A Feature-based System for Classifying Semantic Roles

TOM MYLNE
Griffith University
T.Mylne@mailbox.gu.edu.au

1. Introduction

Most theories of syntax include a theory of an argument list in which the (syntactic) arguments of a predicate are linked to an ordered list of variable names. Each variable name corresponds to an argument in the syntax, and to a "semantic role player" (Napoli, 1989) in the situation denoted by the predicate. The roles played by the semantic role players are commonly described in terms of THEMATIC ROLES.

Many authors (e.g. Bresnan and Kanerva, 1989; Grimshaw, 1990; Simpson, 1991) assume that thematic roles form a natural precedence set, usually called the THEMATIC ROLE HIERARCHY, which is syntactically relevant in that the order of the arguments in the argument list is identical to the order in the thematic role hierarchy of the roles they denote. The thematic role hierarchy may also be relevant to determining the (syntactic) status of an argument as external or internal (Grimshaw, 1990).

For the sake of brevity, I assume without argument that the following statements about thematic roles are true:

- A predicate "assigns" one or more ARGUMENT ROLES in the situation denoted by the predicate to the entities denoted by its argument(s).

- THEMATIC ROLES are categories of argument roles.

- It is not known how many thematic roles are required to fully categorise all possible argument roles (Dowty, 1989:70).

- There is no way to independently justify the assignment of a particular noun phrase in a particular sentence to a particular thematic role (Dowty, 1989:70).

- Particular thematic role names are used by different authors to name different categories of argument roles.

- Different authors present thematic role hierarchies which differ both as to the set of role names included and as to their order.

In brief, the problem with thematic role theory is not that it recognises the existence of different categories of argument roles, or that it attempts to formalise the "hierarchical" relationship between argument roles. A formal theory is required which will do both these things. The problem with thematic role theory lies in the formalism itself: the attempt to categorise argument roles using a system in which categories are not mutually exclusive, not exhaustive, and not intrinsically ordered.

In this paper I propose an alternative feature-based approach to the hierarchical classification of the argument roles associated with any particular predicate. The classification system proposed is tightly constrained: the number of categories is defined (six), and the hierarchical order of the categories can be related directly to the definition of the categories themselves. The classification system predicts (for English, at least) whether a particular argument role will map to an internal or an external argument, and provides a principled basis for justifying the assignment of a particular argument role to a category in that the features which define the categories function as criteria for analysis. Since the same features are used to define all the categories, the categories are by definition mutually exclusive and exhaustive.

2. An alternative to named thematic roles

Consider the alternation (1) which is more or less parallel to the pairing *fear ~ frighten* but represents a true alternation in the argument structure of a single verb.

(1)   a.   Michael worries about the situation in Bosnia. (cf. *fear*)
       b.   The situation in Bosnia worries Michael. (cf. *frighten*)

Grimshaw (1990:16) describes the arguments of both *fear* and *frighten* as an EXPERIENCER and a THEME, and argues strongly (19-20) against any reanalysis of the relative **thematic** prominence (i.e. the relative "height" on the thematic role hierarchy) of the two roles. In order to account for the inverse order of the argument list of *frighten*, Grimshaw proposes that the THEME argument of *frighten* (but no argument of *fear*) is linked to the CAUSE element in another "dimension": the **aspectual** hierarchy, which takes precedence over the thematic hierarchy in determining the order of the arguments in the argument list:

- Analysis in terms of thematic and aspectual hierarchies:
     fear   (EXPERIENCER, THEME)
     frighten   (THEME**&cause**, EXPERIENCER)

Once we accept that two "dimensions" of analysis are required, we can see that the essence of Grimshaw's analysis of *fear* and *frighten* can be reduced to the following two questions: which entity (if any) is or comes to be in a state

specified by the predicate, and which entity (if any) has a controlling role in the situation? This information can be expressed in terms of features, and the alternation in (1) can be expressed as a "switch" in the value of a single feature:

- Analysis in terms of features:
  fear/worry about   ([–Control, +SpecState], [**–Control**, –SpecState])
  frighten/worry   ([**+Control**, –SpecState], [–Control, +SpecState])

The facts about argument prominence which are captured by both of the analyses above are that a [+Control] argument is always more prominent than a [–Control] argument, but that in the absence of a [+Control] argument, a [+SpecState] argument is more prominent in the syntactic argument list than a [–SpecState] argument. These facts are captured in the Role Engagement Scale (2) in which four "zones" are defined by the interaction of the two features [Control] and [SpecState].

(2)   Role Engagement Scale (preliminary):

+Control		–Control	
–SpecState	+SpecState		–SpecState

The combination [+Control, +SpecState] is not exemplified by *fear* and *frighten*, or by *worry*, but this combination accounts (for example) for the argument role associated with *run*, which Bresnan and Kanerva (1989:24) describe as combining the thematic roles AGENT and THEME. Preliminary analysis suggests the following definitions for the two features.

An entity with a [+Control] role in a situation denoted by the verb either:
- initiates or causes (whether consciously or not) the situation as a whole;
- reacts to the situation, thereby bringing about another aspect of the situation;
- possesses an entity which is also a participant in the situation; or
- contains an entity which is also a participant in the situation (in a strong sense which excludes simply being the location at which the entity is to be found).

An entity with a [+SpecState] role in a situation denoted by the verb takes part in the situation as an "insider" to that situation: the entity is, or comes to be, in a state or location (etc.) specifically entailed by the predicate.

For two features to define four zones on a linear scale, one value of one of them must be discontinuous. In the case of (2), the discontinuity of the zone represented by [–SpecState] reflects the well-known fact that "input-only" [–SpecState] roles are found at both extremes of the thematic role hierarchy: AGENT at the [+Control] end, and roles such as LOCATION and PERCEPT at the [–Control] end. The outcome-experiencing and state-specified roles such as

BENEFICIARY, EXPERIENCE, PATIENT, THEME and LOCATUM are found towards the middle of the hierarchy, between these two extremes.

The idea that a switch in the value of a single feature can drive an alternation is not new. For example, Bresnan and Kanerva (1989:25ff) propose that locative inversion in Chichewa results from such a switch. As example (3) shows, the fronted locative in Chichewa has the subject property of triggering subject agreement on the verb, and is therefore arguably the subject of the sentence.

(3)   a.   *A-lendô-wo*   *a-na-bwér-á*   *ku-mu-dzi.*
           2-visitor-2THOSE   2-RECPAST-come-IND   17-3-village
           Those visitors came to the village.   (Bresnan & Kanerva, 1989:#2b)

      b.   *Ku-mu-dzi*   *ku-na-bwér-á*   *a-lendô-wo.*
           17-3-village   17-RECPAST-come-IND   2-visitor-2THOSE
           Those visitors came to the village.   (Bresnan & Kanerva, 1989:#1b)

According to Bresnan and Kanerva (1989:24) grammatical relations are defined in terms of two features: [±r] (thematically restricted or not restricted) and [±o] (object or non-object). The combination [+r, –o] **defines** the grammatical relation OBLIQUE, *and* [–r, –o] **defines** the grammatical relation SUBJECT. Locatives are assigned [–o] as a result of their "intrinsic meaning", and are usually assigned the feature [+r] by default, with the result that they surface as syntactic obliques, as in (3)a. Switching [+r] to [–r] causes the locative argument to surface as the subject of the verb, as in (3)b.

The principle of driving an alternation by switching a single feature can be translated to the analysis of the *worry* alternation (1). Michael has a [+SpecState] role and the situation in Bosnia has a [–SpecState] role in both alternants. A reversal of the hierarchical relation can be triggered by assigning the feature [+Control] to the situation in Bosnia in (1)b as opposed to [–Control] in (1)a.

(4)   Role engagement analysis of *worry* alternation (first approximation):

Predicate	+Control		–Control	
	–SpecState	+SpecState		–SpecState
worry (a):		Michael		Bosnia
worry (b):	Bosnia	Michael		

It is not necessary for the purpose of determining the hierarchical order of arguments in these clauses to inquire what value of the feature [Control] should be assigned to Michael, but this feature is relevant to the internal or external status of this argument. Like all intransitive verbs showing the causative~inchoative alternation, *worry* in (1)a is an unaccusative verb in which *Michael* is an internal argument, as *Michael* certainly is in the transitive

alternant (1)b. Grimshaw's (1990:39) explanation for the existence of unaccusative verbs is that the single argument of such a verb cannot be an external argument because its role is not "maximally [i.e. sufficiently] prominent" on the thematic role hierarchy. The spirit of this suggestion can be transferred to (2) by proposing that only (and all) roles which rate as [+Control] are mapped to external arguments, whereas a role which is [–Control] does not rate highly enough and must be mapped to an internal argument even if there is no argument higher on the scale. In other words, *Michael* denotes a [–Control] role in both (1)a and (1)b, and the alternation (1) corresponds to the following role engagement analysis:

(5)   Role engagement analysis of *worry* alternation:

Predicate	+Control		–Control
	–SpecState	+SpecState	–SpecState
worry (a):		Michael	Bosnia
worry (b):	Bosnia	Michael	

The grammatical relations in (1) can be predicted from the role engagement analysis above. The [+SpecState] role in (1)a maps the subject grammatical relation, since it is the most prominent. However the subject is an internal argument, since its role is marked [–Control]; the other role must therefore surface in an oblique grammatical relation. The argument list for (1)b is that of an ordinary transitive verb, and includes one external and one internal argument.

The analysis of the *worry* alternation applies equally well to the Chichewa locative inversion in (3). Like Michael in (1), the visitors in (3) have [+SpecState, –Control] roles in both alternants. (Bresnan and Kanerva (1989) show that locative inversion occurs only with unaccusative verbs.) Like the situation in Bosnia in (1), the village in (3) has a [–SpecState] role: it is not and does not come to be in a state or location entailed by the predicate. Locations are not normally regarded as exercising control, but the notion of [Control] is evidently broad enough in Chichewa to allow a speaker to choose to assign the feature [+Control] to the location in the absence of another [+Control] role. The location's [–SpecState] status is of course unaltered by the switch in its [Control] status.

(6)   Role engagement analysis of Chichewa locative inversion:

Predicate	+Control		–Control
	–SpecState	+SpecState	–SpecState
'come' (a):		visitors	village
'come' (b):	village	visitors	

The role engagement analyses above show a switch in only one feature of one argument role: the theme/percept or locative argument is [–Control] in one alternant and [+Control] in the other. It is possible that all "argument structure alternations" properly so-called reflect an underlying change in the speaker's evaluation of (exactly) one feature of the level of engagement of (exactly) one semantic role player.

The role engagement scale (2) also accounts for the alternation shown in (7). The ditransitive causative construction (7)a includes two [+Control] (agent-like) roles, but Catherine's role in (7)a is clearly different from the child's role. The difference can be characterised in terms of [SpecState]: Catherine is a [–SpecState] "outsider" agent while the child is a [+SpecState] "insider" agent. In this case, switching the child's role from [–SpecState] to [+SpecState] leaves room for the insertion of another argument.

(7) a. *Mwana   wake a-na-kolol-e   chimanga.*
         child   her   AGR-PAST-harvest-ASP corn
         Her child harvested corn.   (Chichewa)

     b. *Catherine a-na-kolol-ets-a   mwana wake chimanga.*
         Catherine AGR-PAST-harvest-CAUS-ASP   child   her   corn
         Catherine made her child harvest corn.   (Baker, 1988:Ch.4:#3b)

(8)   Role engagement analysis of morphological causative:

Predicate	+Control		–Control	
	–SpecState	+SpecState		–SpecState
'harvest' (a)	child		corn	
'harvest' (b)	Catherine	child	corn	

3. Refinement of the Role Engagement Scale

Whenever the *load* alternation (9) is discussed, it is noted that part of the meaning of the *with hay* alternant (9)b is that the wagon becomes fully loaded, while the *onto the wagon* alternant (9)a does not include this interpretation.

(9) a. Henry loaded hay onto the wagon. (Rappaport and Levin, 1986:#22a)
     b. Henry loaded the wagon with hay. (Rappaport and Levin, 1986:#22b)

In terms of our features, the wagon in (9)b has a [+SpecState] (affected) role whereas the wagon in (9)a has a [–SpecState] (purely locative) role. The hay, by contrast, is a [+SpecState] participant in both alternants: a change in location is

entailed by the predicate in both cases. A preliminary role engagement analysis of the alternation in (9) is as follows:

(10)   Role engagement analysis of *load* (preliminary):

Predicate	+Control			–Control
	–SpecState	+SpecState		–SpecState
load (a)	Henry		hay	wagon
load (b)	Henry		wagon, hay	

In (9)b both the hay's role and the wagon's role are [–Control, +SpecState] roles and a third feature is required to distinguish the two roles. The obvious choice of name for such a feature is [Affected]. An affected role is "ultra-specified". Not only does the entity experience an outcome of the situation; the outcome is "internal" to the entity.

An entity with a [+Affected] role in a situation denoted by the verb undergoes a change which affects the entity's integrity, or alters the entity's internal or cognitive state, and is specifically entailed by the predicate.

Adding the third feature to (2) yields the fully developed role engagement scale (11). Since [+Affected] entails [+SpecState], the three features combine to yield six (not eight) "generalised thematic roles". [Affected], like [SpecState], is a scale-centripetal feature: a [+Affected] role rates lower than other [+Control] roles, but higher than other [–Control] roles.

(11)   Role Engagement Scale:

Engagement level defined by features	+Control				–Control	
	–SpecState	+SpecState				–SpecState
		-Affected	+Affected		–Affected	
Description	*Proactive*	*Reactive*	*Responsive*	*Patientive*	*Activated*	*Static*
Some standard thematic role labels	*Agent*	*Recipient Beneficiary*	*Experience (Perceiver)*	*Patient*	*Theme Instrument*	*Location Goal Source Path Percept*

We are now in a position to fully analyse the alternation in the [–Control] roles in (9). Whereas the wagon's role in (9)a is a [–SpecState] role, the wagon's role in (9)b is a [+SpecState] role. Since the hay's role occupies the [+SpecState, –Affected] space on the role engagement scale, the wagon can only be conceptualised as [+SpecState] participant if it is also conceptualised as [+Affected].

(12)   Role engagement analysis of *load* alternation:

Predicate	+Control			−Control	
	−SpecState	+SpecState			−SpecState
		-Affected	+Affected	−Affected	
load (*onto*):	Henry			hay	wagon
load (*with*):	Henry		wagon	hay	

Other pairs similar to (9) are given below. The situations described by all eight sentences are like those described in (9) in that a theme entity moves towards a goal entity under the control of an agent entity. The situations are also alike in that the (b) alternants demand an affected patient interpretation for goal of motion, whereas the (a) alternants inhibit such an interpretation.

(13) a.   Betty passed the needle through the cloth.
     b.   *Betty passed the cloth with a needle.

(14) a.   *George cut the knife through the cloth.
     b.   George cut the cloth with a knife.

(15) a.   ?Alison pierced the needle through the cloth.(Levin, 1993 Part 1:#216a)
     b.   Alison pierced the cloth with a needle.   (Levin, 1993 Part 1:#216b)

(16) a.   Rita stabbed the knife into the carcass (?the intruder).
     b.   Rita stabbed the intruder (?the carcass) with a knife.

The (b) alternant is not permitted with *pass* because *pass* does not include the possibility of an affected patient in its meaning, whereas *cut*, which can only be used to describe a situation in which an entity is affected, cannot be used in the (a) alternant. Levin (1993) finds (15)a acceptable, but those speakers who conceptualise *pierce* as entailing "the goal participant is affected" will find (15)a unacceptable for the same reason as (14)a is unacceptable. In the case of *stab*, the (b) alternant is more felicitous when an affected patient interpretation is expected, whereas the (a) alternant favours the simple goal interpretation.

4. Conclusion

Although most thematic role theorists aim at a theory of semantic role categories which is not only exhaustive but also cross-linguistically universal, discussions of thematic role theory often focus on sub-sets of verbs, such as those which assign an agent role, the so-called "psych-verbs" and other such sub-sets. A similar criticism could be levelled at the analysis herein in which the proposed role engagement scale and the three features on which it is based have been illustrated with only a small number of examples, from a small number of languages.

Despite this, the feature-based role engagement scale is offered as a system of categorisation of semantic roles which is tightly constrained as to its operation, perspicuous in its analytic capacity, and worthy of further investigation. An important aspect of the proposal is its parsimony. Distinctions which (by hypothesis) will never be activated, such as that between PERCEPT and LOCATION, both analysed above as [–Control, –SpecState], are not made. Conversely, distinctions which must be made can be made, on the basis of three relatively simple criteria.

References

Bresnan, Joan and Jonni M. Kanerva (1989). Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry 20*:1-50.

Dowty, David (1989) "On the Semantic Content of the Notion of 'Thematic Role'." In Chierchia, Gennaro, Barbara H. Partee, and Raymond Turner. *Properties Types and Meaning Volume II: Semantic Issues*. Dordrecht: Kluwer Academic Publishers, pp.69-130.

Grimshaw, Jane (1990) *Argument Structure*. Cambridge MA: MIT Press.

Levin, Beth (1993). *English Verb Classes and Alternations: A preliminary investigation*. Chicago: U of Chicago Press.

Napoli, Donna Jo (1989) *Predication Theory*. Cambridge: CUP.

Rappaport, Malka and Beth Levin (1986). *What to Do with Theta-Roles*. Lexicon Project Working Papers #11. Cambridge MA: MIT Center for Cognitive Science

Simpson, Jane (1991) *Warlpiri Morpho-Syntax: A lexicalist approach*. Dordrecht: Kluwer Academic Publishers.