# Introducing Theory and Evidence in Semantics

**Erhard Hinrichs & John Nerbonne**

May 26, 2008

**CENTER FOR THE STUDY**
**OF LANGUAGE**
**AND INFORMATION**

# 1

# Introducing Theory and Evidence in Semantics

## 1 The Occasion

This is a collection of papers from a one-day symposium, *Theory and Evidence in Semantics*, held on June 1, 2006 at the University of Groningen. The conference covered a good deal of ground semantically, and the present collection of papers reflects that range well, including papers on the syntax-semantics and morphology-semantics interfaces, computational forays experimenting with quantitative and application-driven approaches, and descriptive papers on problematic phenomena, including coordination, negative and collective predicates.

The broad range of topics addressed in this book will seem only appropriate if one recalls that the one-day symposium was held in honor of David Dowty, on the occasion of his sixtieth birthday. Dowty has been a far-ranging scholar in the field of linguistic semantics, and one of the most significant contributors to this field and to theoretical syntax. His well-received 1972 Texas dissertation on the temporal semantics of verbs (Dowty, 1972), supervised by Robert Wall and Emmon Bach, made some use of formal logic, but the succeeding years saw him whole-heartedly adopt the framework of logic and model-theory for his work in semantics, quickly making his mark in explorations of temporal reference and verbal aspect, culminating in the 1979 book *Word Meaning and Montague Grammar* (WMMG; Dowty 1979). Characteristically, the book was empirically ambitious, uniting analyses of inherent verbal aspect, the paradigms of tense and aspect variation, and various sorts of adverbial modification. Further, the WMMG analyses are spelled out in such complete semantic detail that attention to the syntax being interpreted follows naturally, earning the name 'grammar', and foreshadowing Dowty's later professional steps in which syntax figured

prominently. But the book is likewise reflective, considering the cognitive implications of developing models of linguistic meaning in complex and powerful mathematical logics. In addition, the book was innovative in carrying the program of logic-based semantics from the domain of sentential semantics, in which it was beginning to enjoy success, to the domain of word (lexical) meaning, which researchers up till then regarded largely as the primitives on which the theory was to be built. Finally, WMMG was impressive for its attention to earlier scholarship in linguistic semantics, and for demonstrating how rewarding attention to more traditional work could be. If Dowty had written nothing else, the 1979 book would have assured him a place in scientific history.

Later, farther reaching, contributions confirm this assessment, even elaborating on some of WMMG's themes. The opportunity to extend logic-based semantic theories to lexical processes led to a number of specific empirical analyses of e.g. the relation between active and passive clauses (and active and passive verb forms), the functioning of quantificational adverbs, as well as to a number of influential theoretical conjectures concerning, *inter alia*, the nature of grammatical relations, the proper interaction of syntax and semantics, and the nature of *thematic roles* such as AGENT, THEME, etc., and the relation between deep, licensing grammatical relations and their concrete realizations.

Dowty has been very active professionally, touring, lecturing and appearing in summer schools frequently; serving as editor-in-chief of *Linguistics and Philosophy*, the leading journal in semantics, from 1988 until 1992, and also as associate editor of *Language*, the premier linguistics journal; and serving as chairman of the prominent Department of Linguistics of The Ohio State University for several years.

The foreword to a book in David Dowty's honor is a good place to note his professional contributions and, if we may, speculate on other aspects of Dowty's intellectual impact. For Dowty is known not only for creative and compelling linguistic analysis, but also for exacting empirical and scholarly standards. These standards have led him to frame his linguistic analyses in formalized grammars, grammars in principle capable of serving fairly directly in computational models. Dowty has always distrusted sloppiness and hand-waving, preferring pages of complicated, interacting definitions over the risk of equivocation, imprecision, and implicit appeal to intuition. This was, and still is, a minority stance in the larger field of grammatical analysis, where formalization and its required attention to detail in formulation is felt to distract research from the more profound questions. David Dowty's work is always concrete, detailed, and explicit. These standards have ensured respect for his work even among colleagues with whom he has disagreed.

## 2 The Papers

This section introduces the papers contributed to this volume. The discussion will relate the papers to different facets of David Dowty's research. For this reason, the order in which the papers will be introduced here differs from the book itself, where the authors appear in alphabetical order.

As we noted above, WMMG extended the program of logic-based semantics from the domain of sentential semantics, in which it was beginning to enjoy great success, to the domain of word (lexical) meanings. **Gregory Stump**'s contribution to this volume, "Cells and Paradigms in Inflectional Semantics", may be seen to continue that ambitious program.

Stump is the originator of WORD-AND-PARADIGM MORPHOLOGY, also known as REALIZATIONAL MORPHOLOGY. While morpheme-based approaches develop rules for combining morphemes into inflected forms, or for generating inflected forms from stems, word-based morphology states generalizations that hold between the forms of inflectional paradigms. The approach generally focuses on the way in which a cell "realizes" the distinctions on which the paradigm is based. The theory is laid out in Stump (2001).

In the present paper Stump takes up the degree to which semantic distinctions in paradigms may be seen to be orthogonal. He formulates the SEMANTIC FACTORIZATION HYPOTHESIS (SFH), which requires, in brief, that paradigmatic distinctions—say between present and past tense—be realized by a single semantic operator, which, additionally, must be used in the paradigms of all the words in which the paradigmatic distinction exists—following the example of tense used above, in all the verbs in which there is a present and past tense distinction. Stump further assumes that paradigms must be morphosyntactically motivated, meaning that either morphological or syntactic correlates are sufficient grounds for postulating paradigmatic distinctions, but that a semantic distinction, by itself, constitutes insufficient grounds for postulating a paradigmatic distinction.

The Niger-Congo Twi verbal inflection system includes one verbal tense distinction between the recent and the remote past that is not consistently marked. The prefixes *à-* and *è-* are unexpectedly switched in the negative section of the verbal paradigm. Since there is no syntactic motivation for the distinction (just as there is none in English), and since paradigms must be morphosyntactically motivated, as noted above, the switch is a *prima facie* counterexample to the SFH. Stump examines several alternative avenues of analysis for this complicated data, but

finally concludes that the hypothesis is untenable.

Sanskrit distinguishes active and middle forms depending on whether an action is conducted for the benefit of the agent or for the benefit of someone else. But careful examination of this distinction across several stem classes suggests that the morphological distinction is not semantically constant, again in contradiction to the SFH. Instead, the interpretation of the middle inflectional marker depends on whether the given verb exhibits a distinction between active and middle. Stump notes that both sorts of counterexamples fall within the group of phenomena which Aronoff (1994) has called 'morphomes', phenomena whose only coherent construal is within a morphological component of grammar.

**Jack Hoeksema**'s contribution to this volume, "The *swarm* alternation revisited", re-examines the construction underlying sentences such as (1), focusing on the use of the construction in Dutch, which adds an impersonal variant to the alternation in English. Both languages allow an alternative in which *the bees* is the subject ('The bees are swarming in the garden'):

(1) a De tuin stikt van de bijen
    b The garden chokes of the bees
    c The garden is swarming with bees
    d Het stikt in de tuin van de bijen

Hoeksema reports on a collection of 1250 of these sentences that he has gleaned from Dutch corpora, thereby adding further empirical foundation to the discussion of a theoretically interesting phenomenon. The data collection serves as the basis for comparison with the description given by David Dowty in " 'The garden swarms with bees' and the fallacy of 'argument alternation' " (Dowty, 2000). Dowty dubs his analysis the 'dynamic texture hypothesis' because the variant of the construction in which the garden is subject describes a situation in which there is a small and frequently repeated event that may be found more or less throughout the location, giving it a "dynamic texture".

Hoeksema emphasizes that the predicates he finds are all predicates of abundance, a point he strengthens via appeal to native-speaker intuition, noting that the constructions sound infelicitous in combination with "downtoning" adverbials such as *a bit* or *somewhat*. This leads him to hypothesize that we are dealing with a causative degree construction. In sentences such as (1d), the object of *with*, *bees*, causes the subject *the garden* to exhibit a high degree of a property, in this case presumably the property of appearing rather full. This construal explains why subjects

need not denote locations, as in 'John was bristling with anger'. Interestingly, he likewise notes that the Dutch impersonal construction appears to be avoided when the otherwise personal subject is not locative.

Hoeksema discusses the fact that the objects of the preposition *with* in this construction are normally indefinite in English, a point which follows naturally from Dowty's hypothesis, with its emphasis on the frequently repeated events which make up the texture of the location. Hoeksema's Dutch examples, on the other hand, are frequently "fake definites" such as the one in the example above. (Note that there is no suggestion in the Dutch version of the sentence that one is reporting on a uniquely salient set of bees.) He does not add this to strengthen the case for his alternative, only to admonish that subtle, perhaps stylistic effects may rear their heads in corpus research of this type.

**Craige Roberts**' paper, "*know-how*: A Compositional Approach", is concerned with the *know-how* construction in English, i.e. with verbs that take infinitival question complements as in (2).

(2) Lingens doesn't know how to get out of the library.

Following Ryle (1949), Lewis (1979), and Abbott (2006), Roberts claims that knowing-how and knowing-that should not be treated on a par, as was recently proposed by Stanley and Williamson (2001), who provide propositional accounts for both constructions. Rather, Roberts agrees with Lewis that knowing-how, unlike knowing-that, crucially involves the property of self-ascription, i.e. attitudes *de se*. The paper presents a compositional semantic analysis of the *know how to* construction that draws on contributions from a number of authors: an account of *to* infinitival phrases due to Portner (1997), a semantics of indirect questions developed by Groenendijk and Stokhof (1997), and a semantics of infinitival questions as *hypothetical, unsaturated*, and *appropriate actions* in the sense of Dowty and Jacobson (1991). Roberts argues in detail that these accounts can be "assembled" in such a way that the stated requirements for the semantics of the *know how to* construction are met. Again drawing on the insights of Dowty and Jacobson (1991), Roberts argues that verbs which govern *to* infinitival phrases form a natural class of what she calls *epistemically reflective* predicates and that infinitival questions are unsaturated in the sense that they do not contain a phonologically null PRO subject. This latter assumption distinguishes Roberts' account from the analysis of Stanley and Williamson. As linguistic evidence in support of their propositional account of the *know how to* construction, Stanley and Williamson crucially assume that infinitival questions contain a PRO subject. By contrast, Roberts adopts

and refines a theory of control, first proposed by Jackendoff (1972) and Dowty (1985), which relies on the lexical semantics of the matrix verb, but crucially not on any syntactic element, to determine the controller of the subjectless non-finite complement. On the basis of examples such as (3), Roberts argues that control cannot be determined by the lexical semantics of the verb alone.

(3) a John asked Mary to mow the lawn.
b John asked Mary how to mow the lawn.

Roberts observes that *ask* subcategorizes either for an infinitival VP that denotes a goal, as in (3a), or for an infinitival question, as in (3b), with each choice of complement leading to distinct options for the entailed controller. In the case of (3a), the lexical semantics of the verb and the semantics of the infinitival VP together entail that the referent of the object NP is the one committed to perform the goal denoted by the infinitival VP. (3b), on the other hand, has either a generic interpretation or a subject control reading, but interestingly, no object control interpretation.

Roberts demonstrates that her compositional analysis is superior to Stanley and Williamson's analysis both in terms of descriptive adequacy and in terms of the principled and independent motivation of the various building blocks of the analyses involved.

The division of labor between semantics and pragmatics has been a long standing issue in linguistic theorizing. In his work on the nature of thematic roles (Dowty, 1991), on the *swarm* alternation (Dowty, 1991), and on the temporal interpretation of discourse (Dowty, 1986), David Dowty has addressed this question on a wide range of empirical facts. **Manfred Krifka**'s contribution, "Approximate Interpretations of Number Words: A Case for Strategic Communication", provides a pragmatic account for why "round numbers" in measure terms (such as 100) receive an approximate interpretation, whereas numbers such as 103 have a more precise interpretation.

The paper is a follow-up to previous work by the same author (Krifka, 2002) where the pragmatics of measure terms was investigated within the framework of bidirectional OT. The present paper places the overall issue within the context of strategic communication, which can formally be modeled by game theoretic means. This route has been taken by several researchers from the bidirectional OT community (such as Benz, Dekker, Jaeger, and van Rooij) and is well-motivated both conceptually and empirically.

The present paper expands on Krifka's previous analysis by compar-

ing not just expressions of different complexity, but also scales of different granularity. This solves some of the empirical problems that the previous analysis was faced with. The theory that is put forward makes use of scales that differ insofar as they are more or less fine-grained, and proposes a principle that a number expression is interpreted on the most coarse-grained scale that it occurs on. This principle can be motivated in a game-theoretic setting of strategic communication that factors in the overall likelihood of the message. The emerging theory is refined in various ways. In particular, it is shown that complexity of representations, rather than the complexity of expressions, is of crucial importance. The paper concludes with the discussion of some surprising facts about the influence that the number system of a language has on which numbers are actually expressed in that language.

In a number of very influential papers, David Dowty has shown how the framework of categorial grammar can offer elegant solutions to a wide variety of linguistic phenomena, including bound anaphora (Dowty, 2007), non-constituent conjunction (Dowty, 1988), and the treatment of grammatical relations (Dowty, 1982). Among the papers contained in this volume, the contributions by Pauline Jacobson, Chris Barker, and Neil Whitman relate most closely to this aspect of the Dowty oeuvre.

**Chris Barker**'s paper, "Reconstruction as delayed evaluation", addresses the interaction of binding and quantification in examples such as (4).

(4) [Which of his$_i$ relatives] does everyone$_i$ love _?

Such examples violate the empirical generalization derived from weak crossover facts as in (5).

(5) a. Everyone$_i$ loves his$_i$ mother. $\forall x.\mathbf{loves}\,(\mathbf{mother}\,x)\,x$

b. *His$_i$ mother loves everyone$_i$. $\forall x.\mathbf{loves}\,x\,(\mathbf{mother}\,x)$

The contrast in (5) seems to suggest that a quantifier can bind a pronoun only if it linearly precedes it. In derivational theories of syntax, examples as in (4) have given rise to the idea that at some level of representation, (part of) the bracketed material is reconstructed in the gap position and then preceded by the binding quantifier.

The goal of Barker's paper is to show how a fully compositional account of scope and binding developed by Barker and Shan (2006) can naturally account for weak crossover and reconstruction effects without having to resort to movement or copying. The account assumes a categorial grammar with flexible type assignment and with two distinct modes of combination. Apart from the familiar forward and backward slash

modes, Barker makes use of the two *continuation modes* which lead to additional categories: the functional category 'A⫽B' (stands for: "A is missing a B, with B surrounding A") and'B⧵⧵A' (stands for: "an A missing a B somewhere inside of A"). Informally speaking, the continuation mode allows a subexpression to combine with a context to form a larger expression, with the context either wrapping around (⫽) the subexpression or with the subexpression occurring inside (⧵⧵) the context.[1] Each mode of combination comes with its own set of rule schemata for syntactic and semantic combination. The combination rules for continuation operators together with type-shifting operation of Lift and Lower yield a general mechanism for scope taking of, e.g., quantifiers. Two additional continuation-mode operators ▷ (with type-shifting operator Bind) and ? (with type-shifting operator Front) are introduced to account for pronominal binding and fronting of *wh*-phrases. Weak crossover and reconstruction, including functional answers for examples such as (4), can be derived in such a flexible type system if gaps, pronouns, and *wh*-words are treated as identity functions—an idea due to Jacobson (1999)—and as scope-taking elements—a proposal due to Dowty (2007).

**Pauline Jacobson**'s paper, "Do Representations Matter or Do Meanings Matter: The Case of Antecedent Containment", raises one of the leading questions of semantic theorizing: the issue whether semantic interpretation crucially involves an intermediate level of semantic representations or whether semantic interpretation can be computed directly and without recourse to such an intermediary level representation. Jacobson considers cases of VP ellipsis that are referred to in the literature as *antecedent contained deletion* (ACD) and that have been frequently cited as key evidence in support of the view that semantic representations are crucial for semantic interpretation.

(6) Mary voted for every candidate that Bill will.

Jacobson contrasts two views of ACD and VP ellipsis: (i) the "standard view" that at least at the level of logical form (LF) there is actual linguistic material present in the ellipsis site of an elliptical VP such as *Bill will* in (6), and (ii) the alternative view that there is no material in the ellipsis site and that ellipsis is a case of anaphora, i.e. of reference to a property that is contextually salient in the discourse.

In particular, Jacobson discusses two related phenomena of antecedent containment that show parallel patterns of acceptability and that differ

[1] The concept of a *continuation* originates in computer science, where it is used, *inter alia*, to define control operators and to model the differences in call-by-name and call-by-value evaluation regimes (Plotkin, 1975).

as to whether the modifier *also* is present or absent.

(7) a. Mary voted for every boy that BILL did ~~vote for~~.
    b. *Mary voted for every boy who corresponds with a girl that BILL did ~~vote for~~.
    c. *Mary voted for every girl who corresponds with a boy who lives next door to a boy that BILL did ~~vote for~~.

(8) a. Mary voted for every boy that BILL also voted for.
    b. *Mary voted for every boy who corresponds with a boy that BILL also voted for.
    c. *Mary voted for every boy who corresponds with a girl who lives next door to a boy that BILL also voted for.

Jacobson shows that, under the standard view, the analysis of examples such as (7) and (8) crucially involves the use of variables and representational constraints prohibiting the accidental reuse of variable names. In contrast to such a representational account, Jacobson shows that a directly compositional account can be cast in a variable-free semantics that does not make use of variables in the first place and thus can do without any representational constraints on the use of variables. The account is couched in the framework of Combinatory Categorial Grammar and builds on Jacobson's earlier work on variable-free semantics (Jacobson, 1999). The use of combinators makes it possible to compose meanings in such a way that they directly mirror syntactic composition.

In the present paper, Jacobson first develops such a variable-free, directly compositional account for (8a) and then shows how the contrasts in acceptability exhibited for (8) naturally fall out of such an account. In a second step, Jacobson shows how this analysis can be generalized to the patterns of acceptability for the ACD cases in (7).

**Neil Whitman**'s piece "Right-Node Wrapping: Multimodal Categorial Grammar and the 'Friends in Low Places' Coordination" appears to describe a novel sort of construction, which he christens Right-node wrapping. These coordinations have the form [$A$ conjunction $B$] $C$ $D$ and are understood as if the element $C$ were distributed over both sides of the conjunction, while the element $D$ is interpreted only with respect to the second conjunct. Whitman offers the following example from the *Los Angeles Times*, 16 Oct. 2003:

(9) The blast [upended] and [nearly sliced] a [...] Chevrolet in half

The bracketed phrases are the conjuncts $A$ and $B$, *a Chevrolet* is the distributed object $C$, while the underscored *in half* is understood solely

in combination with the likewise underscored second verb *sliced*, and crucially not with the first conjunct *upended*. Whitman provides a long list of examples from actual use, demonstrating the existence of the construction, in spite of the suspicion which Whitman himself confesses to having felt when he first encountered it. Coordination has been studied intensively in several grammatical frameworks, and especially within categorial grammar, so that it is surprising to see a new sort of coordination discovered, even more so one which is readily instantiated in newspaper prose (and elsewhere).

Whitman's work is a clear continuation of other work on coordination in categorial grammar, most specifically work on non-constituent coordination, the earliest examples of which we are aware of being Dowty (1988) and Steedman (1985, 1990). Dowty (1988) based his account of non-constituent coordination on functional composition and type raising. In a sentence such as (10), the objects *Mary* and *Bill* are first raised from the type NP to the type (VP/NP)\VP which then compose leftwardly with the VP\VP-category adverbs *yesterday* and *today*:

(10) John saw [Mary yesterday] and [Bill today]

This paves the way for straightforward cancellation with respect to the VP/NP transitive verb *saw* and the subject.

Whitman formalizes his analysis within multi-modal categorial grammar, using a Gentzen-style rule system with an accompanying semantics. It turns out that it is sufficient to add a single rule of "mixed associativity", which is assumed not to be universal, but rather specific for English. The author contrasts this with an alternative analysis which makes uses of a unary constructor. Although both analyses cover a good deal of the data, Whitman notes some overgeneration in both analyses, as well as undergeneration of data with respect to the first.

David Dowty's work has always been reflective about methodological issues, so it is only fitting that **Peter Lasersohn** picks up on the nature of the compositionality requirement, a subject Dowty devoted a lengthy paper to (Dowty, 2007). Lasersohn's piece, "Compositional Interpretation in Which the Meanings of Complex Expressions are not Computable from the Meanings of their Parts", examines the nature of the compositionality requirement on semantic interpretation, the requirement that the meaning of a complex expression be a function of the meaning of its parts. Lasersohn notes that the requirement that meanings be compositional is occasionally defended as if it amounted to the requirement that meanings of complex constituents be computable on the basis of the meanings of their parts. This sounds psychologically plausible—after all,

one might ask, how else could speakers and hearers be able to use complex expressions which they have never heard if they cannot compute them effectively? They must have some way of computing the meanings of the complex expressions based on the meanings of their constituents. But, as Lasersohn notes, there is actually ample reason to favor the weaker requirement that there merely be a homomorphism from syntax to semantics, the formulation inherited from Frege and Montague. In this form the compositionality requirement guarantees the substitutability of expressions of the same meaning, *salve veritate*. Computability goes beyond this.

Janssen (1997) had earlier proved that compositional semantic mappings are available for a very wide range of semantically interpreted systems, and Zadrozny (1994) had noted earlier that the judicious use of pointwise defined functions, which in fact are used widely in linguistic semantics, while preserving compositionality in the homomorphic sense, allowed extremely counterintuitive mappings. As an example, he showed how the numerals might be defined in a left-branching grammar and nonetheless be interpreted compositionally. So it has long been clear that compositionality, taken by itself, is a fairly tolerant requirement.

But Lasersohn works from the other side, noting that there are non-computable functions which nonetheless might serve as the basis for semantic interpretation, even suggesting that human speakers and hearers may communicate using them as a basis. The core technical idea is straightforward: we begin with the certainty that there are non-recursively enumerable sets. But then any function mapping one-to-one to this set will be non-computable, and mathematically, such one-to-one functions must exist. To the extent that we include non-recursively enumerable sets in our domain of discourse, and we are assured that one-to-one functions exist mapping into them, we are operating with non-computable predicates. The paper is valuable for showing how plausible it is that we indeed can do so. Lasersohn provides a simple example where we straightforwardly determine the truth value of a statement in a language talking about non-recursively enumerable sets.

This leads to a reflection about the nature of the compositionality requirement and to the argument that the familiar homomorphic requirement is on target, and to a rejection of the proposal that semantic operations be computable. Lasersohn notes *inter alia* that we are comfortable with the idea that we can determine the denotation of many complex expressions even without knowing the complete denotations of all their component constituents.

The papers by Hinrichs and Wunsch and by Nerbonne and Van de

Cruys consider semantic phenomena from a computational linguistics perspective (Nerbonne, 1996). These authors use statistical methods to induce semantic regularities from large text corpora. **Erhard Hinrichs** and **Holger Wunsch**'s contribution, "Selectional Preferences for Anaphora Resolution", explores ways of extracting selectional restrictions from German text corpora for the task of computational anaphora resolution, i.e. the task of choosing the correct antecedent for a personal or reflexive pronoun in a discourse.

(11) The child munched on a cookie and smiled. It must have been {tasty/happy}.

Example (11) illustrates the importance of lexical semantics for this task. The pronoun *it* can potentially refer back to either *the child* or *a cookie*. The choice of antecedent depends crucially on the selectional restrictions of the predicate of the clause in which the pronoun appears: cookies can be tasty and children happy, but not vice versa.

While the importance of incorporating lexical semantics, particularly selectional restrictions, into anaphora resolution systems is generally recognized in computational linguistics, it is an open research question how this information can best be acquired by data-driven means.

Hinrichs and Wunsch apply latent semantic clustering (LSC), an unsupervised learning technique due to Rooth (1998), to two German newspaper corpora of radically different sizes: the manually annotated TüBa-D/Z treebank with 27,125 sentences and 473,747 lexical tokens and the automatically annotated TüPP-D/Z corpus with approx. 11.5 million sentences and 204,661,513 lexical tokens. The purpose of this comparison is to determine the quantity of data that is required to obtain meaningful results by means of unsupervised methods such as LSC.

Hinrichs and Wunsch show that the LSC method can yield reliable verb clusters only for the very large corpora TüPP-D/Z corpus. At the same time, automatic annotation of partial syntactic structure in combination with annotation of grammatical functions as in TüPP-D/Z suffices for LSC methods, as long as the annotation is sufficiently accurate and contains relevant information about clause structure.

**John Nerbonne** and **Tim Van de Cruys**'s contribution, "Quantitatively Detecting Semantic Relations: The Case of Aspectual Affinity", applies techniques from computational linguistics as well as from the statistics of dimension reduction to explore the degree to which aspect may be reflected in text. The investigation is inspired by the decades of work on temporal semantics by David Dowty and many more following him.

'Aspect' refers to the temporal make-up of predications, and in particular to whether predications implicitly refer to a completion or not. Those referring to a completion are known as TELIC and the others as ATELIC. Scholars of aspect have long noted that aspect is not marked explicitly in many languages, including Dutch and English. Thus *die* and *sing a song* are telic while *lie dying* and (intransitive) *sing* are not. The former imply completions in a way that the latter do not. The importance of the aspectual distinction is reflected in some aspects of grammar, e.g. the meaning of the progressive, as well as some inferences that depend on aspect, e.g., those involving completion.

But aspectual distinctions are not unambiguously marked on each verb, nor are verbs partitioned neatly into a small set of aspectual classes. If we wish to detect the implicit aspectual class of each verb, we must be prepared to sift through larger amounts of material, enough so that underlying dispositions are reflected in some concrete co-occurrence patterns. The "aspectual affinity" mentioned in Nerbonne and Van de Cruys's title refers to the different preferences of telic vs. atelic verbs for different classes of adverbials of duration. In brief, the telic predications take place *in, within a length of time*, while the atelic predications last *for a length of time*. This sort of affinity is also not unambiguous, however, for reasons which the semantic literature has carefully documented, and which the paper briefly reviews. The problems are reflected in the fact that one and the same predication may appear with a telic or an atelic adverbial of duration.

Nerbonne and Van de Cruys, working on Dutch, exploit a parsed corpus of 500 million words of newspaper text (approximately 27 years of daily newspapers) in order to seek the tell-tale affinities that should distinguish verbs that tend to be used to make telic as opposed to atelic publications. They build in this on the work of their colleagues in Groningen, Gertjan van Noord and Gosse Bouma, who have developed a dependency-based parser which assigns the correct dependencies in a little more than 90% of the cases (Beek et al., 2002). But the full corpus was parsed automatically and thus includes both the 90% that was parsed correctly as well as the 10% that was not.

The authors extract all the instance of clauses which include adverbials of duration by essentially listing all the prepositions which head such adverbials. They collect these into a large frequency table in which each row represents the main verb and each column the adverbial sort. The cells then represent the frequency with which the (column) adverbial occurred with the (row) head. They apply a two-dimensional version of factor analysis (singular value decomposition) in order to extract the most general affinities. While the authors are not satisfied with the

degree to which aspectual distinctions are reflect in the results, they interestingly point out association strengths of some verbs with durational adverbials that appear to reflect tendencies of use rather than aspectual affinities.

# Bibliography


Abbott, Barbara. 2006. Linguistic solutions to philosophical problems. Paper given at the annual meeting of the Pacific APA, Portland, March, 2006.

Aronoff, Mark. 1994. *Morphology by itself: Stems and Inflectional Classes*. Cambridge: MIT Press.

Barker, Chris and Chung-Chieh Shan. 2006. Types as graphs: Continuations in type logical grammar. *Journal of Logic, Language and Information* 15(4):331–370.

Beek, Leonoor van der, Gosse Bouma, and Gertjan van Noord. 2002. Een brede computationele grammatica voor het Nederlands. *Nederlandse Taalkunde* 7(4):353–374.

Dowty, David. 1972. *Studies in the Logic of Verb Aspect and Time Reference in English*. Ph.D. thesis, University of Texas.

Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Dowty, David. 1982. Grammatical relations and Montague Grammar. In P. Jacobson and G. K. Pullum, eds., *The Nature of Syntactic Representation*, pages 79–130. Dordrecht: Reidel.

Dowty, David. 1985. On recent analyses of the semantics of control. *Linguistics and Philosophy* 8(3):291–331.

Dowty, David. 1988. Type raising, functional composition, and non-constituent conjunction. In R. Oehrle, E. Bach, and D. Wheeler, eds., *Categorial Grammars and Natural Language Structures*, pages 153–197. Dordrecht: Reidel.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3).

Dowty, David. 2000. 'The Garden Swarms with Bees' and the fallacy of 'Argument Alternation'. In Y. Ravin and C. Leacock, eds., *Polysemy.*

*Theoretical and Computational Approaches*, pages 111–128. Oxford: Oxford University Press.

Dowty, David. 2007. Compositionality as an empirical problem. In C. Barker and P. Jacobson, eds., *Direct Compositionality*, pages 23–101. Oxford: Oxford University Press.

Dowty, David and Pauline Jacobson. 1991. Infinitival questions. Unpublished manuscript for a talk delivered in colloquia at The University of California at Santa Cruz and Cornell University. The Ohio State University and Brown University.

Dowty, David R. 1986. The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy* 9(1):37–61.

Groenendijk, Jeroen and Martin Stokhof. 1997. Questions. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, chap. 19, pages 1055–1124. Amsterdam: Elsevier Science/MIT Press.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–184.

Janssen, Theo. 1997. Compositionality. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, pages 417–473. Cambridge: MIT Press.

Krifka, Manfred. 2002. Be brief and vague! And how bidirectional optimality theory allows for Verbosity and Precision. In D. Restle and D. Zaefferer, eds., *Sounds and Systems. Studies in Structure and Change. A Festschrift for Theo Vennemann*, pages 439–458. Berlin: Mouton de Gruyter.

Lewis, David K. 1979. Attitudes *de dicto* and *de se*. *The Philosophical Review* 88:513–543.

Nerbonne, John. 1996. Computational semantics—linguistics and processing. In S. Lappin, ed., *Handbook of Contemporary Semantic Theory*, pages 459–82. London: Blackwell Publishers.

Plotkin, G. D. 1975. Call-by-name, call-by-value and the $\lambda$-calculus. *Theoretical Computer Science* 1:125–159.

Portner, Paul. 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5(2):167–212.

Rooth, Mats. 1998. Two-dimensional clusters in grammatical relations. In M. Rooth, S. Riezler, D. Prescher, S. Schulte im Walde, G. Carroll, and F. Beil, eds., *Inducing Lexicons with the EM Algorithm*, vol. 4.3 of AIMS, pages 7–24. Universität Stuttgart.

Ryle, Gilbert. 1949. *The Concept of Mind*. London: Hutchinson.

Stanley, Jason and Timothy Williamson. 2001. Knowing how. *The Journal of Philosophy* 98(5):411–444.

Steedman, Mark. 1985. Dependency and coordination in the grammar of english and dutch. *Language* 61(3):523–568.

Steedman, Mark. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13(2):207–263.

Stump, Gregory. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Zadrozny, Wlodek. 1994. From compositional to systematic semantics. *Linguistics and Philosophy* 17(4):329–342.