

法政大学学術機関リポジトリ

HOSEI UNIVERSITY REPOSITORY

Weighted Underspecification of Aspectual Properties : Against Brennan & Pylkkänen(2008)

著者	ISHIKAWA Kiyoshi
出版者	法政大学文学部
journal or publication title	Bulletin of Faculty of Letters, Hosei University
volume	64
page range	11-21
year	2012-03-15
URL	http://hdl.handle.net/10114/6956

Weighted Underspecification of Aspectual Properties: Against Brennan & Pylkkänen (2008)*

ISHIKAWA, Kiyoshi

Abstract

Brennan and Pylkkänen (2008) claim to have obtained experimental evidence for aspectual coercion (with a specific direction). However, their conclusion is invalid. For one thing, they misconstrued the issue. For another, their result does not support their conclusion; rather, in the absence of a hitherto unknown new interpretation, a comparison of their result with the previous findings in the literature supports a noncoercion analysis that Brennan & Pylkkänen believe to have succeeded in dismissing.

1. Introduction

Semanticists sometimes talk about “aspectual mismatches,” that is, mismatches between the assumed aspectual properties of predicates and modifiers that do not result in unacceptability. For example, consider *jump*, as used in (1); it accepts punctual modifiers as in (1a), in which case it expresses an instantaneous event, but yet it also accepts non-punctual, durative modifiers, as in (1b), in which case it denotes a repeated series of jumping, which as a whole is a non-punctual event:

- (1) a. Susan jumped at seven o'clock.
b. Susan jumped until dawn.

The observation that the verb denotes either a punctual moment or a duration depending on the modifier supports the idea that different kinds of temporal modifiers can only attach to predicates with different aspectual properties (in this case, instantaneous and durative). However, if we accept the idea that the aspectual properties of predicates are uniquely specified in the lexicon, *jump* should lexically express either an instantaneous event or a durative event, not both, and whichever option is adopted, we would have to say that either one of (1a–b) constitutes an instance of aspectual mismatch.

* This paper is based on the ideas I came up with in regular meetings with ISHII So, SAIZEN Akira, and OBA Ryo (and also KADO Tsubasa). I would like to thank the members. The usual disclaimer applies.

However, this conclusion hinges upon the assumption that *jump* is lexically specified for its (non-)punctuality; if it is lexically ambiguous between a punctual and a durative interpretation (or the punctual/durative distinction is lexically underspecified)¹, neither of (1a–b) will constitute aspectual mismatch. Should we, then, say that *jump*, for example, lexically selects an instantaneous or a durative event interpretation, with the lexically unselected interpretation derived on-line in a sentential context by some sort of coercion operation (aspectual mismatch), or that the lexical entry for *jump* simply does not care about (non-)punctuality (underspecification)?

Brennan & Pykkänen (2008) contrast the following three as the logical possibilities:

Underspecification: Predicates such as *jump* are lexically underspecified as to the punctuality of their denotations.

Iterative Coercion: The denotations of predicates such as *jump* are lexically specified as punctual events but can be shifted on-line to iterations of such punctual events, which take time.

Punctual Coercion: The denotations of predicates such as *jump* are lexically specified as durative events, but can be shifted on-line to single, punctual events.

Brennan & Pykkänen claim to have obtained behavioral evidence supporting **Iterative Coercion**.

The purpose of this paper is to argue that they misconstrued the issue and misinterpreted the result of their own behavioral experiment².

The organization of the paper is as follows. Section 2 gives a critical illustration of the background for Brennan & Pykkänen's (2008) behavioral experiment. Section 3 summarizes their behavioral experiment in question. Section 4 points out the defect of their interpretation of the experimental results. Section 5 concludes the paper.

2. Background

2.1 On the Previous Theoretical Discussions

Iterative Coercion and **Punctual Coercion** are instantiations of the aspectual mismatch analysis, and in the absence of constraints on how the assumed coercion operation works, they predict that both instantaneous and durative interpretations should be freely available, a result that **Underspecification** predicts without the help of an additional lexical (non-)punctual specification or an additional coercion operation. Thus, unless some constraints are pointed out on the availability of either instantaneous or durative interpretations, theoretical parsimony would definitely favor **Underspecification**.

Since Brennan & Pykkänen present the three logical possibilities not only as logical possibilities but also as options that have actually been proposed in the theoretical literature, we would expect that some empirical motivations, in the form of the unavailability of either instantaneous or durative interpretations, for an aspectual mismatch analysis have been presented in the literature they cite. A look at the literature they cite, however, reveals that their review of the theoretical literature is rather surprisingly wrong.

First, Brennan & Pykkänen cite Moens & Steedman (1988) as claiming **Underspecification**. As seen above, this would be the most economical analysis, and hence, it would not be surprising if such a proposal was made. However, Moens & Steedman in fact propose not **Underspecification** but rather **Iterative Coercion**; they propose an ATN-style coercion network, according to which the kinds of durative (repetitive) interpretations in question should be derived via coercion from instantaneous interpretations. Moens & Steedman's motivations for such an analysis for the durative interpretations are (i) their intuition that predicates such as *jump* or *wink* denote instantaneous events by default, (ii) the observation that those predicates do allow progressive forms, and (iii) the assumption that the semantics of the progressive requires that non-progressive forms should denote durative events. Their intuition (i) leads them to assume that the lexicon specifies instantaneous interpretations for those predicates, and the observation (ii) and the assumption (iii) lead them to posit an instantaneous-to-durative coercion operation for those predicates.

Putting aside whether the assumption (iii) is justified or not³, Moens & Steedman's treatment would be justified if the durative interpretations are available only when required by something, such as the progressive form. In fact, in the theoretical literature, **Iterative Coercion** seems most widely assumed⁴, probably based partially on theoreticians' intuition and partially on the assumption that such predicates denote repetitions (durative events) only when required by, say, durative modifiers, an assumption which may or may not be correct. (See the next paragraph.)

Second, Brennan & Pykkänen attribute a proposal of **Punctual Coercion** to Rothstein (2004). Conceptually, **Punctual Coercion** is rather unnatural, given that *jump* refers to a repetition of jumping in its durative interpretation, each member of the repetition being also expressed by *jump* under its punctual interpretation; the durative interpretation is conceptually more complicated by the punctual interpretation. Thus it would be very interesting if somebody proposed **Punctual Coercion** on theoretical grounds. However, Rothstein's treatment of the predicates in question is in fact an instance of **Underspecification**.

Rothstein separates those predicates from achievement predicates and calls them "semelfactives." Rothstein adopts the assumption (iii) for the semantics of the progressive, and hence, for her, any predicate that allows progressive forms should be able to denote a durative event. In the light of the observation that (some) achievement predicates accept progressive forms, she proposes **Iterative Coercion** for achievement predicates. In other words, durative interpretations are posited for achievement predicates only in order to accommodate, under the assumption (iii), the fact that they allow progressive forms; they do not seem to have durative interpretations in non-progressive forms, and hence, achievement predicates are assumed to denote punctual events lexically. In contrast, according to Rothstein, semelfactives allow both instantaneous and durative interpretations in non-progressive forms (in contrast to the assumption mentioned in the above paragraph). If so, semelfactive predicates should accommodate both interpretations lexically, and hence, naturally, she assumes **Underspecification** for semelfactive predicates⁵.

It is clear from the design of the experiments that Brennan & and Pykkänen targeted those predicates which can denote both single instantaneous events and durative repetitions of such

events, and they are exactly what Rothstein calls semelfactive predicates. Since Rothstein (2004) is the only literature that Brennan & Pykkänen cite as a claim of **Punctual Coercion**, we have to say that **Punctual Coercion** was proposed by nobody, as far as Brennan & Pykkänen's review of the theoretical literature is concerned.

2.2 On the Previous Experimental Results

Note that **Underspecification** itself only claims that (semelfactive) predicates lexically allow both instantaneous and durative interpretations and hence *is* compatible with the idea that different predicates have different preferences for instantaneous or durative interpretations (to different degrees). Thus we can distinguish between two versions of **Underspecification**, namely **Total Underspecification** and **Weighted Underspecification**. **Total Underspecification** claims that no predicate has a bias toward instantaneous or durative interpretations, in contrast to **Weighted Underspecification**, according to which different predicates have different preferences. For example, assume that *jump* pragmatically prefers a punctual interpretation. Upon encountering *jump*, the parser keeps both a punctual and a durative interpretation as possibilities, but puts a heavier weight on the former than on the latter, according to **Weighted Underspecification**,⁶ since a durative interpretation would require promotion of the less weighted durative interpretation over the more weighted punctual interpretation. **Weighted Underspecification** predicts an additional cost of the less weighted interpretation (in this case, the durative one) just as a mismatch analysis (in this case, **Iterative Coercion**) predicts.

The predictions of the two versions of **Underspecification** differ not with respect to the ranges of possible interpretations but rather with respect to how each interpretation is derived on-line. Under **Weighted Underspecification**, non-favored interpretations should induce some processing costs, in contrast to favored interpretations, whereas **Total Underspecification** predicts that both instantaneous and durative interpretations are accessible to the parser with equal ease. Thus it is natural that the distinction between **Total** and **Weighted Underspecification** is not made in the theoretical literature.

In fact, **Total Underspecification** will sound very implausible to experimentalists; every practicing experimentalist must be well aware that the parser is affected by a variety of factors, including familiarities with, and frequencies of, the words in the sentence to be parsed. It is only natural to expect different (semelfactive) predicates to be used under either instantaneous or durative interpretations in different frequencies, so it is very inconceivable that the parser is immune to such frequency effects. However, this is only a conceptual speculation, if not confirmed experimentally.

Empirically, **Weighted Underspecification** and the two **Coercion** analyses predict some processing cost differences between instantaneous and durative interpretations, while **Total Underspecification** denies such differences. Furthermore, while **Iterative** and **Punctual Coercion** predict more costs for durative and instantaneous interpretations respectively, **Weighted Underspecification** per se makes no prediction about which of durative and instantaneous interpretations are more costly; it simply depends on which predicates are biased toward which.

Thus, it would be relatively easy to argue against **Total Underspecification** based on experimental evidence, because an observation of processing cost differences between instantaneous and durative interpretations would be enough, but it would be relatively hard to argue against Weighted Underspecification, because it could cope with processing cost differences with whichever direction.

This difficulty to tease apart the predictions of **Weighted Underspecification** and a **Coercion** analysis was explicitly mentioned by Todorova et al. (2000). They conducted a self-paced reading experiment within the stop-making-sense paradigm and obtained results that durative interpretations are more costly than instantaneous interpretations. As they point out explicitly, such results are compatible with either **Weighted Underspecification** (with the assumption that punctual interpretations are preferred) or **Iterative Coercion**; their goal was modest in the sense that they only meant to confirm the existence of processing cost differences, which would argue against **Total Underspecification**. However, Brennan & Pykkänen interpret Todorova et al. (2000) as claiming to have obtained experimental support for **Iterative Coercion**, in spite of Todorova et al.'s explicit remark that their results are compatible with both **Iterative Coercion** and **Weighted Underspecification**. Clearly Brennan & Pykkänen interpreted Todorova et al.'s results this way because they did not consider the weighted version of **Underspecification** (although Todorova et al. explicitly mentioned it).

In addition to Todorova et al. (2000), Brennan & Pykkänen also mention Piñango et al. (1990), who, according to Brennan & Pykkänen, claim to have obtained experimental evidence for **Iterative Coercion**. As far as Piñango et al.'s own claim is concerned, this time Brennan & Pykkänen's interpretation is correct. Piñango et al. observed additional processing costs for durative interpretations in a cross-modal lexical decision experiment. For Piñango et al., **Iterative Coercion** is the only theoretically available option, and hence they interpreted this result as confirmation of **Iterative Coercion**.

However, as noted above, such results could well be interpreted in terms of **Weighted Underspecification**. Indeed, Piñango et al. considered the possibility that the results are due to plausibility, i.e., the possibility that their stimuli requiring durative interpretations were semantically more implausible than those requiring punctual interpretations, and that was responsible for the experimental result. They dismiss this possibility with a non-significant result of a questionnaire comparing semantic plausibilities of the two classes of stimuli. However, the questionnaire only examined off-line judgments, not on-line processes through which the stimulus sentences are processed. Less weighted interpretations, under **Weighted Underspecification**, only predict an additional processing cost, not offline semantic implausibility, and hence, their questionnaire does not argue against **Weighted Underspecification**, which still remains as a possible explanation of the result of their main experiment.

If we follow Brennan & Pykkänen and ignore the **Weighted** version of **Underspecification**, the above experimntal results would both support **Iterative Coercion**, in which case there would have been no motivation for Brennan & Pykkänen to conduct further experiments. However, Pickering et al. (2006) report results which, according to Brennan & Pykkänen, were compatible

only with **Total Underspecification**, and hence Brennan & Pykkänen intended to resolve the conflicting results of Piñang et al. (1999) and Todorova et al. (2000) on the one hand, and of Pickering et al. (2006) on the other.

Pickering et al. asked whether the observed additional processing costs associated with durative interpretations would generalize to normal comprehension situations, because in the previous experiments the participants were required to comprehend the stimuli sentences in a rather unnatural situation; in Piñang et al.'s experiment, participants were required to comprehend spoken sentences while conducting a secondary lexical decision; in Todorova et al.'s experiment, the stop-making-sense task required participants to make unnaturally early elaborated semantic decisions that would not be made in normal comprehension situations. Thus Pickering et al. examined whether Piñang et al.'s and Todorova et al.'s results would be replicated in usual self-paced reading and eye-tracking experiments, which were assumed to be closer to normal comprehension situations. Pickering et al. found no significant differences between punctual and durative interpretations, which, according to Pickering et al., suggest that **Underspecification** is the right picture for normal comprehension processes; Piñang et al.'s and Todorova et al.'s results diverted from the predictions of **Underspecification** because of the unnatural experimental settings.

Brennan & Pykkänen intended their own behavioral experiment to resolve the conflict between Piñang et al.'s and Todorova et al.'s results on the one hand, and Pickering et al.'s results on the other. However, since Pickering et al.'s had already offered an account of the seeming conflict, it was already resolved as far as Pickering et al.'s account is valid. Thus, another resolution would mean a rejection of Pickering et al.'s account and a proposal of an alternative. The question, then, will be whether the results of Brennan & Pykkänen's behavioral experiment constitute evidence against Pickering et al.'s account and a proposal of an alternative.

Before moving on to an examination of Brennan & Pykkänen's experiment, a final remark should be made about Piñang et al.'s and Todorova et al.'s experiment. The instantaneous/durative interpretations were manipulated with lexical choices of the verbal predicates in Piñang et al.'s experiment, whereas they were manipulated with a completely different means in Todorova et al.'s experiment, namely the number specifications of the grammatical objects. For example, in their experiment, sentences such as (2a) were used in the "iterative coercion" condition, in contrast to sentences such as (2b) used in the "no coercion" condition:

- (2) a. Even though Howard sent a large check to his daughter for many years, she refused to accept his money.
- b. Even though Howard sent large checks to his daughter for many years, she refused to accept his money.

The assumption here was that the underlined portion of (2a) should normally denote a single sending event but refers to an iteration in this case because of the durative modifier *for many years*, just as a lexically single-event-denoting *jump* denotes an iteration if accompanied by a

durative modifier. However, this assumption may not necessarily be correct; in classical modeltheoretical semantics, *for many years* is a (generalized) quantifier, and whether or not *a large check* is treated also as a (generalized) quantifier or as a free variable, a scope interaction is expected between *a large check* and *for many years*, in which case two different scope relations (3a–b) should be possible for (2a), which should result in interpretations (4a–b) of the first clause:

- (3) a. a large check > for many years
 b. for many years > a large check
- (4) a. There exists a large check such that, for many years, Howard sent it to his daughter.
 b. For each of the many years in question, there exists a large check such that Howard sent it to his daughter that year.

(4b) is nothing but the intended iterative interpretation. In contrast, even though the bare plural *large checks* is analyzed as a (generalized) quantifier, (2b) amounts to the same iterated interpretation irrespective of how the assumed scope ambiguity is resolved. Thus an additional processing cost associated with sentences like (2a), as compared to sentences like (2b), could well be interpreted in terms of the assumption that object (or argument) quantifiers take wider scope than modifier (or adjunct) quantifiers by default. Under such an interpretation, the difference between a single event interpretation and an iteration interpretation of *sent a large check* is a scope issue distinct from the difference between a single event interpretation and an iteration interpretation of *jump*, which is a matter of how lexical verbal predicates are to be interpreted. Thus there remains the possibility that Piñang et al.'s and Brennan & Pylkkänen's experiment on the one hand, and Todorova et al.'s experiment on the other, are dealing with different linguistic phenomena (although we will not be concerned with this issue in the rest of the paper).

3. Defects of Brennan & Pylkkänen's Experiment

This section first illustrates the design and the results of Brennan & Pylkkänen's behavioral experiment, after which two defects of their experiment will be pointed out.

3.1 Brennan & Pylkkänen's Behavioral Experiment

The design of Brennan & Pylkkänen's behavioral experiment closely mirrors that of Pickering et al.'s (2006) first experiment (an examination of the interpretations of verbal predicates, with a self-paced reading task), except for the choice of the verbs to be employed as stimuli. While verbs were not selected on a principled basis in Pickering et al.'s experiments, Brennan & Pylkkänen first conducted a pre-test, asking the participants to rate the naturalness of punctual or iterated interpretations of supposedly punctual verbs. They constructed stimuli for the main experiment employing those verbs which received high ratings for punctual interpretations in the pretest. Following Pickering et al.'s stimuli manipulations, temporal modifiers were put at

the beginning of stimuli sentences in the main experiment; thus the verbs came after the temporal modifiers. Self-paced reading was the experimental task, and significantly longer reading times were observed in the verb regions when the punctuality specifications by the modifiers and by the verbs did not coincide than when they coincided. Brennan & Pykkänen interpret this result as support for **Iterative Coercion**.

3.2 No Resolution of the Seeming Conflict

As noted above, Brennan & Pykkänen intended their experiment to resolve the seeming conflict between Piñang et al.'s and Todorova et al.'s results on the one hand, and Pickering et al.'s on the other. Such a resolution should tell us why significant processing costs were observed in the former but not in the latter. Pickering et al. had already offered such an account, but one could argue against their account and propose an alternative.

For example, one interesting characteristic of Pickering et al.'s account concerns what requires an early semantic decision. According to Pickering et al., Todorova et al. observed processing cost differences because the stop-making-sense task required an early decision, while in their own experiments, significant processing costs were not observed because nothing required such an early decision. However, overt temporal modifiers signaling instantaneous or durative interpretations were presented to the participants as parts of the stimuli sentences, and it does not make good conceptual sense that such modifiers did not function as a prompt for early semantic decisions while the Todorova et al. task does. Thus Pickering et al.'s account may not necessarily be without a problem.

However, although they claim to have resolved the seeming conflict between Piñang et al.'s and Todorova et al.'s results on the one hand, and Pickering et al.'s results on the other, Brennan & Pykkänen simply produced results similar to Piñang et al.'s and Todorova et al.'s experiment, without telling us why similar results were not observed in Pickering et al.'s experiments. This is not a resolution of the seeming conflict between Piñang et al.'s and Todorova et al.'s results on the one hand, and Pickering et al.'s results on the other.

3.3 What Hypothesis Is or Is Not Supported

On the other hand, putting aside the resolution issue, how should Brennan & Pykkänen's own experimental results be interpreted? In order to evaluate their results, we first have to distinguish three different versions of **Weighted Underspecification**.

Under **Weighted Underspecification**, which of punctual and durative interpretations are most costly depends on which are favored by verbal predicates, and nothing requires that all verbal predicates have the same preference. Thus, if all verbal predicates are assumed to favor punctual interpretations (**Version A**), **Weighted Underspecification** will predict more costs with durative interpretations than with punctual interpretations, a prediction also made by **Iterative Coercion**; if all verbal predicates are assumed to favor durative interpretations (**Version B**), it will predict more costs with punctual interpretations than with durative interpretations, a prediction also made by **Punctual Coercion**; if different predicates are assumed to have different

preferences, some favoring punctual interpretations and others favoring durative ones (**Version C**), it will predict that both punctual and durative interpretations can be more costly than the others, depending on the predicates, a prediction made by no other hypothesis.

Now, if Brennan & Pykkänen have indeed obtained processing evidence for the existence of a bias toward punctual interpretations, that would constitute an argument not only against **Total Underspecification** but also against **Punctual Coercion** on the one hand, and **Versions B and C of Weighted Underspecification** on the other. However, given their experimental design, this is not the case.

For the sake of argument, assume **Version C of Weighted Underspecification**. Their choices of the verbs to be employed in the stimuli for the main experiment was based on the results of their pre-test, in which participants rated the biases of verbal predicates to single-event or iteration interpretations; only those verbs which were found to be highly biased toward single-event interpretations were employed in the main experiment. However, under **Version C of Weighted Underspecification**, the observation of more costs with durative interpretations than with punctual interpretations, then, would be hardly surprising, because only those verbs which were expected to result in such an observation were chosen in the first place. Thus, while Brennan & Pykkänen's result is incompatible with **Version B of Weighted Underspecification**, it is compatible not only with **Version A** but also with **Version C of Weighted Underspecification**; their result does not constitute empirical evidence against **Weighted Underspecification**.

This problem also relates to the resolution issue. The only difference between the Brennan & Pykkänen experiment on the one hand, and the Pickering et al. experiment whose design was borrowed by Brennan & Pykkänen on the other, was that the verbal predicates were not chosen in a systematic way in the latter. If two experiments have produced seemingly conflicting results, the need to resolve the conflict arises. In this case, the most plausible account of the conflict is that the two experiments produced conflicting results due to the choice of the verbal predicates. This effect of the manipulation of verb choice is very easily interpreted under **Version C of Weighted Underspecification**; Pickering et al. employed both verbs with a punctual bias and verbs with a durative bias, and the processing costs of those biases canceled each other in the experimental results; Brennan & Pykkänen employed only those verbs with a punctual bias and hence observed more costs with durative interpretations. In contrast, there does not seem an obvious way to interpret the effect of the manipulation not only under **Version A of Weighted Underspecification** but also under **Iterated Coercion**. Thus, unless an unknown interpretation of the conflict is provided either under **Version A of Weighted Underspecification** or under **Iterative Coercion**, the overall results rather point toward **Version C of Weighted Underspecification**, rather than **Iterative Coercion**. However, Brennan & Pykkänen provide no such interpretation.

4. Conclusion

Indeed, **Weighted Underspecification** resembles a coercion analysis in that some processing

costs with particular interpretations are similarly predicted. For example, **Version A** of **Weighted Underspecification** predicts more costs with durative interpretations than with punctual interpretations, just as predicted by **Iterative Coercion**. Indeed, it will be very hard to tease apart the predictions of the two. However, conceptually they are different; the former claims parallel semantic parsing (i.e., multiple interpretations, or a range of interpretations, are kept in memory at one time), whereas the latter claims serial semantic parsing (i.e., only one specific interpretation is kept in memory at one time).

As far as the discussion in this paper goes, a comparison between the Pickering et al. results and the Brennan & Pytkänen result point toward a way to tease apart the predictions of **Iterative Coercion** on the one hand, and a specific version of **Weighted Underspecification** (i.e., **Version C**) on the other. However, the interpretation of the conflict between the two along the line suggested in the last section remains a speculation at this point, because it is not known empirically at this point whether the verbal predicates used in the Pickering et al. experiment are biased toward which interpretations on the one hand, and the effects of different degrees of such biases have not directly been examined (which we are planning to examine in the near future). However, one thing is certain: it is not as easy to decisively tease apart the predictions of different hypotheses as Brennan & Pytkänen assumed.

References

- Brennan, J., & Pytkänen, L. 2008. Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain & Language*, 106, 132–143.
- Kamp, H., & Reyle, U. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.
- Moens, M., & Steedman, M. 1988. Temporal ontology and temporal reference. *Computational Linguistics*, 14, 15–28.
- Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. J. 2006. Aspectual coercion and underspecification. *Discourse Processes*, 42, 131–155.
- Piñango, M. M., Zuruf, E., & Jackendoff, R. 1999. Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistics*, 28, 395–414.
- Rothstein, S. 2004. *Structuring events: A study in the semantics of lexical aspect*. Oxford: Blackwell.
- Todorova, M., Straub, K., Badecker, W., & Frank, R. 2000. Aspectual coercion and the online computation of sentential aspect. In *Proceedings of the 22nd annual conference of the cognitive science society*, Mahwah, NJ, USA: Lawrence Erlbaum Associates, pp. 3–8.

Notes

- 1 Speaking precisely, lexical ambiguity between *X* and *Y* on the one hand, and underspecification of the *X/Y* distinction in the lexicon on the other, are distinct claims; only in the former is the *X/Y* distinction part of the the grammar (in this case, the lexicon). However, the empirical predictions of the two are rather hard to tease apart, and nothing in this paper hinges upon the ambiguity/underspecification distinction. Thus the terms “ambiguity” and “underspecification” are used interchangeably throughout in this paper.
- 2 They also conducted a physiological experiment, but it will not be extensively discussed in this paper because the argument in this paper concerns the way they constructed the stimuli, which they fully specify for the behavioral experiment, but not for the additional stimuli employed in the physiological experiment.

- 3 In fact, Kamp & Reyle's (1993) analysis of the progressive, for example, does not employ the assumption (iii). Thus (iii) is not an assumption that everybody accept.
- 4 See Piãngo et al. (1999) for references.
- 5 She has more justification for distinguishing achievement and semelfactive predicates (e.g., their distinct behavior with respect to what is called the Imperfective Paradox in the modelsemantic literature), and for treating semelfactive predicates as a special class of activity (not achievement) predicates (based on her event lattices). However, they are beyond the scope of this paper.
- 6 This description might give the impression that an ambiguity conception, as opposed to underspecification, is assumed. However, an analogy will help see that underspecification, as well as ambiguity, is compatible with this description. Consider the word *politician*, for example. In a society where the majority of politicians are male, the default interpretation of this word would be a male, as opposed to female, politician. However, nobody would be willing to claim gender ambiguity for the lexical entry for this word; it would simply be underspecified for sex but has a bias toward a male interpretation.