linguistic repertoire for describing attitudes of ourselves and of other agents can help deepen our understanding of general agency. And also at a more technical level, there are many striking similarities in modus operandi. For instance, later on, we will briefly discuss analogies between dynamic–epistemic logic and 'dynamic semantics'.

### 4.2 Static–Dynamic Vocabulary Pairs

The logical dynamics of agency also suggests new lines in the study of natural language. Given its emphasis on how statics and dynamics belong together, we should work in tandem, and look for pairs of expressions where one describes a static attitude and the other an accompanying activity that produces and modifies the attitude. For instance, there have been many semantic studies of static verbs like "know", but few of the accompanying dynamic verb "learn". But the true functioning of natural language might be revealed only when we analyze both, and the inferences that connect them.[20] Of course, even charting the family of dynamic verbs around static "know" is not a trivial matter. One might want to include expressions like "see" (in the epistemic sense), "observe", "find out", "convince", and so on. These may seem very different in semantics, but in the perspective of logical dynamics, they seem to be about related informational processes.[21] If this is true, many current linguistic studies might need extension.[22,23] The same is true for expressions involving beliefs, where static–dynamic pairs might involve various acts that trigger belief changes, such as "suggest", "persuade", or on the negative side, "disabuse", "doubt" and so on.

[19] Holliday and Icard (2013) switch seamlessly to logic of agency mode in discussing Yalcin (2007).

[20] Even innovative logicians looking at activity verbs like "see" have suppressed its dynamics, construing it as a static relation between an agent and a situation (Barwise and Perry 1983).

[21] With epistemic modals, we would need the action expressions that make us say that something "must" or "might" be the case, with candidates such as the verbs "conclude" or "suspect".

[22] Some words in natural language might be ambiguous between a static and a dynamic reading, witness the 'product–process ambiguity' noted for many expressions in van Benthem (1996).

[23] One of the referees has pointed at work at the interface of linguistics and cognitive science that may be relevant to the issues raised here, including studies of evidentiality (Papafragou et al. 2007) and the acquisition of attitude verbs (Anand and Hacquard 2013).

### 4.3 Enriching the Logic of Agency

But natural language can also provide fresh impetus for logics of agency. The basic vocabulary for these logics is seldom informed by an analysis of how we actually formulate our thinking about attitudes, decisions, and actions. But as we saw with epistemic modals, it makes sense to assume that these encode basic ways in which we engage in these things, and one is tempted to assume that there is a general merit to the refined ways in which we can describe stronger or weaker informational attitudes, such as being certain, convinced, harboring a suspicion, and so on.[24] This intrigues me especially in the area of decision and action, as driven by entangled informational-evaluative notions. I have worked a bit on qualitative high-level descriptions of underlying mathematical theories of rational action in terms of deontic "ought" or "may" (van Benthem 2014), but it still seems to me that we are still missing the key. In particular, my hunch would be that, to understand human behavior, we want to understand the semantics of not just, and perhaps not even primarily, "know" and "believe", but of entangled qualitative terms such as "hope", "fear", and "regret".

## 5 Making the General Linguistic Process Explicit

The next key feature of agency that seems underplayed in natural language semantics (but less so in pragmatics) is the nature of the larger process taking place. To be sure, discourse representation theory and dynamic semantics are important innovations that describe representation stages, as well as transition steps that arise to form linguistic scenarios of interpretation or communication. Indeed, many linguistic theories employ a rich notion of 'context' that can pack any feature of use. But that is still a far cry from explicitly representing a process in the sense of computer science or game theory.[25]

### 5.1 Temporal Protocols

As for single steps in the linguistic process, we have seen that they can be diverse: changes in information, steps of persuasion changing beliefs, changing preferences. Now I want to draw attention to another feature of processes, their long-term temporal horizon. Conversations consist of streams of assertions, and these histories may satisfy constraints depending on the type of process. In the dynamic–epistemic literature and beyond (van Benthem et al. 2009; Fagin et al. 1995), the relevant notion is that of a *protocol* stipulating constraints such as telling the truth, answering questions eventually, not repeating oneself, and so on. Now, as pointed out by Joe Halpern, this poses a challenge to standard semantics. 'The meaning' of an assertion can often not be understood locally: one needs to know more about the protocol to see its real information content or other dimensions. It may even be impossible to have good

[24] Lenzen (1980) studies variety of basic attitudes beyond knowledge and belief in epistemology.

[25] Game-theoretic semantics does provide more process structure (Hintikka and Sandu 1997). See van Benthem (2008) for a survey of several different games proposed for natural language.

inferential intuitions about what is entailed by a given sentence without knowing the relevant protocol.[26] Still, even lacking a general account, one could mine the current literature on semantics and pragmatics for major protocols, even when these are not considered an official part of the modeling job.[27]

### 5.2 Making the Agents Explicit

But more is involved here. It is hard to think of an explicit process underlying natural language without also being explicit about the linguistic *agents* involved in that process. In most studies of natural language (and the same is true for dynamic–epistemic logics of agency), the agents themselves are not part of the formal modeling. But this is strange if we want to model realistic linguistic agents and their computational resource limitations in language processing (Icard 2013). There have been occasional proposals for doing this, for instance, in the semantic automata of van Benthem (1986) and their modern follow-up, but nothing systematic has developed in terms of the kinds of agents that perform the tasks associated with language. Even game-theoretical semantics for natural language leaves the structure of the players open—though it has to be said that modern game theory has started to take the nature of the players seriously as an essential additional parameter in specifying a game.

Thinking about the nature of linguistic agents raises interesting questions. For instance, most studies of language focus on speakers and hearers, that is, individuals interacting. But it could also be held that conversation typically involves the formation of a *group* as a third entity, and then we might want to formulate our intuitions about language use in terms of "you", "me", and "us". Another interesting issue is the undeniable diversity of agents in terms of information, attitudes, goals, and linguistic or other abilities. While theories of language tend to allow for diversity of the first kinds, they tend to assume more uniformity in terms of ability than seems warranted by the facts.[28] Allowing for diversity of competence in communication, while natural in learning and other natural scenarios, raises problems of its own. At least we also need to account for the fact that natural language is a universal medium where diverse agents, less or more competent, manage to synchronize their information and actions, smoothening out differences.[29]

[26] Similar points have been made in Halpern (2003), Bovens and Ferreira (2010) on the impossibility of solving standard puzzles like Sleeping Beauty in the philosophical literature about probability and decision. Without more process information, discussions remain fatally open-ended.

[27] A case in point is the semantics of questions in terms of decision problems in van Rooij (2003). Also relevant is the treatment of pragmatics in Jäger et al. (2005).

[28] Agent diversity in realistic communication drives complex phenomena beyond standard semantics. For a case study of the complexity arising even in the small test realm of logic puzzles, see Liu and Wang (2012) on logical scenarios where liars and truth-tellers meet and interact.

[29] Recall the point by Stalin in the early 1950s, reported in Klaus (1959), who argued that thinking of language as a class-dependent medium would make serious class struggle impossible, since all one could say would be that capitalists and proletarians are talking at cross-purposes.

Which processes underlie linguistic communication seems a moot point.[30] But I hope to have persuaded the reader that it is a natural issue that will amply repay attention.

## 6 Natural Language: Code Plus Practice

Our discussion so far has many loose ends, but a major one may be what we mean by 'natural language'. Readers may feel that I have insidiously expanded the scope of language to include many connected activities, creating disturbing open ends, and in the end, making natural language coextensive with all cognitive agency. This seems a legitimate worry, but I will only make a few observations about it here. Many people identify language with syntax, and even standard semantics may favor an approach where we take syntax as given, and then put anything we want to say about linguistic processes into the denotations used, perhaps with some 'context' thrown in. But other views are equally widespread, of language as a broader medium for communication with multi-media channels, involving also intonation, gestures, looks, and accompanying processing in non-language areas of the brain.[31] My sympathies lie with the latter view, but of course, one has to take care that my main thesis about the utility of logics of general agency for the study of natural language does not become vacuously true.[32]

### 6.1 Grammar Plus Logic

One relatively modest richer conception of language goes back to a strand that has been present in semantics since the 1970s: the idea that 'the inferences are part of the data'. If we take this seriously, this is more than just using valid or non-valid inference patterns as constraints on semantic modeling. Instead, the idea emerges that a language is not just a *grammar* generating (interpreted) expressions describing the world, but also, intertwined with this, an *inferential practice*. And then, different 'languages' may consist in different mixtures of grammar and inference engine. There are precedents for this. Kameyama (1995) develops an appealing view that languages consist of a tight mixture of linguistic assertion plus logical default inference, and she suggests that languages like English and Japanese, both functioning completely successfully, differ as to what they put in syntax, and what in reasoning routines. I find this view very congenial, and it fits well, for instance, with my program of natural logic as coming closely linked to natural grammar (cf. van Benthem 1986, and for modern work in this line, Moss 2010). Notice also that a dynamic inferential component of a language will have to use more fine-grained syntactic information, as discussed earlier on, perhaps in the form of suitably compacted discourse representations.

[30] Interesting proposals for 'the mother process of natural language' are the view of language as abduction in Hobbs et al. (1993), the game theoretical analysis of conversation in Feinberg (2008), or the neuroscience-inspired probabilistic lambda calculus of Goodman and Lassiter (2013).

[31] Cai (2013) ties this to the original semiotic program of Peirce or De Saussure for language.

[32] One might also think that the 'language' of semantics is narrower than the broader 'language' of pragmatics, but this is a subtle debate I do not want to enter into here: cf. Stojanovich (2008).

But even such a small extension of the notion of language has interesting consequences. Consider the basic notion of *translation*: this should now be a transformation from one language into another that links not just the possible assertions but also the inferences. As it happens, many translations $\tau$ in logic from a language $L_1$ into $L_2$ have such a surplus. Given the natural induced map $t$ taking $L_2$-models $\boldsymbol{M}$ to $L_1$-models, they ensure that (a) $\boldsymbol{M} \models \tau(\varphi)$ iff $t(\boldsymbol{M}) \models \varphi$, connecting assertions about given situations—but also, given mild assumptions on the model transformation map $t$, they yield the equivalence (b) $\varphi \models \psi$ in $L_1$ iff $\tau(\varphi) \models \tau(\psi)$ in $L_2$, allowing us to connect inferences as well. However, this only works with inferences defined in the usual style of truth preservation. If we define a language as a grammar for assertions plus a proof system for inferences, then the existence of translations in our extended sense of connecting richer practices becomes a much more delicate, but also perhaps practically more important, matter. They now come to stand for successful simulations between types of behavior.[33]

## 7 Methodological Issues

Our juxtaposition of research traditions also raises some interesting methodological issues with a general thrust. We merely mention one, suggested by a specific analogy in topics and techniques between studies of natural language and of general agency.

### 7.1 Dynamic Logic and Dynamic Semantics

Recall the main idea of dynamic semantics (Groenendijk and Stokhof 1991; Veltman 1996; Ciardelli et al. 2012): the meaning of linguistic expressions consists in the information change they bring about in ideal hearers who accept the truth of what is said. This sounds like a basic scenario for general information-driven agency, and what is more, technical logics of dynamic semantics often have a model-transforming flavor similar to that in dynamic–epistemic logics. How should we understand the similarities and differences between these two ways of bringing out the agency in language? van Benthem (2013) analyzes the contrast in terms of a broad distinction across logic between *explicit* and *implicit* styles of modeling, going back all the way to the contrast between intuitionistic and epistemic logic as accounts of information and knowledge. Dynamic–epistemic logics identify basic informational or other actions in relevant scenarios, introduce explicit syntax for these, and then give complete recursion laws for changes in agents' knowledge, belief, preference, or other relevant attitudes, mostly leading to conservative extensions of the original static logic. In contrast, dynamic semantics keeps the original language as is, but enriches its notion of meaning in terms of information change, issue change, plausibility change, or other parameters.

[33] On our full agency view, however, translation will have to be a much richer notion still, involving different agents to communicate successfully across their grammar + reasoning practices. Moreover, given the variety of basic informational actions, we may also want a translation to mimic update steps on models for agents using the two languages. It is not hard to see that a logical translation of the usual sort will do this, for instance, for public announcement updates, but we leave the details of this sort of extended correlation of behavior to a future occasion.

It then arrives at non-classical logics whose deviations from classical logic (including the absence of certain classical laws) are the very locus for seeing what the dynamics does—though systems may have sublanguages that still satisfy classical principles, where the dynamics is absent or muted.

This is just a sketch, and many more things can be said on this choice of methodology.[34] But even at this level of generality, how should we assess the two options, implicit and explicit? I cannot do justice to this issue in this short compass, but do want to point out an arguably deeper view of the contrast that seems to underlie much of the modern literature—and that may offer a slightly different perspective on the topic of this paper.

### 7.2 Logic 'of' and Logic' as'

In the background of the above, there is a distinction that has been emerging recently at interfaces of logic, process theory, and game theory. Let me illustrate it with the case of games (it is in fact a major thread in van Benthem 2014). On one view, suitably expressive standard logics serve to analyze the structure of games, leading to a conception of logic *of* games. But on another dual view, games themselves redefine the meaning of the original logical constants (say, as choices for conjunction and disjunction, or as role switch between players for negation), leading to a conception of logic *as* games. The same may be true for approaches to natural language. Do we think of standard expressions as embodying acts of information change or other forms of update as discussed above, or do we think of them as a medium telling us about such changes? Is our focus to be language as action, or language of action?[35]

### 7.3 Natural Language as a Universal Medium

But I may still be missing an essential point about natural language if I construe the above issue as a battle over its proper analysis in terms of taking sides in some battle of formalisms. Language seems to have features of both implicit and explicit stances in its vocabulary and its constructions, and it even provides shifts in vocabulary indicating, say, when a 'participating' implicit stance gives way to a 'commenting' explicit one. And this is just one case, where natural language has a universality that transcends different fixed positions. Other instances include the ease with which language users can shift between first-person and third-person reports of a scenario, a shift in perspective that would be viewed as a contentious major system choice in designing dynamic

[34] van Benthem (2013) discusses more detailed examples of the explicit–implicit contrast in epistemic logic and intuitionistic logic, dynamic logics of belief change and non-monotonic logics, or game logics and logic games—and draws general lines, including possibilities for borrowing, formal translation, and system merging between the two approaches. Holliday and Icard (2013) is an interesting case study relating the two approaches in the area of epistemic modals.

[35] One of the referees has pointed at areas of linguistics where similar dualities seem to play, including studies of presupposition (Beaver 1997), vagueness, and 'procedural semantics'.

logics of action.[36] This versatility may hold the key to a deeper understanding of the functioning of natural language and its amazing powers. In particular, to do justice to it by a logician's lights, the full success story of natural language may need both implicit and explicit approaches in the above sense.[37] I would love to discuss these issues further, but they are a topic for a paper all on its own.

## 8 Conclusion

I have compared the high-level agenda of semantics and pragmatics of natural language with that for dynamic logics of action. My starting point was a distinction between bare truth about the world and epistemic agency—though it should have become clear that both are needed and must proceed in tandem, since agency can only work in the world. In the process of developing my comparisons, I have suggested many questions, both conceptual and technical, that might be of mutual benefit. Semantics and pragmatics may profit from a rich general view of the linguistic process and its agents, but just as well, logics of agency might pay more attention to state of the art semantics and pragmatics of the linguistic expressions that accompany and guide human behavior.

Still, what I am suggesting may sometimes sound too reasonable, or self-evident. So let me remind the reader that approaching natural language through the lens of agency is a particular take that has competitors. One could also seek the essence of language in agent- and attitude-free patterns, as in behavioristic theories of commitment, or in statistical studies of bulk behavior without any individual agent model at all. And also, I would not wish to make semantics/pragmatics and logic of agency too similar. We might lose a creative tension between making sense of natural language as it happens to be, and designing new logical languages for conceptual reasons without empirical constraints. This tandem works well in philosophical logic, and it may also do so here.

## References

Anand, P., & Hacquard, V. (2013). Epistemics and attitudes. *Semantics and Pragmatics*, *6*, 1–59.

Baltag, A., Smets, S., & Rodenhäuser, B. (2012). Doxastic attitudes as belief-revision policies. In *Proceedings of the ESSLLI workshop on strategies for learning, belief revision and preference change*, University of Opole.

Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: The MIT Press.

Beaver, D. (1997). Presupposition. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 939–1008). Amsterdam: Elsevier.

Bovens, L., & Ferreira, J. L. (2010). Monty Hall drives a wedge between Judy Benjamin and the sleeping beauty. *Analysis*, *70*, 473–481.

Cai, S. (2013). The value of semantics and pragmatics: In the view of linguistics, psychology, and cognitive science. In J. van Benthem & F. Liu (Eds.), *Logic across the University: Foundations and applications*. London: College Publications (to appear).

Ciardelli, I., Groenendijk, J., & Roelofsen, F. (2012). Inquisitive semantics: A new notion of meaning. *Language and Linguistics Compass* (to appear).

[36] Language even refuses to be drawn into controversies between formal and natural language. It has a creative ability to absorb formal language as needed, witness the smooth absorption of technical terminology and notations in special fields like mathematics, but also more generally.

[37] Ciardelli and Roelofsen (2013) discusses such merged systems in the area of questions.

Ciardelli, I., & Roelofsen, F. (2013). *Inquisitive dynamic–epistemic logic*. Working paper, Institute for Logic, Language and Computation, University of Amsterdam.
Clark, R. (2012). *Meaningful games*. Cambridge, MA: The MIT Press.
Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2012). Tolerant, classical, strict. *Journal of Philosophical Logic*, *41*, 347–385.
Fagin, R., Halpern, J., Moses, Y., & Vardi, M. (1995). *Reasoning about knowledge*. Cambridge, MA: The MIT Press.
Feinberg, Y. (2008). Meaningful talk. In K. Apt & R. van Rooij (Eds.), *New perspectives on games and interaction* (pp. 105–119). Amsterdam: Amsterdam University Press.
Gärdenfors, P., Warglien, M., et al. (2006). Cooperation, conceptual spaces, and the evolution of semantics. In P. Vogt (Ed.), *Symbol grounding and beyond* (pp. 16–30). Heidelberg: Springer.
Goodman, N., & Lassiter, D. (2013). *Probabilistic semantics and pragmatics: Uncertainty in language and thought*. Department of Psychology, Stanford University.
Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, *14*, 39–100.
Halpern, J. (2003). *Reasoning about uncertainty*. Cambridge, MA: The MIT Press.
Hintikka, J., & Sandu, G. (1997). Game-theoretical semantics. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 340–361). Amsterdam: Elsevier.
Hobbs, J., Stickel, M., Appelt, D., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, *63*, 69–142.
Holliday, W., & Icard, Th. (2013). *Logic, probability, and epistemic modality*. Department of Philosophy, Stanford University, and Department of Philosophy, University of Californa at Berkeley.
Hoshi, T. (2009). *Epistemic dynamics and protocol information*. Ph.D. thesis, Department of Philosophy, Stanford University (ILLC-DS-2009-08).
Icard, Th. (2013). *The algorithmic mind: A study of inference in action*. Ph.D. thesis, Department of Philosophy, Stanford University.
Jäger, G., Benz, A., & van Rooij, R. (Eds.). (2005). *Game theory and pragmatics*. New York: Palgrave McMillan.
Jäger, G., & van Rooij, R. (2007). Language structure: Psychological and social constraints. *Synthese*, *159*, 99–130.
Kameyama, M. (1995). *Indefeasible semantics and defeasible pragmatics*. Artificial Intelligence Center. Menlo Park: SRI International.
Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, Th. Janssen, & M. Stokhof (Eds.), *Truth, representation and information* (pp. 277–322). Dordrecht: Foris.
Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Heidelberg: Springer.
Klaus, G. (1959). *Einführung in die formale logik*. Berlin: Deutscher Verlag der Wissenschaften.
Lewis, D. (1969). *Convention*. Cambridge MA: Harvard University Press.
Lewis, D. (1979). Score-keeping in a language game. *Journal of Philosophical Logic*, *8*, 339–359.
Lenzen, W. (1980). *Glauben Wissen und Wahrscheinlichkeit*. Library of exact philosophy. Wien: Springer.
Liu, F. (2012). *Dynamics of preference change*. Dordrecht: Springer.
Liu, F., & Wang, Y. (2012). Reasoning about agent types and the hardest logic puzzle ever. *Minds and Machines*, *23*, 123–161.
Montague, R. (1976). *Formal philosophy*. New Haven: Yale University Press.
Moss, L., et al. (2010). Natural logic and semantics. In M. Aloni (Ed.), *Logic, language and meaning*. Lecture notes in computer science (Vol. 6042, pp. 84–93). Berlin: Springer.
Papafragou, A., Li, P., Choi, Y., & Han, C. (2007). Evidentiality in language and cognition. *Cognition*, *103*, 253–299.
Piattelli-Palmarini, M. (Ed.). (1980). *Language and learning: The debate between Jean Piaget and Noam Chomsky (the Royaumont debate)*. Cambridge, MA: Harvard University Press.
Schlenker, Ph. (2008). Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics*, *34*, 157–212.
Shoham, Y., & Leyton-Brown, K. (2008). *Multiagent systems: Algorithmic, game theoretic and logical foundations*. Cambridge, UK: Cambridge University Press.
Stalnaker, R. (1970). Pragmatics. *Synthese*, *22*, 272–289.
Stalnaker, R. (1973). Presuppositions. *Journal of Philosophical Logic*, *2*, 447–457.
Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, *25*, 701–721.
Stokhof, M. (2007). Hand or hammer? On formal and natural languages in semantics. *Journal of Indian Philosophy*, *35*, 597–626.

Stojanovich, I. (2008). The semantics/pragmatics distinction. *Synthese*, *165*, 317–319.
van Benthem, J. (1986). *Essays in logical semantics*. Dordrecht: Reidel.
van Benthem, J. (1996). *Exploring logical dynamics*. Stanford: CSLI.
van Benthem, J. (2008). Games that make sense: Logic, language and multi-agent interaction. In K. Apt & R. van Rooij (Eds.), *Proceedings KNAW colloquium on games and interactive logic* (pp. 197–209). Amsterdam: Amsterdam University Press.
van Benthem, J. (2011). *Logical dynamics of information and interaction*. Cambridge: Cambridge University Press.
van Benthem, J. (2013). *Implicit and explicit stances in logic*. Report PP-2013-02, Institute for Logic, Language and Computation, University of Amsterdam.
van Benthem, J. (2014). *Logic in games*. Cambridge, MA: The MIT Press.
van Benthem, J., Gerbrandy, J., Hoshi, T., & Pacuit, E. (2009). Merging frameworks for interaction. *Journal of Philosophical Logic*, *38*, 491–526.
van Rooij, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, *26*, 727–763.
Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, *25*, 221–261.
Wenzel, Ch. (2010). How pictorial is Chinese? And does it matter? In E. Nemeth, R. Heinrich, & W. Pichler (Eds.), *Papers of the 33rd International Wittgenstein Symposium* (pp. 323–325). Kirchberg: International Wittgenstein Society.
Wooldridge, M. (2001). *An introduction to multiagent systems*. New York: Wiley.
Yalcin, S. (2007). Epistemic modals. *Mind*, *116*, 983–1026.