Many linguists consider semantic theory to be a part of cognitive science. In this respect, there is a tension between the internalist commitments inherited from Noam Chomsky and the externalist commitments inherited from David Lewis. Supposing that the theory specifies an interpretation function, \([\cdot]\), that maps morphosyntactic objects to something else, the tension resolves to the question: Should we understand the somethings to be (i) worldly objects (possible or actual individuals, sets of individuals, etc), or (ii) mental representations? The answer is often (i), likely because this assumption helps to maintain a tight connection between the data (judgments of sentence truth and falsity) and the theory—thus, semantics interfaces with (at least) metaphysics (Bach 1986, Bach and Chao 2012). Yet, recent research at the interface between semantics and cognitive psychology suggests that we (re-)consider option (ii). I explicate the foundational tension, and argue for a new perspective on the connection between formal semantics and the data for semantics under (ii).

In pursuing explanations for semantic phenomena, we often posit more entities than might seem metaphysically plausible (cf. Hobbs 1985). For example, the ring and its constituent gold can’t be considered identical referents for the ring and the gold, lest paradoxes abound (e.g. Parsons 1979). And as the number of types of entities increases, so does their richness; e.g., homomorphic relations must hold within and between at least times, pluralities, events, and degrees (e.g. Krifka 1989, Schwarzschild 2006, Nakanishi 2007, etc). Meanwhile, the specific question of how these entities and ‘the world’ are related is often met with agnosticism: we investigate natural language metaphysics, not metaphysics per se; to say more “would be unethical” (Bach and Chao 2012:192, echoing Montague 1973), even “immoral” (Bach 1986:592; cf. Moltmann 2017). Yet, agnosticism deprives us of an independent theory of \([\cdot]\)’s range (its domain is, of course, constrained by syntactic theory), and ultimately it is unclear how we can say that \([\cdot]\) relates expressions to ‘the world’.

On the other hand, the relative dearth of defendants of (ii) could be due to a wariness about what would constitute a suitably predictive alternative theory. But, advances in non-linguistic cognitive science in the past decades has been remarkable, and we now know a lot more about the cognitive systems that plausibly interact with linguistic cognition, setting the stage for new interdisciplinary collaboration (cf. Spelke’s 2003 review). Perhaps worse is the worry that, if we relate expressions to mind-internal objects, we can no longer see “how mutual understanding can ever be guaranteed or even achieved”, nor “how any truth-conditional account [of meaning] could be involved” in semantics (Pelletier 2011, p33). But, positing a ‘coordination’ problem in semantics duplicates that already present for percepts and concepts, independently of language. Cognitive scientists tend to think that our perception of a common experience is due to a construction of experience that is species-level (in addition to species-specific; cf. Jackendoff 1994, Hoffman 2009).

But how can a theory accepting (ii) build on, rather than reject, the successes of truth-conditional semantics? This question is addressed in recent work at the language-cognition interface. For example, Lidz et al (2011) offer an ‘Interface Transparency Thesis’, whereby truth-conditionally equivalent formalisms of a meaning \(\alpha\) specify hypothetical differences between the mental representations and operations invoked in understanding \(\alpha—[\alpha]\) is a function-in-intension in Church’s (1941) sense. In my case study, I discuss recent results linking the formal property of ‘atomicity’ in language with categorization behavior for static and dynamic entities (e.g., Wellwood, Hespos, & Rips 2017). Assuming ITT as a linking hypothesis, I show how a cognitive interpretation of the atomicity requirement on predicates strongly predicts naıve participants’ categorizations, as well as the correspondences between their categories and varieties of syntax that impose the requirement. I discuss how the results of such experiments are mysterious on a view like (i), but follow from (ii).
References


Schwarzschild 2006. The role of dimensions in the syntax of noun phrases. Syntax, 9(1).