

# Chapter 6

## Sentence Semantics 2: Participants

### 6.1 Introduction: Classifying Participants

In the last chapter we looked at aspects of sentence level semantics: how speakers may choose to characterize situations and express various degrees of commitment to the portrayal. Another set of semantic choices which face a speaker seeking to describe a situation concerns how to portray the roles of any entities involved. Take for example 6.1 below:

6.1 Gina raised the car with a jack.

This sentence identifies three entities, Gina, *the car* and *a jack* related by the action described by the verb *raise*. The sentence portrays these entities in specific roles: Gina is the entity responsible for initiating and carrying out the action, the car is acted upon and has its position changed by the action, and the jack is the means by which Gina is able to cause the action. Such roles have a number of labels in semantics, including participant roles (Allan 1956), deep semantic cases (Fillmore 1968), semantic roles (Givon 1990), thematic relations (Gruber 1976, Jackendoff 1972) and thematic roles (Dowty 1986, 1989, 1991, Jackendoff 1990). Given its wide usage in recent work, we will use the last term here: **thematic roles**.

In this chapter we examine this notion of thematic roles. We begin by sketching the basic picture of these roles that seems to be assumed by much of the syntax and semantics literature. Thus in sections 6.2-6.4 we outline the main contenders for individual types of roles, look at the relationship between thematic roles and grammatical relations, and discuss the idea that verbs must have their thematic role requirements listed in the lexicon. In the second part of the chapter we look more critically at the idea of thematic roles: first in section 6.5 we review criticisms that have been levelled at the notion. Then in 6.6 we review the job these roles do in linguistic description. In the third part of the chapter, section 6.7, we investigate **voice** systems and see how they allow speakers some flexibility in the relationship between thematic roles and grammatical structure: we focus on **passive** voice and **middle** voice. In the final part of the chapter we turn our attention to semantic classification systems that are based on the inherent features of nominals rather than their roles within a predication. In section 6.8.1 we discuss **classifiers** and in 6.8.2 **noun classes**.

### 6.2 Thematic Roles

Each of the writers mentioned above, and others, for example Andrews (1985) and Radford (1988), have proposed lists of thematic roles. From this extensive literature we can extract a list of thematic roles like the following (where the relevant role-bearing nominal is in bold):

AGENT: the initiator of some action, capable of acting with volition, e.g.

6.2 **David** cooked the rashers.

6.3 **The fox** jumped out of the ditch.

PATIENT: the entity undergoing the effect of some action, often undergoing some change in state, e.g.

6.4 Enda cut back **these bushes**.

6.5 The sun melted **the ice**.

THEME: the entity which is moved by an action, or whose location is described, e.g.

6.6 Roberto passed **the ball** wide.

6.7 **The book** is in the library.

EXPERIENCED the entity which is aware of the action or state described by the predicate but which is not in control of the action or state, e.g.

6.8 **Kevin** felt ill.

6.9 **Mary** saw the smoke.

6.10 **Lorcan** heard the door shut.

BENEFICIARY: the entity for whose benefit the action was performed, e.g.

6.11 Robert filled in the form for **his grandmother**.

6.12 They baked **me** a cake.

INSTRUMENT: the means by which an action is performed or something comes about, e.g.

6.13 She cleaned the wound with **an antiseptic wipe**.

6.14 They signed the treaty with **the same pen**.

LOCATION: the place in which something is situated or takes place, e.g.

6.15 The monster was hiding **under the bed**.

6.16 The band played **in a marquee**.

GOAL: the entity towards which something moves, either literally as in 6.17 or metaphorically as in 6.18:

6.17 Sheila handed her licence **to the policeman**.

6.18 Pat told the joke **to his friends**.

SOURCE: the entity from which something moves, either literally as in 6.19 or metaphorically as in 6.20:

6.19 The plane came back **from Kinshasa**.

6.20 We got the idea **from a French magazine**.

STIMULUS: the entity causing an effect (usually psychological) in the EXPERIENCER, e.g.

6.21 John didn't like **the cool breeze**.

6.22 **The noise** frightened the passengers.

Thus to return to our first example, repeated below:

6.23 Gina raised the car with a jack.

we can describe the thematic roles by calling *Gina* the AGENT of the action, *the car* the THEME, and *the jack* the INSTRUMENT.

There is some variation in the use of these terms: for example Radford (1988) treats PATIENT and THEME as different names for the same role. Here we adopt the distinction that PATIENT is reserved for entities acted upon and changed by the verb's action while THEME describes an entity moved in literal or figurative space by the action of the verb, but constitutionally unchanged. Thus the noun phrase *the rock* would be a PATIENT in 6.24 below but a THEME in 6.25

6.24 Fred shattered **the rock**.

6.25 Fred threw **the rock**.

A number of tests for identifying thematic roles have been suggested. Jackendoff (1972) for example provides a test for AGENT: whether the phrases like *deliberately*, *on purpose*, *in order to*, etc. can be added to the sentence. This reflects the fact that an AGENT characteristically displays animacy and volition. The contrast between 6.26 and 6.27 below identifies John as an AGENT in 6.25 but not 6.27:

6.26 **John** took the book from Bill in order to read it.

6.27 **?John** received the book from Bill in order to read it.

Some writers (e.g. Foley and Van Valin 1984, Jackendoff 1990) have suggested that AGENT is a particular type of a more general thematic role ACTOR, where ACTOR 'expresses the participant which performs, effects, instigates, or controls the situation denoted by the predicate' (Foley and Van Valin 1984: 29). So every AGENT is an ACTOR, but not the other way round: in 6.28 below *the car* is an ACTOR but not AGENT since it presumably is neither in possession of a wish to kill nor animate:

6.28 **The car** ran over the hedgehog.

Other simple tests suggested by Jackendoff (1990) include predicting that for an ACTOR (X) it will make sense to ask 6.29 below, and for a PATIENT (Y) that it will be able to occur in the frames in 6.30:

6.29 What did X do?

6.30 a. What happened to Y was . . .  
b. What X did to Y was . . .

So for example 6.31 below the tests would give 6.32-3, identifying *Robert* as the ACTOR and *the golf club* as PATIENT:

- 6.31 Robert snapped the golf club in half.
- 6.32 What Robert did was to snap the golf club in half.
- 6.33 a. What happened to the golf club was that Robert snapped it in half.  
b. What Robert did to the golf club was snap it in half.

Some writers have suggested other thematic roles in addition to those we have discussed. For example a role of FORCE is sometimes used instead of INSTRUMENT for an inanimate entity which causes something, e.g.

- 6.34 a. **The wind** flattened the crops.  
b. The sea wall was weakened by **the waves**.

A role of RECIPIENT is sometimes identified, e.g. by Andrews (1985), as a type of GOAL involved in actions describing changes of possession, e.g.

- 6.35 a. He sold **me** this wreck.  
b. He left his fortune **to the church**.

While these roles, ACTOR, AGENT, PATIENT, EXPERIENCER, THEME, INSTRUMENT etc. may seem intuitively clear, in practice it is sometimes difficult to know which role to assign to a particular noun phrase. For example, in a sentence like 6.36 below *to the lighthouse* is clearly a GOAL, and in 6.37 *him* is a BENEFICIARY, but in 6.38 below is *Margarita* the GOAL/RECIPIENT, or the BENEFICIARY, or both?

- 6.36 Fergus carried the bag **to the lighthouse**.
- 6.37 Sylvie bought **him** a sports car.
- 6.38 **Margarita** received a gift of flowers.

Examples like these raise the difficult question of whether a single entity can fulfil two or more thematic roles at the same time; for example in 6.39 below, are we to say that Mr Wheeler is both AGENT and THEME?

- 6.39 **Mr Wheeler** jumped off the cliff.

These issues are still under investigation in various theoretical approaches. A central claim of Chomsky's Principles and Parameters theory, for example, is the **Theta-Criterion**, which states that there must be a one-to-one

correspondence between noun phrases and thematic roles (see Chomsky 1988, Haegeman 1994). Jackendoff (1972), on the other hand, suggested that one entity might fulfil more than one role. In Jackendoff (1990) the idea that one nominal might fulfil more than one role is elaborated into a theory of tiers of thematic roles: a **thematic tier**, which describes spatial relations, and an **action tier** which describes ACTOR-PATIENT type relations. His examples include the following (1990: 126-7):

- |      |    |               |           |                 |
|------|----|---------------|-----------|-----------------|
| 6.40 | a. | Sue hit       | Fred.     |                 |
|      |    | Theme         | Goal      | (thematic tier) |
|      |    | Actor         | Patient   | (action tier)   |
|      | b. | Pete threw    | the ball. |                 |
|      |    | Source        | Theme     | (thematic tier) |
|      |    | Actor         | Patient   | (action tier)   |
|      | c. | Bill entered  | the room. |                 |
|      |    | Theme         | Goal      | (thematic tier) |
|      |    | Actor         |           | (action tier)   |
|      | d. | Bill received | a letter. |                 |
|      |    | Goal          | Theme     | (thematic tier) |
|      |    |               |           | (action tier)   |

Thus *Fred* in 6.40a is simultaneously the GOAL and the PATIENT of the action. The gaps in a tier reflect instances where the nominal has only one thematic role: thus *the room* in 6.40c has no role in the action tier. Presumably these tiers would divide thematic roles into two types, perhaps as follows:

- 6.41 a. Action tier roles: ACTOR, AGENT, EXPERIENCER, PATIENT, BENEFICIARY, INSTRUMENT.  
b. Thematic tier roles: THEME, GOAL, SOURCE, LOCATION.

To these dimensions of action and space, Jackendoff also proposes a dimension of time, which we will not investigate here. The basic insight is clear: the roles that speakers assign to entities may be more complicated than a single thematic role label. For a detailed discussion of this proposal, see Jackendoff (1990: 125-51).

Having identified these thematic roles, the next question we might ask is: how are such roles identified in the grammar? For our English examples above, the answer is by a combination of syntactic structure and the choice of verb. There are typical matchings between participant roles and grammatical relations. As in our original example 6.22, the subject of the sentence often corresponds to the AGENT, the direct object to the THEME, while the INSTRUMENT often occurs as a prepositional phrase. Though this is the typical case, it is not necessarily so: for example it is possible to omit the AGENT from the sentence and as a result have the INSTRUMENT occupy subject position, e.g.:

6.42 **The jack** raised the car.

We can see the effect of the choice of verb if we try to describe this same situation without either the AGENT or the INSTRUMENT. We cannot simply allow the THEME to occupy subject position as in 6.43; we have to change the verb as in 6.44:

6.43 \*The car raised.

6.44 The car rose.

This is because the verb *raise* requires an ACTOR. The verb *rise* however describes a change of state without any slot for an ACTOR so that while 6.44 above is fine, 6.45 and 6.46 below are not possible:

6.45 \*Gina rose the car.

6.46 \*The jack rose the car.

What this simple example shows is that a speaker's choice of participant roles has two aspects: the choice of a verb with its particular requirements for thematic roles, and within the limits set by this, the choice of grammatical relations for the roles. We look at these choices in the rest of this chapter, beginning with the relationship between thematic roles and grammatical relations: first we describe how various thematic roles may occupy subject position, then we look briefly at the selection of thematic roles as part of a verb's lexical semantics. Later we discuss the role of *voice* in allowing speakers to alter prototypical matchings between thematic roles and grammatical relations.

### 6.3 Grammatical Relations and Thematic Roles

We have seen that while in English there is a tendency for subjects to be AGENTS, direct objects to be PATIENTS and THEMES, and INSTRUMENTS to occur as prepositional phrases, this need not always be the case. There are two basic situations where this is not the case: the first is where roles are simply omitted, and the grammatical relations shift to react to this, as we will discuss in this section; and the second is where the speaker chooses to alter the usual matching between roles and grammatical relations, a choice often marked by an accompanying change of verbal *voice*. We deal with voice later on in section 6.7.

We can begin with a simple example of thematic role omission in 6.47-9 below:

6.47 Ursula broke the ice with a pickaxe.

6.48 The pickaxe broke the ice.

6.49 The ice broke.

This is similar to our example 6.23 earlier: in 6.47 Ursula is the AGENT and subject, the ice is PATIENT and direct object, and the pickaxe, the INSTRUMENT, is in a prepositional phrase. In 6.48 the AGENT is omitted and now the INSTRUMENT is subject; and finally in 6.49 with no AGENT or INSTRUMENT expressed, the PATIENT becomes subject. The verb *breaks* unlike *raise* earlier, allows all three thematic roles to occupy subject position. Several writers have suggested that this process of different roles occupying the subject position is a hierarchical process, not only in English but across many languages. The observation is that when speakers are constructing a sentence, they tend to place an AGENT into subject position, the next preference being for a RECIPIENT or BENEFACTIVE, then THEME/PATIENT, then other roles. From our English examples, it seems that INSTRUMENT is then preferred to LOCATION. This is sometimes described as an *implicational hierarchy*. There are various versions of such a hierarchy proposed in the literature, e.g. in Fillmore (1968) and Givon (1984b), but we can construct a simple example of a universal subject hierarchy like 6.50 below:

6.50 AGENT > RECIPIENT/BENEFACTIVE > THEME/PATIENT > INSTRUMENT  
> LOCATION

This diagram can be read in two equivalent ways: one is that the leftmost elements are the preferred, most basic and expected subjects, while moving rightward along the string gives us less expected subjects. A second way to read this diagram is as a kind of rule of expectation, going from right to left: if a language allows the LOCATION role to be subject, we expect that it will allow all the rest. If, however, it allows the role INSTRUMENT to be subject, we expect that it allows those roles to the left, but we don't know if it allows the LOCATION role as subject. The idea is that languages can differ in what roles they allow to occur as subject but they will obey this sequence of preference, without any gaps. So, for example, we should not find a language that allows AGENT and INSTRUMENT to be subject but not THEME/PATIENT.

It is a little difficult to think of English examples with LOCATION as subject, unless we include sentences like 6.51a-b below:

6.51 a. **This cottage** sleeps five adults.  
b. **The table** seats eight.<sup>1</sup>

but the other positions on the hierarchy occur regularly, as we can see from the following examples:

- 6.52 AGENT subjects:  
**The thief** stole the wallet.  
**Fred** jumped out of the plane.
- 6.53 EXPERIENCER subjects:  
**I** forgot the address.  
**Your cat** is hungry.
- 6.54 RECIPIENT subjects:  
**She** received a demand for unpaid tax.  
**The building** suffered a direct hit.
- 6.55 PATIENT subjects:  
**The bowl** cracked.  
**Una** died.
- 6.56 THEME subjects:  
**Joan** fell off the yacht.  
**The arrow** flew through the air.
- 6.57 INSTRUMENT subjects:  
**The key** opened the lock.  
**The scalpel** made a very clean cut.

See Comrie (1981) and Croft (1990) for discussion of this and other implicational hierarchies.

## 6.4 Verbs and Thematic Role Grids

As we saw earlier with the verbs *raise*, *rise* and *drive*, verbs have particular requirements for their thematic roles. Since this is part of a speaker's semantic knowledge about a verb, we might expect it to be part of the lexical information stored for verbs. Thus we need to know not only how many arguments a verb requires (i.e. whether it is intransitive, transitive, etc.) but also what thematic roles its arguments may hold.

In the generative grammar literature, this listing of thematic roles is often called a **thematic role grid**, or **theta-grid** for short.<sup>2</sup> A simple example might be:

- 6.58 **put** V: <AGENT, THEME, LOCATION>

This entry tells us that *put* is a three-argument, or ditransitive, verb and spells out the thematic roles the three arguments may carry. Here we show

Williams's (1981) suggestion of underlining the AGENT role to reflect the fact that it is this role that typically occurs as the subject of the verb (or 'external argument' in Williams's terminology). Clearly this is just the start of the job that a grammatical description must do of mapping between thematic roles and grammatical categories and structures. Our thematic grid for *put* in 6.58 predicts that this verb, when saturated with the correct arguments, might form a sentence like 6.59:

- 6.59 John<sub>AGENT</sub> put the book<sub>THEME</sub> on the shelf<sub>LOCATION</sub>.<sup>3</sup>

Of course not all nominals in a sentence are arguments of a verb and thus specified in verbal theta-grids in the lexicon. We will make the assumption that one can employ grammatical tests to identify arguments: for example to distinguish between the role of argument played by the prepositional phrase *in the bathroom* in 6.60 below and its status as a non-argument in 6.61:

- 6.60 [S Roland [VP put [NP the book] [PP in the bathroom]]]

- 6.61 [S Roland [VP read NP the book]] [PP in the bathroom]]

The square brackets in 6.60-61 reflect the fact that while *in the bathroom* is an argument of the verb *put*, explaining why it cannot be omitted:

- 6.62 \* Roland put the book.

it is not an argument of the verb *read*, for example, which can form a sentence without it:

- 6.63 Roland read the book.

In grammatical terms, while *in the bathroom* is an argument in 6.60, it is an **adjunct** in 6.61. As well as not being required by the verb, adjuncts are seen as less structurally attached to the verb, explaining why 6.64 below is a much more unusual word order than 6.65, and usually requires a marked intonation pattern:

- 6.64 In the bathroom Roland put a book.

- 6.65 In the bathroom Roland read a book.

See Radford (1988) and Haegeman (1994) for discussion of the grammatical status of arguments and adjuncts. We will assume that all verbs may co-occur with adjuncts (usually adverbials of time, place, manner, etc.) and that requirements need only be listed in the lexicon for arguments.

Another way of making this distinction is to distinguish between **participant roles** and **non-participant roles**. The former correspond to our arguments: they

they are needed by the predication, in the sense we have been discussing; the latter are optional adjuncts which give extra information about the context, typically information about the time, location, purpose or result of the event. Of course only participant roles will be relevant to verbal thematic grids, and our discussion in this chapter focuses on these participant roles.

Listing thematic grids soon reveals that verbs form classes which share the same grids. For example English has a class of TRANSFER, or GIVING, verbs which in one subclass includes the verbs *give, lend, supply, pay, donate, contribute*. These verbs encode a view of the transfer from the perspective of the AGENT. They have the thematic grid in 6.66; 6.67 is an example:

6.66 V: <AGENT, THEME, RECIPIENT>

6.67 Barbara<sub>AG</sub> loaned the money<sub>TH</sub> to Michael<sub>RE</sub>.<sup>4</sup>

Another subclass of these TRANSFER verbs encodes the transfer from the perspective of the RECIPIENT. These verbs include *receive, accept, borrow, buy, purchase, rent, hire*. Their thematic grid is in 6.68, with an example in 6.69, paralleling 6.67 above:

6.68 V: <RECIPIENT, THEME, SOURCE>

6.69 Michael<sub>RE</sub> borrowed the money<sub>TH</sub> from Barbara<sub>so</sub>.

Thematic grids such as these are put to use in the literature for a variety of descriptive jobs. We can look at some of these in section 6.6, when we ask more generally: what purpose do thematic roles serve in linguistic analysis? First though we discuss some of the problems associated with the simple picture of thematic roles we have outlined so far.

## 6.5 Problems with Thematic Roles

In our introductory discussion, we mentioned that the lists of roles given in the literature have varied from author to author. Authors disagree about what if any distinctions are to be made between PATIENT and THEME, for example, or between AGENT and related roles like ACTOR, EXPERIENCER, etc.

We can see these debates as reflections of two general problems with thematic roles (usually abbreviated to ‘theta-roles’, sometimes also called  $\Theta$ -roles). The first problem is really about delimiting particular roles. The extreme case would be to identify individual thematic roles for each verb: thus we would say that a verb like *beat* gives us two theta-roles, a BEATER - role and a BEATEN -role. This would of course reduce the utility of the notation: if we lose the more general role types like AGENT, PATIENT etc., then we cannot make

cannot make the general statements about the relations between semantic roles and grammatical relations discussed earlier, nor put theta-roles to any of the uses we describe in the next section.

But if we are to classify individual theta-roles like BEATER and BEATEN into theta-role types like AGENT and PATIENT, we will have to find some way of accommodating variation within the role type. Let us take the example of PATIENT in a typical grid:

6.70 V: <AGENT, PATIENT, INSTRUMENT>

A typical example would be 6.71:

6.71 The child<sub>AG</sub> cracked the mirror<sub>PA</sub> with his toy<sub>m</sub>.

Earlier we defined the PATIENT as the entity affected by the action of the verb. However, attempts to examine particular verbs, such as Dixon (1991), reveal that both the type of ‘affectedness’ and the role of the INSTRUMENT vary between verb types. For example, Dixon (1991: 102-13) identifies eight types of affectedness: a range including the minimal contact of the verb *touch* in 6.72, where possibly no change occurs in the PATIENT, through *rub* in 6.73, where the surface of the PATIENT might be affected, and *squeeze* in 6.74 where a temporary change of shape in the PATIENT occurs, to *smash* in 6.75, where the PATIENT loses its physical integrity:

6.72 John touched the lamp with his toe.

6.73 The captain rubbed the cricket ball with dirt.

6.74 Henry squeezed the rubber duck in his hands.

6.75 Alison smashed the ice cube with her heel.

The questions which face semanticists here are: do the differences between the affectedness of the PATIENT reduce the usefulness of this label, or can the differences be explained in some way?

The second problem is more general: how do we define theta-roles in general? That is, what semantic basis do we have for characterizing roles? Facing both of these problems, Dowty (1991) proposes a solution where theta-roles are not semantic primitives but are defined in terms of **entailments** of the predicate. In this view a theta-role is a cluster of entailments about an argument position which are shared by some verbs. He gives examples like *x murders y*, *x nominates y*, *x interrogates y*, where:

6.76 entailments they all share include that *x* does a volitional act, that *x* moreover intends this to be the kind of act named by the verb, that *x* causes some event to take place involving *y* (*y* dies, *y* acquires

a nomination,  $y$  answers questions - or at least hears them), and that  $x$  moves or changes externally (i.e. not just mentally). (1991: 552)

Such a set of shared entailments about  $x$  will serve to define the nominal which denotes  $x$  as AGENT. Thus theta-roles are defined in terms of shared verbal entailments about nominal referents.<sup>5</sup> We will see something of how these entailments are used in this approach in the rest of this section.

In this view of theta-roles as clusters of entailments, we can see a solution to the problem of the fuzziness of roles. Dowty proposes that we view  $\pm e$  roles not as discrete and bounded categories but instead as **prototypes**, where there may be different degrees of membership. He suggests that there are two basic prototypes: Proto-Agent and Proto-Patient,<sup>6</sup> each of which would contain characteristic lists of entailments such as those in 6.77 and 6.78 below:

- 6.77 Properties of the Agent Proto-Role (Dowty 1991: 572):
- volitional involvement in the event or state
  - sentience (and/or perception)
  - causing an event or change of state in another participant
  - movement (relative to the position of another participant)
- 6.78 Properties of the Patient Proto-Role (Dowty 1991: 572):
- undergoes change of state
  - incremental theme<sup>7</sup>
  - causally affected by another participant
  - stationary relative to movement of another participant

The idea is that these clusters of entailments would allow various kinds of shading. For example some arguments might have more of the entailments than others. So, for example, *John* in *John cleaned the house* would include all four of the entailments in 6.77 above: volition, sentience, causation and movement. By contrast *John* as an argument of *drop* in *John fainted and dropped the vase* would involve no volition, and *the storm* in *The storm destroyed the house* would involve neither sentience nor volition. We can see that this approach allows variation amongst AGENTS: some will be more typical and involve a greater number of characteristic entailments; others will be more marginal. Similar variation would hold for PATIENTS.

This approach would also allow other forms of fuzziness: some entailments might be viewed as more important than others; or each entailment itself might be fuzzy-edged. As several commentators have pointed out, speakers sometimes blur the distinction between sentient and non-sentient when they talk about computers, saying things like *The computer thinks these are the same file* or *This program doesn't realize that the memory is full*.

These proposals by Dowty to view thematic roles in terms of prototypical clusters of entailments allow flexibility in defining thematic roles. One result of

of his classification is that traditional role types fall out as more-or-less prototypical versions of the two main categories. Thus, as we have seen, a centrally prototypical AGENT like *Maggie* in 6.79a below involves all four entailments in 6.77, while an EXPERIENCER, like *Joan* in 6.79b can be seen as a more marginal AGENT, including sentience but not volition or causation; and an INSTRUMENT like *the scalpel* in 6.79c includes causation and movement but not volition or sentience:

- 6.79
- Maggie** pruned the roses.
  - Joan** felt the heat as the aircraft door opened.
  - The scalpel** cut through the muscle.

Similarly a centrally prototypical PATIENT, like *the roses*, in 6.79a and repeated in 6.80a below, will involve all four entailments in 6.78 above, but a STIMULUS like *the game* in 6.80b does not undergo a change of state nor is causally affected:

- 6.80
- Maggie pruned **the roses**.
  - Roberto watched **the game**.

Having seen something of an attempt to cope with the problem of defining theta-roles on a more systematic basis, in the next section we examine some of the uses of such roles.

## 6.6 The Motivation for Identifying Thematic Roles

From our discussion so far it is clear that linguists employ thematic roles to describe aspects of the interface between semantics and syntax, in particular to characterize the links between the semantic classification of its participants that is inherent in a verb's meaning and the grammatical relations it supports. Thus, to recap our discussion in its simplest terms, when we use an English verb like *feel* in *Joan felt the heat as soon as the aircraft door was opened*, we identify a relationship between an EXPERIENCER and a PERCEPT. This can be viewed as one of many conventional ways of viewing relations that is coded in the language. Grammatically of course the verb *feel* is transitive, taking a subject and direct object. As we have seen, one fact we have to account for is that there is a conventional linkage between the participant roles and the grammatical relations, such that in this case the EXPERIENCER will be subject and the PERCEPT, direct object.<sup>8</sup>

Predicting such linkages, and more general patterns amongst individual cases, is one of the primary functions of thematic roles. To take one example, in Dowty's prototype and entailments approach described in the last section, this linkage is described as below by an argument selection



- b. V: <AGENT, INSTRUMENT & THEME?2 LOCATION>  
 NP NP PP  
 V: <AGENT, LOCATION, INSTRUMENT & THEME>  
 NP NP PP

c. *bang, bash, beat, hit, knock, pound, rap, tap, whack*<sup>13</sup>

6.87 a. The whole community will benefit from the peace process.

The peace process will benefit the whole community.

- b. V: <BENEFICIARY, SOURCE>  
 NP PP  
 V: <SOURCE, BENEFICIARY>  
 NP NP

c. *benefit, profit*<sup>TM</sup>

These alternations are just two of a large range identified for English in Levin (1993). The conditional factors for such alternations are often a mix of semantic information, such as the verb's meaning and its theta-grid (as shown above), and its syntactic environment.

We can look at one further type of justification for thematic roles which comes from another area of grammar: the claim that in some languages they play a role in the morphology of verbal agreement. Mithun (1991: 514) gives examples of the pronominal verbal prefixes in Lakhota (Siouan; USA, Canada). In the transitive verbs in 6.88a below we see a prefix *wa* which marks an AGENT argument and in 6.88b a prefix *ma*, which marks a PATIENT:

- 6.88 a. *awá?u* 'I brought it.'  
*waktékte* 'I'll kill him.'  
 b. *amá?u* 'He brought me.'  
*maktekte* 'He'll kill me.'

We can see that these prefixes do not mark subject or object agreement because a subject, for example, can take either prefix depending on whether it is an AGENT (as in 6.89a below) or PATIENT (as in 6.89b) (Mithun 1991: 514):

- 6.89 a. AGENT subjects  
*wapsica* 'T jumped'  
*wahi* 'I came'  
 b. PATIENT subjects  
*makhua* 'I'm sick'  
*maxma* 'I'm sleepy'

In other words, what would be a subject pronoun in English corresponds to either an AGENT or PATIENT pronoun affix in Lakhota. Thus Lakhota morphological marking is sensitive to theta-roles rather than grammatical relations. Mithun gives similar examples from Guarani (Tupij Paraguay, Bolivia) and the Pomo-

an languages of California. The implication for our discussion is clear: if we need theta-roles to explain morphological patterns, this is strong evidence that they are significant semantic categories.

We have seen then in this section a number of different motivations for identifying thematic roles: to explain linking rules in verbal argument structure, to reflect semantic classes of verbs, to predict a verb's participation in argument structure alternations, and finally to describe morphological rules adequately. For many linguists this utility motivates their continuing use, despite the definitional problems discussed in the last section. In the next section we look at the category of **voice**, which, as we shall see, adds new dimensions to the relationship between theta-roles and grammatical relations.

## 6.7 Voice

### 6.7.1 Passive voice

The grammatical category of *voice* affords speakers some flexibility in viewing thematic roles. Many languages allow an opposition between *active voice* and *passive voice*. We can compare for example the English sentences in 6.90 below:

- 6.90 a. Billy groomed the horses.  
 b. The horses were groomed by Billy.

In the active sentence 6.90a *Billy*, the AGENT, is subject and *the horses*, the PATIENT, is object. The passive version 6.90b, however, has the PATIENT as subject and the AGENT occurring in a prepositional phrase, the structure often associated with INSTRUMENT, as we saw in the last section. This is a typical active-passive voice alternation: the passive sentence has a verb in a different form - the past participle with the auxiliary verb *be* - and it allows the speaker a different perspective on the situation described. This passive sentence (6.90b) allows the speaker to describe the situation from the point of view of the PATIENT rather than that of the AGENT. In some cases indeed passive constructions are used to obscure the identity of an AGENT, as in 6.91 below:

- 6.91 The horses were groomed.

Here the AGENT is so far backgrounded that it becomes merely an implied participant. Many writers describe this foregrounding of the PATIENT and backgrounding of the AGENT in terms of promoting the PATIENT and demoting the AGENT (for example Givon 1990) or as reflecting the speaker's greater empathy with the PATIENT rather than the AGENT (Kuno 1987). There are other lexical and syntactic strategies which alter perspective in this way. For

example in 6.92 below the alternation relies in part on the lexical relation between *in front of* and *behind* while in 6.93 it is accomplished by the syntactic patterns known as **pseudo-cleft** in a and **cleft** in b:

- 6.92 a. The house stood in front of the cliff.  
b. The cliff stood behind the house.
- 6.93 a. What Joan bought was a Ferrari.  
b. It was Joan who bought the Ferrari.

In 6.93 above the same situation is described but in a the speaker is interested in Joan's purchase, while in b she is interested in the Ferrari's purchaser. This kind of choice of perspective presumably depends on a speaker's judgements of conversational salience. We can use the terms **figure** and **ground**<sup>15</sup> to describe this kind of linguistic perspective: if we call the situation described a **scene**, then the entity that the speaker chooses to foreground is the figure, and the background is the ground. So in 6.92a above *the house* is the figure and *the cliff* the ground, and vice versa in 6.92b.

Passive constructions allow the foregrounding of roles other than PATIENT. In 6.94-6 we see English examples of THEME, PERCEPT, and RECIPIENT roles occurring as the subject of passives:

- 6.94 **This money** was donated to the school, (THEME)
- 6.95 **The UFO** was seen by just two people, (PERCEPT)
- 6.96 **He** was given a camera by his grandmother, (RECIPIENT)

The qualifications for foregrounding in a passive in English are complex: partly grammatical, partly semantic and partly due to the flow of discourse and the speaker's choice of viewpoint. The importance of grammatical information can be shown by observing that each of the roles occurring as passive subjects in 6.94-6 above occur in object position in a corresponding active sentence:

- 6.97 Someone donated **this money** to the school.
- 6.98 Just two people saw **the UFO**.
- 6.99 His grandmother gave **him** a camera.

The typical pattern is that a nominal occupying object position is fronted to subject in passives. When a theta-role normally occurs as a prepositional phrase in an active sentence, this is less likely to be foregrounded in a passive. Neither moving the full prepositional phrase nor extracting just the nominal seems to work, as shown below:

- 6.100 a. This house stood **on the corner**, (LOCATION)  
b. **\*On the corner** was stood by this house.  
c. **?The corner** was stood **on** by this house.
- 6.101 a. John built a garage for **her**. (BENEFICIARY)  
b. **\*For her** was built a garage by John.  
c. **?She** was built a garage by John.
- 6.102 a. He opened the door **with this key**. (INSTRUMENT)  
b. **\*With this key** was opened the door by him.  
c. **\*This key** was opened the door **with**.

Some apparent exceptions to this rule are possible however, e.g.

- 6.103 a. Three monarchs lived in **this house**, (LOCATION)  
b. **This house** was lived in by three monarchs.<sup>16</sup>

To further underline this grammatical aspect of passives, i.e. that it is the object position that is relevant to passivization, we can look at a class of English verbs called the *spray/load* verbs. These verbs allow the speaker to select either their THEME role (as in 6.104a and 6.105a) below, or the GOAL (as in 6.104b and 6.105b), to be the verb's direct object and thus be the focus of the effect of the action:

- 6.104 a. He sprayed paint on the car.  
b. He sprayed the car with paint.
- 6.105 a. He loaded hay on to the tractor.  
b. He loaded the tractor with hay.

We can easily show that whichever argument occupies object position can be passivized while the argument in the prepositional phrase cannot: corresponding to 6.104 above we find the patterns:

- 6.106 a. Paint was sprayed on the car.  
b. **\*The car** was sprayed paint on.  
c. The car was sprayed with paint.  
d. **\*Paint** was sprayed the car with.

See Rappaport and Levin (1985, 1988), Jeffries and Willis (1984) and Levin (1993) for further discussion of these *spray/load* verbs.<sup>17</sup>

The discourse factors affecting passives have been described in a number of frameworks: for example, as mentioned above, Kuno (1987: 209-16) employs the notion of speaker empathy. He gives an example of a person relating a story about their friend Mary and her experiences at a party. In the narrative the speaker's empathy is with Mary and thus events are viewed

from her perspective. This explains why a passive is fine in 6.107b below but not in 6.108b (treating these as two independent reports of events):

- 6.107 Mary had quite an experience at the party she went to last night.  
 a. An eight-foot-tall rowdy harassed her.  
 b. She was harassed by an eight-foot-tall rowdy.
- 6.108 Mary had quite an experience at the party she went to last night.  
 a. She slapped an eight-foot-tall rowdy in the face.  
 b. \*An eight-foot-tall rowdy was slapped in the face by her.

The passive construction works in 6.107b because the fronted nominal refers to the entity the speaker empathizes with, but not in 6.108b where the other participant is fronted.

Passive constructions have received a great deal of attention in the linguistics literature. This is not surprising: even from our brief discussion, we can see that while the general effect of passive is to allow a shift in linkage between theta-roles and grammatical relations, the process is subject to a complex of grammatical and discourse factors. It is this interdependence of different levels of analysis that makes passives an interesting arena for theoretical debate.

### 6.7.2 Comparing passive constructions across languages

While many languages have passive-type constructions, the comparison of passives across languages reveals that there is considerable variation around the pattern of the English passive outlined in the last section, i.e. where the AGENT is demoted from subject position, a non-AGENT role is promoted to subject, and the verb shows a distinct form which agrees with the promoted subject: the total package being what we have called **passive voice**. Often languages have more than one passive construction: in English for example, it is possible to distinguish between *be*-passives and *get*-passives, as in 6.109 (R. Lakoff 1971, Givon and Yang 1994):

- 6.109 a. Mary was shot on purpose.  
 b. Mary got shot on purpose.

As noted by Lakoff these sentences differ in the amount of control over the event associated with Mary.<sup>18</sup>

Other languages have a special type of passive, often called the **impersonal passive**, which does not allow the AGENT to be mentioned in the sentence. In Irish, for example, we can distinguish between one type of passive associated with verbal noun constructions as shown in the active/passive pair in 6.110 below, and another, the impersonal passive, with verbs, as is shown in 6.111 (Noonan 1994: 282-6):

- 6.110 a. Bhi si ag bualadh Sheain.  
 was she at hit-NOMIN John-GEN  
 'She was hitting John.'  
 b. Bhi Sean a bhualadh aici.  
 was John to+his hit-NOMiN at-her  
 'John was being hit by her.'
- 6.111 a. Thug siad Siobhan abhaile inniu.  
 brought they Joan home today  
 'They brought Joan home today.'  
 b. Tugadh Siobhan abhaile inniu.  
 brought- IMPERS Joan home today.  
 'Joan was brought home today.'

This impersonal passive in 6.111 does not straightforwardly correspond to the translation given: i.e. to an English passive where no AGENT is expressed. In 6.111b we can see how both in Irish and in the English translation the passive verb form is differentiated from the active, and how in both the AGENT is often omitted. However the Irish passive in 6.111b differs from its English translation because the THEME, *Siobhan*, remains in its original position as an object while in the English passive *Joan* becomes subject. In other words, the PATIENT is not promoted to subject in the Irish impersonal passive in 6.111b, but the AGENT is omitted. See Noonan (1994) for discussion.

This example from Irish is of a transitive impersonal passive. In many languages the term *impersonal passive* is used to describe passives of intransitive verbs: Kirsner (1976:387) gives the following pair of examples from Dutch:

- 6.112 a. De jongens fluiten.  
 the boys whistle.  
 'The boys whistle/are whistling.'  
 b. Er wordt door de jongens gefloten.  
 there becomes by the boys whistling  
 'By the boys (there) is whistling.'

In 6.112b the AGENT is backgrounded, but there is no other argument to be foregrounded and subject position is taken by the word *er* 'there', which does not refer directly to any entity and which has no theta-role. It is also possible to delete the AGENT altogether in this passive, giving:

- 6.113 Er wordt gefloten.  
 there becomes whistling  
 'There is whistling/People whistle/Someone whistles.'

Similar impersonal passives have been reported for other languages, including German, Welsh and Latin; see Perlmutter (1978) and Perlmutter and Postal (1984) for discussion.

These impersonal passives imply that in comparing languages we need to separate out the two functions of the passive: firstly, the demotion of agents, and secondly, the promotion of non-AGENTS. Thus an English passive like *Spike was arrested by the police* combines both functions: the agent argument is demoted to a prepositional phrase, and the patient is promoted to subject. We can see the related sentence *Spike was arrested* as a special case of this, where demotion reaches its extreme in the suppression of the agent. In the Dutch impersonal passives in 6.112 b on the other hand we see a passive strategy which just embodies the first function: demotion of agent, with no concomitant promotion function. Since this example has an intransitive verb, the further step of suppressing the agent leaves a sentence with no theta-role bearing nominal as in 6.113.

The third characteristic of English passives described in the last section was a special verb form and associated verbal agreement with the promoted subject. This too is subject to cross-linguistic variation. Passive verbs are often semantically distinguished from their active counterparts, for example by being more stative, though this is not always so, and they may show agreement with the promoted non-AGENT nominal (as in English), or the demoted agent, or neither, since agreement inflections may be neutralized; see Givon (1990: 563-644) for discussion of variations along this parameter as well as along the parameters of AGENT demotion and non-AGENT promotion.

One conclusion from comparing passives across languages seems to be that the phenomenon is typically a cluster of functions: in each case following the general pattern of allowing the speaker planning her discourse some variation in the linkage between thematic and grammatical roles, but with considerable variation in the associated semantic and grammatical elements of the cluster. In most active-passive systems the active form is usually grammatically simpler and we may ask why this should be so. It has been argued that we as humans naturally view situations from the point of view of any human beings involved, and if there are none, of other living creatures. This preference, sometimes called an **animacy hierarchy** (see for example Dixon 1979, Hopper and Thompson 1980), is coded into the lexical semantics of a language so that a verb like *drive*, for example, in 6.114 sets up a thematic role frame which requires an agent as the subject:

6.114 Ann drove the truck across the field.

and since agency, as we have seen, requires *wilful* action, AGENTS are typically people, or higher animals. It is difficult to think of a verb which describes the action in 6.114 from the point of view of the truck. We might say:

6.115 The truck carried Ann across the field.

but this sentence has a different meaning: we have not specified that Ann was driving. So it seems that the meaning of the verb *drive* is set up to

prioritize the role of any human or volitional agent. Passive voice allows the speaker to get around this in-built bias, so that to switch the viewpoint from Ann to the truck, or to the field, she can use passive constructions as in 6.116-17:

6.116 The truck was driven across the field by Ann.

6.117 The field was driven across by a truck (\*by Ann).

We can see that in 6.117 there is no longer a slot for the AGENT, Ann. So passive constructions do allow a change of perspective but the conventional bias towards animate subjects means that the active *drive* is grammatically simpler than the passive *was driven*.

### 6.7.3 Middle voice

While very many languages display this active/passive voice contrast, some languages have a three-way distinction between active, passive and **middle voice**. As we might expect, the use of middle voice varies from language to language but a central feature is that middle forms emphasize that the subject of the verb is affected by the action described by the verb. This **affect- edness**, as it is often termed (e.g. Klaiman 1991), can be of several types, and we can select four typical uses as examples: neuters, bodily activity and emotions, reflexives, and autobenefactives. Though we will use examples from several languages, to keep the discussion brief we will concentrate on two unrelated languages, well separated in space and time: classical Greek and the modern Cushitic language Somali.<sup>19</sup> In both these languages middle voice is marked by verbal inflection.

#### *Neuter intransitives*

This type of middle is where the subject undergoes a non-volitional process or change of state. The external cause is not represented but can often be shown in a related active form, as shown in 6.118 below, an example from Sanskrit (Klaiman 1991: 93):

- 6.118 a. So namati dandam  
he-NOM bends-3sg ACTIVE stick-ACC  
'He bends the stick?'
- b. Namate dandah.  
bends-3sg MIDDLE stick-  
NOM 'The stick bends.'

Middle voice verb forms of this neuter type, where the subject undergoes a process over which it has no control, occur in classical Greek, as shown in 6.119 (Bakker 1994: 30) and Somali,<sup>20</sup> as in 6.120:

6.119	<b>phu-e-sthai</b>	'grow'
	<b>treph-e-sthai</b>	'grow up'
	<b>sep-e-sthai</b>	'rot'
	<b>tek-e-sthai</b>	'melt'
	<b>rhegnu-sthai</b>	'break'
6.120	kab-o	'recover, set (of a bone)'
	qub-o	'fall (of leaves and fruit)'
	dhim-o	'die'
	haf-o	'drown'
	garaads-o	'reach maturity'

### *Bodily activity and emotion*

In some languages the verb occurs in a middle voice when the activity involves the body or emotions of the subject. These would seem to be clear cases of affectedness since the subject is so overtly involved. Examples of such middle voice verbs are in 6.121-2:

6.121	Classical Greek (Bakker 1994)
	<b>klin-e-sthai</b> 'lean'
	<b>hed-e-sthai</b> 'rejoice'
6.122	Somali (Saeed 1999)
	fadhiis-o 'sit down'
	baroor-o 'mourn, wail'

### *Reflexives*

In some languages the middle is used where the subject's action affects the subject himself, or a possession or body part of the subject. To take another example from classical Greek (Barber 1975: 18-19):

6.123	Lou-omai.
	wash 1sg MIDDLE
	'I wash myself.'

This use means that in many languages verbs of grooming occur in the middle voice, with no need for a reflexive pronoun as object; see 6.124 for some further examples from Somali, and examples from other languages in 6.125 from Kemmer (1994: 195):

6.124	feer-o	'comb one's hair'
	maydh-o	'wash oneself, bathe'
	labbis-o	'dress up, put on one's best clothes'

6.125	Latin	orno-r	'adorn oneself'
	Quechua	arma-ku-y	'bathe'
	Turkish	giy-in	'dress'
	Hungarian	mosa-kod-	'wash oneself'

### *A utobenefactives*

This type of middle is used to signify that the action of the subject is done for his or her own benefit. Once again this use occurred in classical Greek as in 6.126 (Barber 1975: 18), and in a regular process in Somali, as 6.127 shows (Saeed 1993: 58):

6.126 a.	hair-o	moirah.
	take-1sg-ACTIVE	
	share	'I take a share.'
b.	hari-oumai	moiran.
	take-1sg-MIDDLE	share
		'I take a share for myself.'

#### Active verbs:

wad	'to drive'
beer	'to cultivate'
qaad	'to take'
sid	'to carry'

#### Middle verbs:

wad-o	'to drive for oneself'
beer-o	'to cultivate for oneself'
qaad-o	'to take for oneself'
sid-o	'to carry for oneself'

In the examples so far, middle voice has been marked by verbal inflection. In some languages a pronoun marks middle forms, often the same form as a reflexive pronoun, e.g. German *sich*, French *se*, Spanish *se*, or a closely related form, e.g. Russian reflexive *sebjä*, middle *-sja*, Dutch reflexive *zichzelf* middle *-zelf* (Kemmer 1994). In such languages the overlap between middle voice and reflexivity, seen in examples 6.121-7 above, becomes overt. In French and Spanish for example, we might identify our first three types of middle:

### 6.128 French middle reflexives

a. neuter:	s'ecrouler	'collapse'
	s'evanouir	'vanish'
b. bodily activity:	s'asseoir	'sit down'
emotion:	se plaindre	'complain'
c. reflexive:	s'habiller	'dress oneself'
	se peigner	'comb one's hair'

### 6.129 Spanish middle reflexives

a. neuter:	helarse	'freeze (intr.)'
	recuperarse	'get well'
b. bodily activity:	tirarse	'jump'
emotion:	enamorarse (de)	'fall in love (with)'

c. reflexive:	afeitarse	'shave'
	quitarse	'take off (clothes)'

However, even in languages where the middle and reflexives are marked by the same pronoun, there are usually clear cases where the meaning distinguishes between true reflexives and the middle, e.g. in German (Kemmer 1994: 188):

6.130	Er sieht <b>sich</b>	'He sees himself'	(Reflexive)
	Er furchtet <b>sich</b>	'He is afraid'	(Middle - emotion)

In English there is no inflectional or pronominal marker of the middle: the distinction is only shown by alternations between transitive active verbs and intransitive middle verbs, where the agent is omitted, e.g.

6.131	a. They open the gates very smoothly. (Active)	
	b. The gates open very smoothly. (Middle - neuter)	

These intransitive middles in English are often used to describe the success of a non-AGENT in some activity, e.g.

6.132	a. These clothes wash well.
	b. This model sells very quickly.
	c. These saws don't cut very efficiently.

See Dixon (1991: 322-35) for more examples of this type of construction in English. Because of the similar suppression of the AGENT in this type of middle and in the passive, some writers use the term **medio-passive** to cover both.

## 6.8 Classifiers and Noun Classes

So far in this chapter we have been exploring the ways that participants may be assigned semantic roles relative to the action or situation described by the verb. In this section we look at semantic characterizations that are based on inherent properties of the entities referred to by noun phrases. Many languages have overt systems for marking how referents fit into a semantic classification system. We divide our brief discussion of these into first, classifiers, and then, noun classes.

### 6.8.1 Classifiers

Noun classifiers are morphemes or lexical words that code characteristics of the referent of the noun, allowing the speaker to classify the referent

according to a system of semantic/conceptual categories. They may show up grammatically in different guises. Some, termed **noun classifiers**, occur with nouns. Dixon (1977) describes the noun classifiers of the Australian language Yidiji as a closed set of around 20 members, which he divides into two general types, each containing several subtypes. The first type, inherent nature classifiers, includes as subtypes of classifiers: human; animals; vegetation; natural objects (like the classifier *walba* 'stone'); and artefacts (like the classifier *baji* 'canoe'). The second type, functional classifiers, divides entities into: meat food; non-meat food; drinkable things; movable; habitable; and 'purposeful noise'. Dixon (1982) reports that two classifiers can be used with the same nominal as long as they come from the two different general types, for example (where CL = classifier):

6.133	bulumba walba malan
	CL:HABITABLE CL:STONE flat.rock
	'a flat rock for camping'
	(Dixon 1982: 200)

In many languages classifiers occur in specific grammatical constructions or locations, for example **numeral classifiers**, which occur when the entity is being counted, and **possessive classifiers**, which occur in constructions describing possession. Numeral classifiers occur in Japanese as in shown in example 6.134 below:

6.134	Classifiers in a Japanese shopping list (cited in Aikhenvald 2000: 2)			
	<b>Shopping list</b>	<b>Numeral</b>	<b>Classifier</b>	<b>Meaning of classifier</b>
	<i>nasu</i> (eggplant)	<i>nana</i> (7)	<i>-ko</i>	CL: SMALL, EQUIDIMENSIONAL
	<i>kyuuri</i> (cucumber)	<i>hachi</i> (8)	<i>-hon</i>	CL: ELONGATED
	<i>hamu</i> (ham)	<i>juu</i> (10)	<i>-mai</i>	CL: SHEETLIKE

As we can see, these classifiers relate to a classification based on shape.

Possessive (or genitive) classifiers may characterize the possessed item, as in the Fijian example in 6.135 below; or classify the type of possession relation involved, as in 6.136 from Hawaiian:

6.135	Fijian possessive classifiers (Uchtenberk 1983: 157-8)	
	a.	na me-qu yaqona
		ART CL:DRINKABLE-my kava
		'my kava (which I intend to drink)'
	b.	na no-qu yaqona
		ART CL:GENERAL-my kava
		'my kava (that I grew, or that I will sell)'

- 6.136 Hawaiian possessive classifiers (Lichtenberk 1983: 163)
- a. k-o-'u inoa  
ART-CL-my name  
'my name (that represents me)'
  - b. k-a-'u inoa  
ART-CL-my name  
'my name (that I bestow on someone)'

A further type is **verbal classifiers**, where the classifier occurs as a morpheme attached to the verb and serves to classify (intransitive) subjects or objects: see for example:

- 6.137 Dogrib (Athapaskan) (cited in Allen 2001: 309)
- a. let'e niyeh-tsi  
bread I.pick.up-PERF. CL: FLAT, FLEXIBLE, ENTITY  
'I pick up a slice of bread'
  - b. let'e niyeh-?a  
bread I.pick.up-PERF. CL:ROUND.ENTITY  
'I pick up a loaf of bread'

Wherever they are marked grammatically classifiers tend to exploit a fixed set of semantic distinctions. Though there is large variation, it is possible to identify some prototypical distinctions, as Allan (2001) does below:

- 6.138 Prototypical classifier categories (Allan 2001: 307)
- a. *Material make-up*: e.g. human (-like), animate, female, tree (-like)
  - b. *Function*: e.g. piercing, cutting, or writing instruments; for eating, drinking
  - c. *Shape*: e.g. long (saliently one-dimensional), flat, round
  - d. *Consistency*: e.g. rigid, flexible, mass
  - e. *Size*: including diminutives and augmentatives
  - f. *Location*: inherently locative entities such as towns
  - g. *Arrangement*: e.g. a row of, a coil of, a heap of
  - h. *Quanta*: e.g. head of cattle, pack of cigarettes

## 6.8.2 Noun classes

Noun classes are agreement-based noun systems that seem, at least historically, to be based on semantic classifications somewhat similar to those we have seen for classifiers. One famous example occurs in the Bantu languages of Africa, where nouns belong to a pattern of classes, related variously in the modern languages to an ancestral system that is characterized by

Aikhenvald (2000) as follows, (where class pairs 1/2 etc. are singular and plural):

- 6.139 Noun classes in Proto-Bantu (Aikhenvald 2000: 282)
- | Class | Semantics  |
|-------|--|
| 1/2   | Humans, a few other animates                                     |
| 3/4   | Plants, plant parts, foods, non-paired body parts, miscellaneous |
| 5/6   | Fruits, paired body parts, miscellaneous inanimates              |
| 7/8   | Miscellaneous inanimates   |
| 9/10  | Animals  |
| 11/10 | Long objects, abstract entities                                  |
| 6     | Small objects, birds   |
| 14    | Masses   |
| 15    | Infinitives  |

The key feature of noun class systems is that other elements in the sentence agree with the noun in terms of its class. See for example (6.140) below from the modern Bantu language Swahili:

- 6.140 Swahili class 8 (Allan 2001: 310):
- Vi-su vidogo viwili hi-vi amba-vy-o nili-vi-nunua ni  
knife vz'-small vz-two this-z>i which-vz l.s-w-buy be  
vi-kali sana  
t>z-sharp very  
'These two small knives which I bought are very sharp'

Here the noun class prefix, marked in bold, is copied as an agreement feature by other elements in the noun phrase headed by *visu* 'knife' and in the sentence in which the noun phrase is subject.

In the modern Bantu languages the assignment of nouns to classes is not always as semantically transparent as the classes in 6.139 suggest. Often the classes are much more heterogeneous and membership may be more conventionalized.

Gender systems, familiar from Indo-European languages, in which nouns are assigned to two or three classes - male, female and perhaps neuter - are a type of noun class system. Indeed Corbett (1991) extends the term *gender* to cover all noun class systems. As may be the case with more complex noun class systems, gender in languages like German or Hindi is a grammatical distinction only loosely connected to biological sex. Humans and animals may be typically (though not exclusively) assigned to genders on the basis of biological sex, but other nouns are assigned by a mixture of criteria, some of which have no semantic basis, for example phonological shape.

Noun class systems may be differentiated from classifiers by a number of features, some of which are summarized by Dixon (1986) as follows:

Differences between noun classes and classifiers (Dixon (1986))

	<b>Noun classes</b>	<b>Classifiers</b>
Size	Small finite set	Large number
Realization	Closed grammatical system	Free forms
Scope	Marking is never entirely within the noun word	Never any reference outside the noun phrase

However the large degree of variation within both types of system means that any simple characterization is only suggestive of typical cases.

## 6.9 Summary

In this chapter our main focus has been on the ways in which a speaker may portray the roles of participants in a situation. We outlined a classification of such semantic roles, termed **thematic roles** or **theta-roles**, including AGENT, PATIENT, THEME, etc. and described the relationship between these roles and grammatical relations like **subject** and **object**. It has been claimed that as part of its inherent lexical specification a verb requires its arguments to be in specific thematic roles, and that this can be reflected by formulating thematic role **grids**, or **theta-grids**. We discussed the difficulties there are in fixing tight definitions for individual thematic roles, and presented one approach, from Dowty (1991), which seeks to provide a solution in terms of fuzzy categories. This difficulty with precision notwithstanding, it seems that the notion of thematic roles has proved a useful descriptive tool in a number of areas of the semantics-grammar interface.

The grammatical category of **voice** allows speakers different strategies for relating thematic roles and grammatical relations. We concentrated on relations with **subject** position, in particular the way in which **passive voice** allows the foregrounding of non-AGENT roles to subject and the backgrounding of AGENT roles away from subject. We also looked at **middle voice**, which reflects the **affectedness** of the subject in the action of the verb: thus offering a different view of the relationship between subject and verb from the **active voice**.

Finally we looked at **classifiers** and **noun classes**: systems where nouns identifying entities are classified by inherent semantic features, though membership of the relevant classes may only be partially semantically determined.

## FURTHER READING

An important study of thematic roles is Dowty's (1991) article. Palmer (1994) is a survey of thematic roles, the different ways they are grammaticalized and the role

of passive and middle voice. Dixon (1991) discusses the ways in which the grammar of English verbs reflects semantic distinctions, and includes sections on thematic roles, and voice. Levin and Rappaport Hovav (2005) provide further discussion of the problems with thematic roles identified in this chapter. Givon (1994) is a collection of studies on argument structure changing processes, including passive. Keenan (1985) reviews passive constructions in a range of languages, while Klaiman (1991) does a similar job for middle voice. Wilkins (1988), Grimshaw (1990) and Williams (1994) shed light on the interaction of thematic roles and grammatical processes. These works are quite technical, however, and require some background in syntactic theory. Aikhenvald (2000) provides a comprehensive cross-linguistic overview of classifier systems; and Corbett (1991) discusses norm class systems.

## EXERCISES

- 6.1 On the basis of the informal definitions in section 6.2, try to assign a single **thematic role** label to each of the expressions in bold in the following sentences:
- Helen** drove to the party.
  - He** swatted the fly with a newspaper.
  - The baboon** was asleep on the roof of my car.
  - Joan** drank the yard of ale.
  - Campbell** saw the gun first.
  - George** gave the doorman a tip.
- 6.2 For each of the **theta-roles** below, construct an English sentence where an argument bearing that role occurs as **subject**. Use simple active sentences, avoiding for the present exercise passive constructions and complex sentences.
- EXPERIENCER
  - PATIENT
  - THEME
  - INSTRUMENT
  - RECIPIENT
- 6.3 For each of the **theta-roles** below, construct an English sentence where an argument bearing that role occurs as **object**.
- PATIENT
  - THEME
  - BENEFICIARY
  - RECIPIENT
  - STIMULUS

6.4 As we saw, Jackendoff (1990) proposes a distinction between a **thematic tier** of thematic roles (relating to movement and location) and an **action tier** (relating to ACTOR-PATIENT type relations). An argument may have a role at each level and thus fulfil two roles. For example the underlined argument in *The car smashed into the shop window* can be analysed as both PATIENT and GOAL. For each of the combinations of roles below, try to invent a sentence where a single argument fulfils the combination:

- a. AGENT and GOAL
- b. PATIENT and THEME
- c. AGENT and SOURCE
- d. AGENT and THEME

6.5 In sections 6.5 and 6.6 we discussed proposals from Dowty (1991) to characterize thematic roles in terms of **clusters of entailments**, and to describe the rules linking thematic roles and grammatical relations like subject and object in terms of **argument selection principles**. Using the selection principles in 6.81 in the chapter and the properties of Proto-roles in 6.77 and 6.78, discuss the selection of subject and object positions in the following sentences:

- a. The butler is polishing the silverware.
- b. The dogs will smell the food.
- c. The train hit the cow.

What problems are posed for these principles by the selection of subject and objects in the pairs of sentences below?

- 1
  - a. He fears AIDS.
  - b. AIDS frightens him.
- 2
  - a. Patricia resembles Maura
  - b. Maura resembles Patricia.
- 3
  - a. Joan bought a sportscar from Jerry
  - b. Jerry sold a sportscar to Joan.

6.6 We saw how **passive** allows the foregrounding of non-AGENT theta-roles into subject position. Compare for example the active sentence 1 below with the passive equivalent in 2:

- 1 Craig<sub>AG</sub> devoured the ice-cream<sub>PA</sub>.
- 2 The ice-cream<sub>PA</sub> was devoured by Craig<sub>AG</sub>.

Assume 2 is formed from 1 by a simple rule: (a) Place the non-AGENT argument at the beginning of the sentence; (b) change the active verb to a passive verb (e.g. *devoured* → *was devoured*); (c) Place the word *by* in front of the AGENT and place the AGENT at the end of the sentence. Below are some active sentences with a non-subject argument underlined. For each one, use our simple rule to try to create a corresponding passive where the underlined non-AGENT argument becomes subject.

- a. The court fined Emma five hundred francs.
- b. Aliens abducted me in the middle of my examination.
- c. The professor mailed the answer to the student.
- d. The professor mailed the answer to the student.
- e. The professor mailed the student the answer.
- f. The professor mailed the student the answer.

Were any of the resulting passives ungrammatical? If so, what explanation can you give?

6.7 The rule of passive formation is not entirely regular. Using the rule in the last exercise as your model, try to create passives by foregrounding the underlined arguments below. What problems do you meet? Can you think of any semantic reasons to explain the results of the passive rule in these cases?

- a. A pleasant smell filled the room.
- b. Roy likes linguine in a clam sauce.
- c. Your paper included a nice conclusion.
- d. The fans crowded the hall.
- e. We were just watching television.
- f. Fritz hated war.
- g. He watched himself.

6.8 As we saw, in some languages (e.g. Somali) when a speaker describes a reflexive act of grooming, say for example the equivalent of *I wash myself* the verb occurs in a **middle voice** form with no object. In others (e.g. French) a **reflexive pronoun** is used as the object. In English we find another strategy: some verbs which are normally transitive allow the speaker to omit the object in order to convey a reflexive meaning. For example, we know that *hide* is normally a transitive verb because of sentences like *She hid the money*; however *She hid* means of course *She hid herself*. So English has verbs like *hide*, which by omitting an argument can convey an **understood reflexive object**. Unlike Somali though, the English verb in these construct-

-ions do not have a special middle voice ending. Below are some verbs which describe what we could call acts of grooming. Decide which of these allow an understood reflexive object.

undress	towel
wash	bathe
brush	shampoo
soap	shave
strip	lather

Is there any semantic differences between those verbs which allow this understood reflexive object and those which do not? If you think there is, test your hypothesis with other verbs from this semantic field of grooming.

6:9 Design **lexical theta-grids** for the verbs in bold in the sentences below. For example a theta-grid for *buy* in *Dee-dee **bought** the car for his mistress* would be: **buy** <AGENT, THEME, BENEFICIARY>.

- Brenda **reported** the incident to her boss.
- Frogs **fell** from the sky.
- Our headquarters will **remain** in London.
- Batman **received** a commendation from the mayor.
- Harvey **noticed** a strange smell.

## NOTES

- One might also think of examples like: *In the village stands a pump*. But here the subject still seems to be *a pump* rather than *in the village*<sup>^</sup> as can be shown by the pattern of agreement in: *In the village stand several pumps*. But see Levin and Rappaport Hovav (1995: 261-4) for arguments, couched in the theory of Lexical-Functional Grammar (e.g. Bresnan 1994), that the preverbal PP is, at some level of analysis, a subject.
- See the introductory discussion of theta-grids in Haegeman (1994: 33-73).
- Hereafter we will use just the two first letters of a thematic role with this subscript notation, e.g. Joan<sub>AG</sub> for Joan<sub>AGENT</sub>.
- In Jackendoff's (1990) two-tier representation described earlier, these 'transfer' verbs would have a more complicated thematic grid: we could, for example, assign both AGENT and SOURCE roles to *Barbara* in 6.67.
- Note that in this view, theta-roles convey a speaker's classifications of things in the world: in other words, the roles are borne by real-world entities rather than grammatical elements like NPs. See for example the following example and comment from Laduslaw and Dowty (1988: 63):

- Fido chased Felix.
  - Felix was chased by Fido.

. . . The only sense in which it is reasonable to think of the subject NP of (1a) as the Agent is the sense in which it is shorthand for saying that the object (in the world) referred to by the subject is the Agent in the action described by the sentence. What makes Fido an agent in the event described by (1a) and (1b) is information about Fido and his role in the event, not about the grammatical category or function of anything in the sentence.

- For a related idea, see Foley and van Valin's (1984) theory of **macro-roles**, where all thematic roles fall into two main categories: **actor** and **undergoer**.
- This term arises from Dowty's (1991) examination of different types of what he calls THEME roles, some of which would be PATIENT roles in our classification. He proposes a class of **incremental themes** for the THEME/PATIENT roles of achievement and accomplishment verbs, e.g. *mow the lawn eat an egg build a house demolish a building*. The observation is that the action (for example, the mowing action) and the state of the associated THEME/PATIENT (e.g. the lawn) are in a proportional relationship: some mowing cuts some of the grass, more mowing, more of the grass, etc. until completing the action cuts all of the grass. Dowty extends this idea of incremental themes to other types of role, e.g. *swim from England to France*, where the path is incrementally affected, and *memorize a poem*, where there is a similar incremental relationship between the action and a representation of the THEME entity. See Dowty (1991) for further details.
- In our discussion we focus on languages like English which have three grammatical relations, **subject** and **object**. We therefore leave aside the different pattern of mapping between theta-roles and grammatical relation shown by **ergative** languages. Briefly, in a typical ergative system one grammatical relation, called **absolutive**, is used for the single argument of an intransitive verb, whatever its theta-role (and in this resembles English subject), but is also used in ditransitive verbs for the PATIENT argument (and here resembles English object). A second grammatical relation, called **ergative**, is used for the AGENT/EXPERIENCER in ditransitive verbs (as is English subject). There is therefore no correspondence between the absolutive/ergative distinction and the subject/object distinction. They represent two different strategies for mapping between theta-roles and grammatical relations. See the following simple example of an ergative system from Tongan (Austronesian: Tonga), given by Anderson (1976):

- na'e lea 'a etalavou.  
PAST speak ABS young.man  
'The young man spoke.'
- na'e alu 'a Tevita ki Fisi.  
PAST go ABS David to Fiji  
'David went to Fiji.'
- na'e tamate'i 'a Kolaiate 'e Tevita.  
past kill ABS Goliath ERG David  
'David killed Goliath.'
- na'e ma'u 'e siale 'a e me'a'ofa.  
past receive ERG Charlie ABS DEF gift  
'Charlie received the gift.'

- Note that in these Tongan sentences the verb comes first in the sentence, and the case-marking particles (in bold) precede their nominals. Sentences a and b have intransitive verbs and the verb's only argument is in the absolutive case. Sentences c and d have transitive verbs. Here the AGENT in c and the RECIPIENT in d are in the ergative case. The PATIENT in c and the THEME in d are in the absolutive case. The reader may compare this with the mapping for subject - object languages like English. Ergative languages are found all over the world and include Basque in southern Europe, the Australian language Dyirbal, Tongan from the Pacific, and the Inuit languages of Canada, Greenland, etc. See Dixon (1979) for discussion and Croft (1990) and Palmer (1994) for cross-linguistic overviews.
- 9 Note that Dowty's hierarchy here has INSTRUMENT and PATIENT in reverse order to our earlier hierarchy. We won't try to arbitrate between these claims here: compare the discussion in Dowty (1991) and Croft (1990).
- 10 These are labels commonly used in the literature for the thematic roles associated with these verbs. We leave aside discussion of how these roles would correlate with the Agent-properties and Patient-properties in a Dowty-style approach.
- 11 See Grimshaw (1990) and Levin (1993) for discussion of these classes of psychological verbs.
- 12 Here we follow Jackendoff (1990) in allowing one argument to have two theta - roles, as described earlier.
- 13 See Dowty (1991: 594-5), Levin (1993: 67-8).
- 14 See Levin (1993: 83).
- 15 This is similar to the use of 'figure' and 'ground' in the analysis of motion verbs by Talmy (1975), and others, as discussed in chapter 9. There the *figure* is the entity in motion and the background is called the *ground*.
- 16 But only under some special conditions, which have been much debated in the literature. Levin and Rappaport Hovav (1995: 143-4), for example, discuss examples of this type like This platform has been stood on by an ex-president under the label prepositional passives. They provide a restriction on the construction in English that mixes grammatical and semantic factors: that it is only possible with unergative verbs which take an animate subject. Unergative is a term introduced by Perlmutter (1978) for intransitive verbs like sit and stand whose single argument is an agent and whose grammatical behaviour contrasts with unaccusative verbs which are intransitive verbs like grow or drown and whose single argument is essentially a patient. Dixon (1991: 298-321) on the other hand proposes syntactic restrictions, which include the absence of a direct object in the active sentence, and a lack of an alternative active construction in which the passivized NP could occur as direct object. For an in-depth study of these prepositional passive constructions see Couper-Kuhlen (1979).
- 17 Other English verbs allow alternations into object position, e.g.:
- 1 a. He wrapped cling-film around the food.  
b. He wrapped the food in cling-film.
  - 2 a. David gave the keys to Helen.  
b. David gave Helen the keys.
  - 3 a. She bought some flowers for her husband.  
b. She bought her husband some flowers.

Alternations like 2 and 3 are often called Dative Shift. Givon (1984a) describes these, and similar alternations in other languages, as promotion to object, a process paralleling passive. By comparison with passive, though, the process is more restricted to particular verbs and is less likely to be marked on the verb by a distinct inflection of voice.

18 Though this is less true of pairs like:

1 Mary was killed.

2 Mary got killed.

See Givon and Yang (1994) for a discussion of the English get-passive; and Weiner and Labov (1983) for a sociolinguistic approach.

19 For a survey of the meanings of middle voice in Somali, see Saeed (1995).

20 Note that not all neuter middles in Somali have an active form: the verbs *jabo*, *qubo*, *hafo* do, but *garaadso* does not, and the middle verb *dhimo* 'to die' has as its active equivalent a different lexical verb *dil* 'to kill'. It seems that all languages which have a middle voice have some verbs that are inherently middle and have no morphologically related active forms. See Klaiman (1991) for discussion.