

### What an *average* semantics needs

This paper uses the semantics of sentences that describe numerical averages, such as (1), to argue for two general conclusions about the syntax-semantics interface.

- (1) Another study, by British researchers, stated that men had 12.7 heterosexual partners in their lifetimes *on average* and women had 6.5. (*New York Times*, 8.12.07)

First, numerals must be able to take scope independently of the nouns they modify. Second, it is not always the case that the function created by abstracting over the base position of a scope-bearing element composes with the scopal term (as in Heim and Kratzer 1998); instead, this function may compose with a third element intervening between the scopal term and its base position (Barker in press, Bhatt and Takahashi 2007).

Numerical averages can be described in various ways. The first conjunct of the embedded clause in (1) can be paraphrased by the variants in (2a-b), for example.

- (2) a. A man's average number of heterosexual partners over his lifetime is 12.7.  
 b. Men have an average of 12.7 heterosexual partners in their lifetimes.

These examples can be handled fairly straightforwardly by assuming a basic meaning for *average* that is looking for three arguments: a measure function, a set, and a number. In (2a), for example, we could derive intuitively correct truth conditions by assuming that *number of heterosexual partners* denotes a relation between individuals and the number of sexual partners they have (i.e., type  $\langle d, et \rangle$ , the standard type posited for gradable predicates), that *a man* denotes a property on the basis of which a set can be built, and that *average* denotes the relation AVG in (3), where  $g_m = \lambda y. \max\{d \mid g(d)(y)\}$  and  $S_f = \{z \mid f(z) = 1\}$ .

$$(3) \quad \text{AVG} = \lambda g_{d,et} \lambda f_{et} \lambda n. \frac{\sum_{x \in S_f} g_m(x)}{|S_f|} = n$$

Composition will give us back true for (2a) just in case the sum of the number of sexual partners for each man divided by the number of men equals 12.7, which is what we want. Similar results can be obtained straightforwardly for (2b) by assuming that *an average of 12.7* is a constituent that can take scope at the VP level, and by defining a new meaning for nominal *average* based on (3) but with a different order of argument composition (first the number, then the degree relation created by movement to VP, then the property).

The adverbial construction in (1) presents more of a challenge, however. Such examples are discussed by Higginbotham (1985) and Carlson and Pelletier (2002) as providing a potential solution to the problems presented by examples like (4), which differs from (2c) in that *average* combines with a nominal that denotes a property rather than measure nominal.

- (4) The average man has 12.7 heterosexual partners in his lifetime.

Cases like (4) have been argued to present a challenge to a referential theory of meaning (Hornstein 1984, Chomsky 2000), as there is no plausible analysis of *the average man* here as referring to a unique individual. Higginbotham suggests that this problem can be overcome if (4) can be given a semantics that renders it equivalent to (1). What he fails to notice, however, is that this move fails to explain a more basic fact about these sentences: neither

entails the existence of individuals with 12.7 heterosexual partners. This is unexpected if the DP constituent *12.7 heterosexual partners* denotes a (weak or strong) generalized quantifier, since such a meaning will end up generating such entailments. While fractional sexual partners may not be logical impossibilities, they are not part of the meanings of (1) or (4); instead, both sentences have truth conditions that are equivalent (in the relevant respects) to the more transparent instances of numerical averaging in (2).

Our solution to this problem involves two primary analytical claims. First, a numeral can take scope independently of the noun phrase in which it appears, leaving a variable over amounts/degrees (type  $d$ ) in its place, an assumption that is common in work on comparatives. Second, this variable can be bound not only at the scope position of the numeral, but at any node in the LF representation, creating a function of type  $\langle d, \alpha \rangle$ . This second assumption is crucial, as it lets the *average* term intervene between its number argument (the scoped numeral) and its degree relation argument (the result of binding the base position of the numeral), allowing for a compositional interpretation based on (3).

We illustrate with (1), starting from the assumption that *on average* attaches to VP and denotes the relation in (3). We assign the number sentential scope and bind the variable it leaves behind just below the adverb, as indicated in the LF-representation in (5a); composition gives us (5b), which will give us the truth conditions we want.

- (5) a.  $[_{TP} 12.7 [_{TP} \text{men} [_{VP} \lambda d [_{VP} \text{have } d \text{ htrsxl partners in their lifetimes}]] \text{ on average}]]]$   
 b.  $\text{AVG}(\lambda d \lambda x. x \text{ has } d \text{ heterosexual partners in } x\text{'s lifetime})(\lambda z. z \text{ is a man})(12.7)$

(4) can be handled similarly by deriving a meaning from AVG that switches the order of its first two arguments (and, with Carlson and Pelletier, that *the* is vacuous here); again it is crucial that the variable left behind by the number be bound at the VP (below the subject) to ensure that the constituent that the subject combines with denotes a function of type  $\langle d, et \rangle$ . As we show in detail elsewhere, this analysis handles the challenges for a referential semantics posed by Hornstein and Chomsky; here we focus on the fact that it provides a fully compositional account of how (4) and (1) come to have meanings that are equivalent to more transparent examples of numerical averaging such as the sentences in (2).

Crucial to our analysis is the hypothesis that the function created by abstracting over the base position of a scope-taking element need not compose directly with that element, but can correspond to a constituent lower down in the representation. This hypothesis also plays a crucial role in (one version of) the analysis of the functional adjectives *same* and *different* in Barker (in press) and in the analysis of phrasal comparatives in Heim 1985 and Bhatt and Takahashi 2007; the fact that we see the same compositional operation at work in the semantics of averaging (which involves both a ‘functional’ adjective and the semantics of degree) provides further evidence that it is a grammatical option. Moreover, the adverbial construction in (1) provides new insights on how this option should be characterized. In Barker (in press) and Bhatt and Takahashi (2007), the relevant configuration is created by inserting an expression between a scopal element and the  $\lambda$ -term that represents its scope (hence Barker’s term ‘parasitic scope’). While this analysis makes sense for ‘interveners’ that have scopal properties of their own, it seems less plausible for examples like (1), where the intervener is an adverb. We therefore pursue the hypothesis that this option is freely available, and provide implementations both for a semantics with variables (free attachment of ‘binding indices’) and without (based on Jacobson’s semantics for bound pronouns).