

Thematic Proto-Roles and Argument Selection

David Dowty

Ohio State University

As a novel attack on the perennially vexing questions of the theoretical status of thematic roles and the inventory of possible roles, this paper defends a strategy of basing accounts of roles on more unified domains of linguistic data than have been used in the past to motivate roles, addressing in particular the problem of *argument selection* (principles determining which roles are associated with which grammatical relations). It is concluded that the best theory for describing this domain is not a traditional system of discrete roles (Agent, Patient, Source, etc.) but a theory in which the only roles are two cluster-concepts called *Proto-Agent* and *Proto-Patient*, each characterized by a set of verbal entailments: an argument of a verb may bear either of the two proto-roles (or both) to varying degrees, according to the number of entailments of each kind the verb gives it. Both fine-grained and course-grained classes of verbal arguments (corresponding to traditional thematic roles and other classes as well) follow automatically, as do desired "role hierarchies". By examining occurrences of the "same" verb with different argument configurations (e.g. two forms of psych predicates and object-oblique alternations as in the familiar *spray/load* class), it can also be argued that proto-roles act as defaults in the learning of lexical meanings. Are Proto-Role categories manifested elsewhere in language or as cognitive categories? If so, they might be a means of making grammar acquisition easier for the child, might explain certain other typological and acquisitional observations, and may lead to an account of contrasts between unaccusative and unergative intransitive verbs that does not rely on deriving unaccusatives from underlying direct objects.

1. Introduction

There is perhaps no concept in modern syntactic and semantic theory which is so often involved in so wide a range of contexts, but on which there is so little agreement as to the nature and definition of the concept, as *thematic role* (or *thematic relation*) and its derivative, *theta-role* in Government-Binding Theory (GB). In addition to the "argument-indexing function in GB (see below), thematic roles have been invoked in the statement of multifarious syntactic generalizations in that and in other syntactic theories, and the existence of thematic roles is so taken for granted that psycholinguists now attempt to study their role in mental processing experimentally (Carlson and Tanenhaus 1988, Stowe 1989), and an introductory text in formal semantics (Chierchia and McConnell-Ginet 1990) offers a technique for formalizing roles while presupposing their necessity in a linguistic theory.

Yet apart from some syntactic correlates of thematic roles, there is in fact a notable absence of consensus as to what thematic roles are (cf. below). At best, they are obviously creatures of the syntax-semantics interface, and thus require a sound semantic theoretical basis as well as a syntactic one (and mutually consistent

ones) in order to be considered respectable parts of a linguistic theory. But at worst, appeal to them can be a confusion of notions from the syntactic, semantic and pragmatic domains, or a "thinly disguised wild card to meet the exigencies of syntax" (Jackendoff 87:371). Despite the mention of Thematic roles in the Chierchia and McConnell-Ginet text, they have hardly been studied at all in formal semantics¹, which seems to have gotten by up to this point without any significant purpose for them to serve. Ray Jackendoff, the only semanticist who has studied the concept extensively (non-model-theoretically) and who is ritually cited by syntacticians at their first mention of the notion, has evolved a detailed understanding of thematic roles within his own theoretical framework (Jackendoff 1972, 1976, 1983, 1987) that is clearly quite different from and inconsistent with that of GB or of many current syntacticians (Jackendoff 1987).

Though the term *thematic relation* (later *role*) was introduced by Gruber (1965) and made widely known by Jackendoff (1972), as semantic categories they obviously corresponded to a great extent to the (semantic) *Deep Cases* of Fillmore's contemporaneous Case Grammar (1966, 1968) --- and this in turn harks back to ideas of structuralists such as Frank Blake (1930), and ultimately to Panini's *karakas* --- but Deep Cases played a quite different part within his theory from that of thematic roles for Gruber-Jackendoff or \bar{U} -Roles do in GB. Chomsky (1981: 35), in introducing \bar{U} -roles into GB and citing precedents for the idea, claimed thematic roles such as Agent had been primitives of Donald Davidson's event-logic (Davidson 1967a), but he was in error: Davidson did not analyze events in terms of Agent and Patient but in fact rejected Hector Casteñeda's suggestion (Casteñeda 1967) that the Davidsonian event analysis be modified to do so (in Davidson 1967b:125)². Though many linguists seem to assume that linguistic theory should include a finite (and short) language-universal canon of thematic roles (that includes the familiar members such as Agent, Patient, Goal, Source, Theme, Experiencer, Instrumental, etc.), no one that I know of has ever attempted to propose a complete list³. There is disagreement even on the most familiar roles (e.g. whether Theme, usually "something that moves or changes state", can be

¹ The exceptions I know of being Chierchia (1984), Carlson (1984), and Dowty (1989), the last discussed below.

² That is, what Davidson did propose was that adjuncts (temporal, locative and adverbial modifiers) were predicates of an existentially-qualified event variable in logical form, but subject and object were not: they are traditional "arguments", related to the event variable by the n-place predicate denoted by the verb.

³ The most comprehensive list I have seen is also the earliest: Blake (1930), who argued that semantically-defined "case relationships" (clearly a similar notion to today's thematic role-types) are "numerous but not infinite; they are not indefinite and subjective, depending on the lucubrations of the individual mind, but objective, definite, and determined once for all by general grammatical principles and the laws of thought", offered as a "pioneer study" an organized system of 87 temporal and locative roles and 26 other roles, including such subsequently ignored roles as *additional* ("he gave him a sum of money *besides the cattle*"), *substitutive* ("he gave me promises *instead of money*") and *similative* ("he barked *like a dog*").

'assigned by' a stative predicate; and whether Theme is the same role as Patient or whether they are distinct), and new candidates for Thematic Roles are being proposed all the time (e.g. *Figure and Ground* in Talmy (1985a), *Neutral* in Rozwadowska 1988), *Landmark* in Jackendoff (1982), even *Subject* in Baker 1985). A paper such as this cannot begin to do justice to all the diverse literature on the subject, and a warning to this effect to the reader, plus apologies to the authors who are omitted or overlooked, are hereby issued.⁴

Among the various understandings of thematic roles, we can distinguish two kinds. What I will call the *argument-indexing* view of thematic roles is demanded by the *Ú-Criterion* of GB: each NP argument of a predicate is assigned exactly one *Ú*-Role, and the same *Ú*-Role is not assigned to two NP arguments of the same predicate (Chomsky 1981:36, 139). By clear implication, the *Ú*-Roles that Chomsky originally had in mind to fulfill this criterion were the familiar Agent, Patient, etc. from Gruber, Jackendoff, and others. By virtue of the *Ú*-Criterion, *Ú*-roles served (originally at least) two main purposes in the GB theory, (i) distinguishing "real", semantically contentful arguments of a predicate from dummy arguments such as *it* and *there*, and (ii) helping to keep track of identity and distinctness of NPs as particular semantic arguments of a predicate during the course of a derivation. From the structure of the early Case Grammar theory (Fillmore 1968), it is obvious that Deep Cases also served an argument-indexing function there, since in Deep Structure each NP argument bears exactly one case label (Agentive, Instrumental, Dative, Locative, etc.), and subsequent syntactic transformations are stated in terms of these labels, not arbitrary or tree-structurally positioned NPs, and this seems to presuppose that there is not more than one Agentive (etc.) NP per clause. (This was modified later: cf. below.)

In order for such systems to work, in an account in which the roles Agent, Theme, Goal, etc. are given explicit semantic content, the meanings of all natural language predicates must turn out to be of a very particular sort: for every verb in the language, what the verb semantically entails about each of its arguments must permit us to assign the argument, clearly and definitely, to some official thematic role or other--it cannot be permitted to hover over two roles, or seem to "fall in the cracks" between roles---and what the meaning entails about every argument must always be distinct enough that two arguments clearly not fall under the same role definition. This is a very strong empirical claim about possible meanings of natural language predicates, and, as soon as we begin to try to be precise about exactly what Agent, Patient, etc. "mean", all too subject to

⁴ In this paper I have tried to follow the practice of citing papers in which, in my view, the essence of a proposal or insight was first made, but not necessarily later discussions of the insight (under the same or different terminology) unless I believe they contributed something new that is relevant here. Hence relatively more references are made to early literature by Fillmore, Jackendoff and their contemporaries, and relatively fewer references to recent literature on roles, than is sometimes found elsewhere.

difficulties and apparent counterexamples.

Doubts as to whether the familiar short lists of Roles/Deep Cases (or refinements thereof) would ever really work this way already arose in the Case Grammar days (e.g. Cruse 1973, Huddleston 1970, Mellema 1974, Fillmore 1971a, and many others). Later GB writers saw the danger too, and proposed to circumvent the problem by refraining from committing themselves to the traditional roles, what I will henceforth call *thematic role types*, and instead invoked *individual thematic roles*---these terms from (Dowty 1989)⁵. That is, we simply call the thematic role of the subject of the verb *hit* the "hitter role", that of the subject of *kill* the "killer role", of *build* the "builder role", and so on, with no assumption made that there is one thematic role type common to these arguments (Marantz 1984, van Riemsdijk and Williams 1986)---though the possibility that role types also exist need not be ruled out, either. Trivially then, there will be enough distinct \bar{U} -roles around (i.e., the individual roles) to permit the \bar{U} -Criterion to be satisfied, i.e. to preserve the argument-indexing view of thematic roles.

Nevertheless, many syntacticians working within the GB framework have continued to appeal to the traditional thematic role types to state syntactic generalizations (Rappaport and Levin 1988, Nishigauchi 1984, Belletti and Rizzi 1986, etc.). And appeal to a particular *hierarchy* of thematic roles, such as Nishigauchi (1984) does in stating control principles (i.e. Source >...), requires ALL arguments of predicates (at least those that ever occur in control relationships), to have roles mentioned in the hierarchy, i.e. a role-type, not an individual role. Thus for such hypotheses, it IS a crucial question whether there is a small set of distinguishable role-types that effectively index all arguments.

In contrast to the argument-indexing view, Jackendoff's research on thematic roles is of a fundamentally different kind. For him, thematic relations (the term he prefers to *roles*) are most importantly notions of conceptual structure, as elucidated in Jackendoff (1983, 1987), rather than basically syntactic or interface notions; they are not theoretical primitives but are defined by particular configurations of primitive operators such as GO, STAY and CAUSE in conceptual structure; one discovers their nature and distribution empirically by looking at certain lexical and syntactic patterns in natural language in relation to their meanings, e.g. the distribution of prepositions in particular (though not, perhaps surprisingly, by psychological experiment). And the thematic roles one finds by this method do not by any means turn out to obey the theta-criterion: some verbs turn out to assign more than one role to the same argument, others assign the same role to two different arguments, and some verbs "have" thematic roles that they do not assign to any NP, e.g. *to butter* has both a Theme and a Goal role, but the Theme is

⁵ For convenience, I will continue to use *Thematic Role* for role types, when no confusion between role types and individual roles can arise.

"completely expressed by the verb" (1987:387). Whether *all* arguments of all verbs receive one of the thematic relations he has mentioned may not be stated, but his view of roles would not seem to require that they all do. In short, Jackendoff's interest in thematic roles arises purely from his desire to describe semantic patterns in lexical subcategorization and in syntax (that to him reveal conceptual structure), not to index arguments, and that thematic biuniqueness does not seem to result is of no concern. The individual-thematic-role escape hatch is of no interest to the Jackendoffs and Fillmores (or to this writer), for it ignores precisely the semantic generalization of role-type across verbs that gives the notion its interest. It should be added that Fillmore's later work on Case Grammar also permitted more than one case per argument (Fillmore 1977) (and of course he had never advocated a one-to-one relation between Deep (semantic structure) Cases and *surface* constituents).

Alas, this paper is not going to solve all these problems and does not purport to offer a theory of thematic roles that serves everyone's needs perfectly---nor does it attempt to demolish the notion once and for all. Its goal are more modest: (1) to lay out some methodological groundwork for studying thematic roles with the tools of model-theoretic semantics, and to propose some new strategies for attacking the area one step at a time, (2) to propose one new account of thematic roles (not unrelated to some other recent proposals) that seems to have merit at the "first step" in the strategy, and (3) perhaps most important of all, to make syntacticians and all linguists recognize the dangers of continuing to take this notion for granted and of assuming thematic roles are as well-motivated as phonemes or phrase-markers, and to encourage others, by this one example, to invent and explore other novel theories of thematic roles. And though this is not a psycholinguistics article and I am not a psycholinguist, I believe the linguist making a theoretical proposal about an area such as this has the responsibility to point out what psycholinguistic implications his proposal could have (the extent is it correct) and what questions it raises; thus the paper will include some speculations of this kind.

As is customary in model-theoretic semantics, I begin with the question of what *logical type* thematic roles should have, summarizing briefly the results of Dowty (1989) in §2. As the traditional empirical difficulties with arriving at a well-motivated set of role-types (most of all an argument-indexing set) may not be well-known today, I survey these in §3, including some pitfalls of misidentifying roles. It is argued in §4. that a fundamental methodological problem is that we have no agreement on what *kind* of linguistic evidence is appropriate for identifying a role-type correctly, and as a remedy, a strategy is defended of examining first the domain of *argument-selection* alone. As a further constraint on legitimate kinds of roles, I argue in §5 that *event-dependent* but not *perspective-dependent* roles be admitted. The inventory of role-types must, in view of the definitions in §4, be widened to involve a new kind of

role, *Incremental Theme* (§6.). With this preparation, I introduce a new theory of roles, in which roles are "prototypes", here called *thematic proto-roles*, rather than discrete categories (§7); the argument-selection principles for this theory and their workings are discussed in §8. Most interesting for this account of roles are three cases of subtly-contrasting argument selection: partially symmetric interactive predicates (§9.1), psychological predicates (§9.2), and the *spray-load* alternations (§1.3). Comparisons of the present view of roles with related proposals in the literature are made in §10. Some psycholinguistic implications that this account suggests for the place of thematic roles in the acquisition of grammar and of lexical meanings are considered in §11, and finally, what this account might imply about the so-called "unaccusative" phenomenon is considered in §12. The paper concludes with a brief summary of its proposals in §13.