

Theta-Roles, Logical Form, and Argument Structures¹

1 Introduction

Semantic roles are the participant roles assigned by a semantic predicate. For example, the predicate adjective tall assigns one participant, or semantic, role. The role is the thing or object that is being described as tall. There are a number of possible semantic labels that could be used to name this role. We will not attempt to go through them here. What is important is how these semantic roles are incorporated into the syntax of a language.

Theta-roles (θ -roles) are semantic roles that are assigned to A-positions in syntax. Grimshaw (1990) argues that there are no distinct θ -roles. They do not occur in the syntax. In the syntax each verb assigns a set of arguments. Each argument is individually identified as 'x', 'y', etc. For the sake of convenience, she continues to use the traditional names. We will do likewise here.

Some examples of semantic roles include the following:

- | | | | |
|-----|----|--|-------------|
| (1) | a. | The <u>rock</u> is falling. | theme |
| | b. | <u>John</u> is happy. | experiencer |
| | c. | <u>John</u> is cold, <u>John</u> feels cold. | experiencer |
| | d. | <u>John</u> is/feels cold (to the touch). | theme |
| | e. | <u>John</u> is sleeping. | experiencer |
| | f. | <u>The door</u> squeaked. | source |
| | g. | <u>The ice</u> formed overnight. | goal |
| | h. | <u>The fish</u> is swimming. | actor |
| | i. | <u>The fish</u> bit through the string. | agent |
| | j. | The fish bit through <u>the string</u> . | patient |
| | k. | <u>the rock</u> broke the window. | instrument |
| | l. | Mary died from <u>the poison</u> . | trigger |
| | m. | It was done <u>for Mary</u> . | benefactive |
| | n. | John left <u>to buy some beer</u> . | purpose |
| | o. | John left <u>because he was cold</u> . | reason. |

¹ This paper is derived from a course on θ -roles taught at Simon Fraser University in 1990.

In (1f) the door emitted a squeak and in (1g) the ice is the final or goal state. The θ -role of noun complements is hard to classify though theme seems to be the best term:

- | | | | |
|-----|----|-------------------------------------|------------|
| (2) | a. | the top of the table | theme |
| | b. | the side of the house | theme |
| | c. | the bottom of the mountain | theme |
| | d. | the colour of the car | theme |
| | e. | the form of an egg | theme |
| | f. | the size of a box | theme |
| | g. | the gift of a book <u>to Mary</u> . | recipient. |

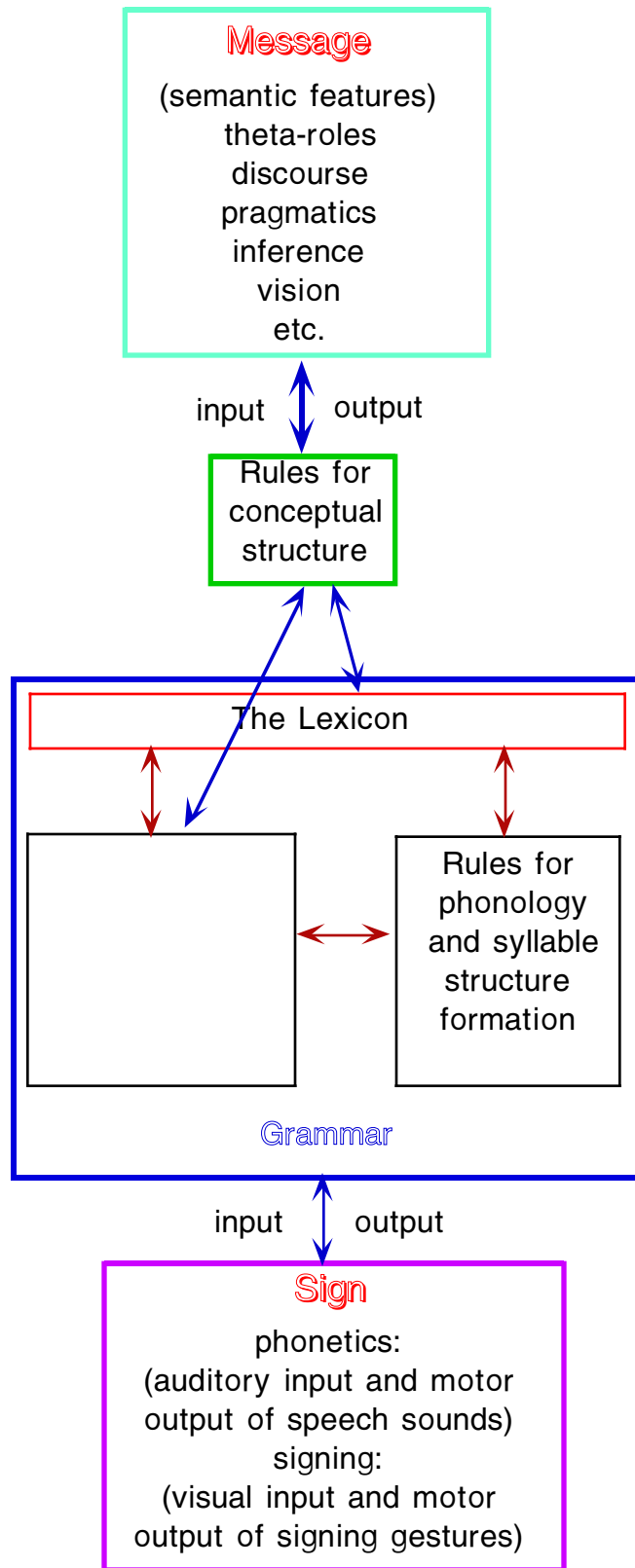
In the first three examples the head noun is a part of some entity. Let us assume that the θ -role that the head N subcategories is “entity.” Example (8c) is ambiguous—bottom could be a member referring to the bottom part of the mountain, or it could be location referring to the area at the base of the mountain. In the latter three examples, the head noun refers to some attribute of the noun. Suppose we consider the θ -role of the complement to be entity as well. Top unambiguously refers to part, and colour to attribute. No single lexical item can be simultaneously a part and an attribute.

The following list contains some of the characterizations of the known θ -roles:

- | | | |
|-----|----|--|
| (3) | | Semantic Role:Characterization |
| | a. | agent: deliberate cause (external change) |
| | b. | actor: deliberate act (no external change) |
| | c. | trigger: an object or event that triggers some event. |
| | d. | instrument: an effector that is used to change something |
| | e. | theme: an object in motion |
| | f. | theme: object being described |
| | g. | experiencer: animate being which experiences an emotion |
| | h. | source: source of MOVE (motion, transfer, etc.) |
| | i. | intermediate: intermediate points of MOVE |
| | j. | goal: goal of MOVE. |
| | k. | topic: the subject matter of a text. |

The following diagram represents the relations between conceptual meaning, form, and sign. It is based on Jackendoff (1991):

(4)



1.1 Logical Form of Sentences

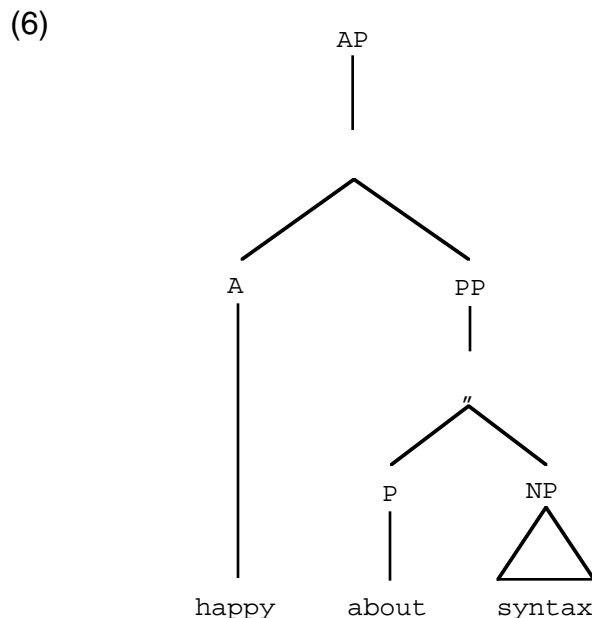
The discussion below follows from the model of syntax and grammar presented in L322. The system is basically a predicate and argument model, where predicates are subdivided into lexical items and operators. Notes for these are listed on my website:

(5) <http://www.sfu.ca/~dearmond/course.outline.322.htm>

The most important topics are propositions, logical form of the sentence, verbal operators, and event classes. Students who are not familiar with these topics or who are need to refresh their memory are strongly encouraged to read these sections.

1.2 Some Notes on Categories

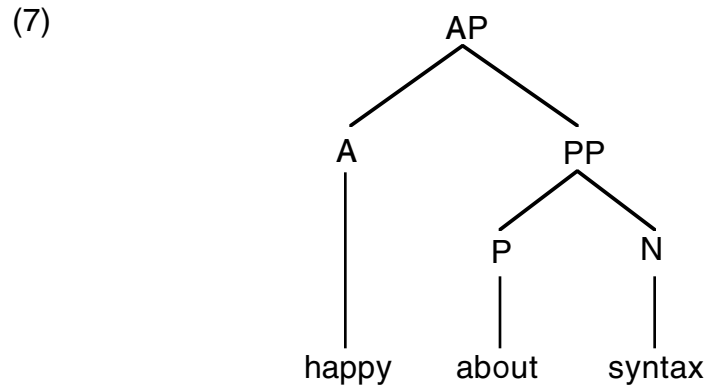
The standard expansion of the \bar{X} -system is the three-level \bar{X} -system: $XP - \bar{X} - X$. This is applied to all categories. However, based on θ -role assignment we will replace the 3-level system with a multilevel system gradually. Initially, we will adopt the three level system except for a few categories—Det, Agr, and perhaps one or two others. For example, initially happy about syntax will be represented as (6) expanding both ASP and PP:



The external θ -role, the subject of T, is contentious. The external θ -role refers to the θ -role assigned to the subject of T (normally). In earlier views and in the view of most current syntax textbooks, its D-structure position is Spec-T. Following more recent arguments, will argue that its origin is internal—it is assigned within the projection of the category which assigns it.

We consider specifiers a moot point and unnecessary in a minimalist form of grammar. A minimalist grammar eliminates all constructs that are considered unnece-

essary and for which there is precious little evidence. In doing so we will adopt the view that all branching is binary. There is neither unary nor multi-branching. In this case, both \bar{A} and \bar{P} will disappear. In the case of syntax, it is simultaneously a head (N) and a phrase (NP). We will write such single level phrases as N (X), where it is understood if N is maximal it is a phrase. In cases where it is unclear we will write “N=NP”. Figure (10) is replaced with:



There are intermediate levels. We will introduce them and how to represent at the appropriate place.

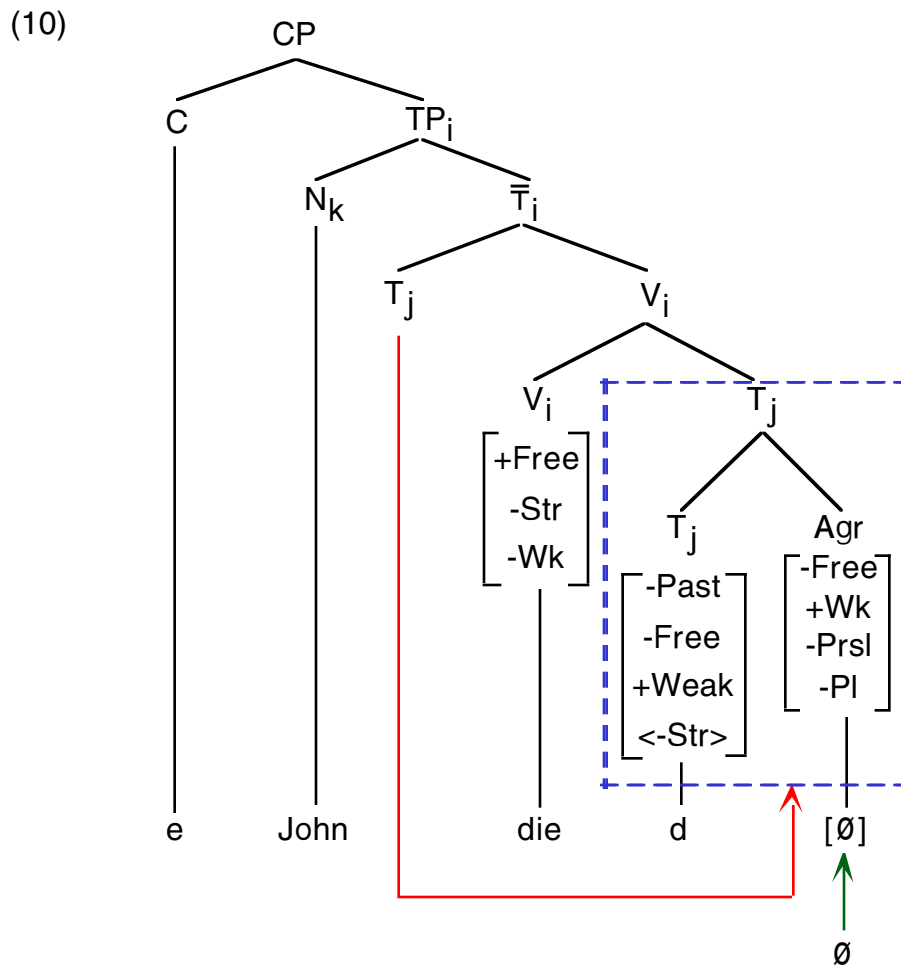
2 Monadic Predicate—Theme or Patient of Verbs.

Let us first start with a monadic (one-argument) predicate. Consider the monadic verb die:

(8) John died.

Die has one argument. In Grimshaw’s (1990) version of argument structure, this argument is an **external** argument in that it is assigned to an external argument position—Spec-T:

(9) John ((x)).



Here, the NP John occurs in the subject argument position. The θ -role assigned to John is called the theme or patient.² We will use this patient here. One lingering question that occurs is how does the θ -role get assigned to a particular A-position? We will return to this problem later. Agr is spelled out as 'd' (/ɪd/).

If we extend the basic concept of minimalism to eliminate null categories, T_j would dominate the features of T_j above plus Agr. This we will spell out below.

Rather than lowering, we propose a theory of feature percolation based on the same in Chomsky (1986). In this theory, 'φ'-features percolate between a head and a modifier of the head, and between the head and the complement of the head. There are certain constraints built into the theory.

In (10) in the older theory 'T' stands for the formal category tense. The even older form 'I' (inflection) is replaced with T and Agr(eement). We assume without argument here that the past tense is formally (/ɪd/), and that agreement is null ([∅]), i.e. unmarked, in the past tense.³ Here we will assume that T is a category whose

² It is controversial whether theme and patient are two distinct theta-roles. Theme usually refers to a steady state and patient that which undergoes an internal change of state.

dominant feature is tense which contains two features [\pm Tense]. First we start with TENSE, a logical operator:

- (11) Expansion of Tense
 [TENSE] \rightarrow [\pm Tense].

[+Tense] refers to an event which is marked for tense, [-Tense] to categories which are not marked for tense; e.g. infinitives and gerunds.

We will differentiate formal categories from lexical items from operators by capitalizing the former (JOHN) and enclosing the latter in caps in square brackets ([TENSE]): [MOOD], [MODALITY], [TENSE], [RELEVANCE], [ASPECT], [VOICE]; BOOK, PEN, WATCH, FOOD, IDEA, NEGATION, RED, BIG, SEE, RUN. These forms do not receive a phonological shape until after certain syntactic rules have applied.

One possible model for the conceptual structure of John died is the following. That John died marks a change of state preceding the time of the speech event. The final state, dead, is a descriptive property of John. The form written in caps is what we will call a predicate (a lexical item) following Reichenbach (1947). It takes one argument—the theta role patient. Arguments are often called theta-roles (or semantic or thematic roles) when they are related to syntax. We will enclose them in angled brackets in the logical representation of the predicate. An incomplete proposition consists of a predicate and its arguments. A complete proposition is that part of a sentence ---proposition includes an eventuality (Bach (1986)), the time of the eventuality, any aspectual modifiers of the eventuality, and other predicates that relate directly to the eventuality. In John died the verbal predicate or eventuality is DIE and JOHN is its argument. The theta role assigned to JOHN is theme, which is written in non-caps:

- (12) DIE <theme: JOHN>.

We assume the Vendler (1957) classification of events as achievement, accomplishment, and a state. Together this set (events plus state) is called an eventuality (Bach (1986)). DIE is an eventuality (E) licenses one argument (a semantic role). E¹ is the predicate containing the eventuality and its arguments. We will merely represent any person named John as JOHN, and put aside the problems associated with naming. Die takes one argument; its lexical entry is revised to the following format:⁴

³ Agreement is marked in the past tense of the verb be—was, were. We assume that both forms are monomorphemic—was includes the feature [-PI], and were includes the feature [+PI].

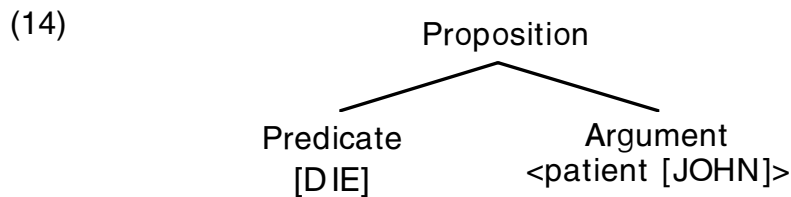
⁴ DIE is not a semantic prime. It is composed of BECOME DEAD, and DEAD may be composed of NOT ALIVE. We will not go into these deeper representations in detail here.

- (13) /da.j/
 +V, -N

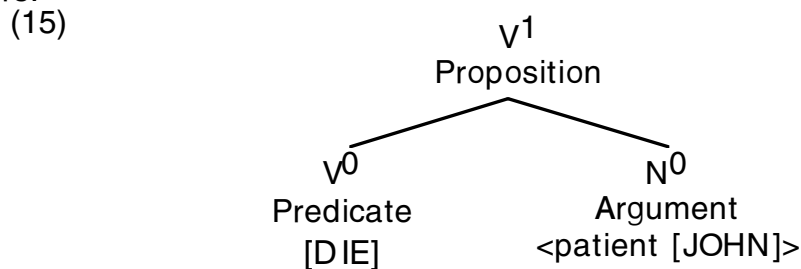
 [BECOME [NOT [ALIVE]]] <theme>

The assigned theta-role is enclosed in angular brackets. Here, we place it on the same line as the set of features that assign it. A single set of angular brackets indicates that there is but one θ -role and it is assigned to the complement of the verb die. Opposed to predicates are objects.

The predicate or object for each category is enclosed in square brackets. For the sake of simplicity, we assume here that die is a simple predicate, not a complex predicate as depicted in (13):

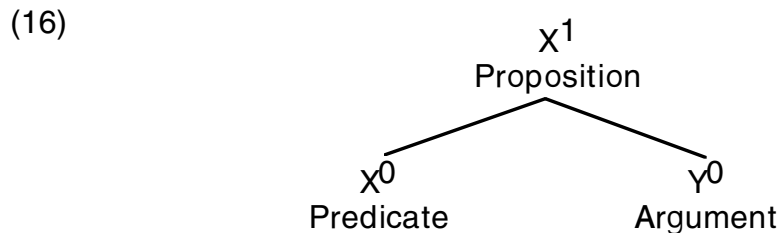


The lexical predicate DIE, which is an eventuality, is spelled out as the head V^0 in the syntax; the lexical entry for DIE contains the information that it is a verb. Logically, DIE takes one argument—the patient John. John is a noun that takes no arguments. Let us suppose that the head N^0 is assigned as the category for John and that if the noun takes no arguments, the head is equivalent to the maximal projection. In the following structure we adopt a variation of \bar{X} syntax. In this revised theory, a D-structure has no more levels of projection than it has complements. That is, Each complement determines a the next higher level of projection. If there are no complements or any arguments of the object, the maximal projection is the same as the minimal projection. We represent this as N^0 , the head level. There are no specifiers. The agent and other external arguments are taken up in subsequent chapters.



It not always necessary to include the terms predicate, proposition, argument, and the theta-roles, since they are implicit except for the theta-roles. By tradition, theta-roles have not been included in tree structures.

To discuss in more detail this alternate version of \bar{X} -theory, let us initially hypothesize that an eventuality (EV) corresponds with a syntactic category X, one argument with the syntactic relation Comp-X, and the proposition¹ with the maximal projection XP. In (16) the maximal projection of X is X¹:



At S-structure, John should be the subject of. Adjectives like nouns do not assign configurational Case. The complement, therefore, is not assigned Case. It must move to NP, where it will be assigned the nominative Case. See also Chapter 9 of Introduction to Syntax, where the D-structure of sentences like are discussed.

We have not accounted for tense in sentences such as (18). We will not cover tense in detail here. The argumentation for tense is found in DeArmond (1993) and (forthcoming). Let us assume here that the argument of the predicate [TENSE] is an incomplete proposition. It cannot be an object:

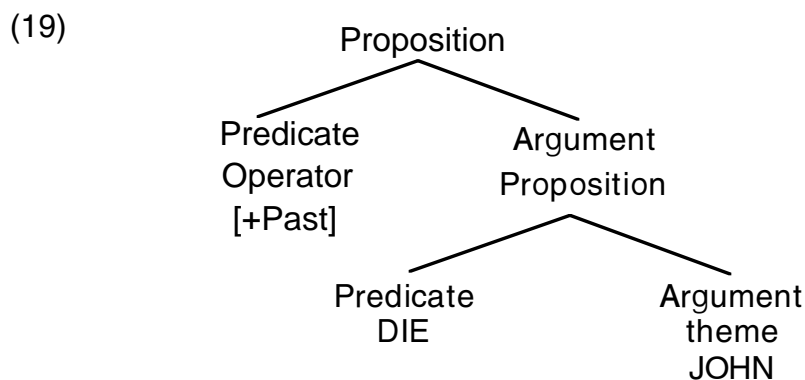
(17) *Carred (\leftarrow [_N car] + [+Past] = [PAST] <[CAR]>).

Only an eventuality can be in the direct scope of [TENSE].

In the following sentence:

(18) John died.

[TENSE] remains when the predicates JOHN and DIE are extracted. The proposition DIE <JOHN> is the argument of [TENSE]

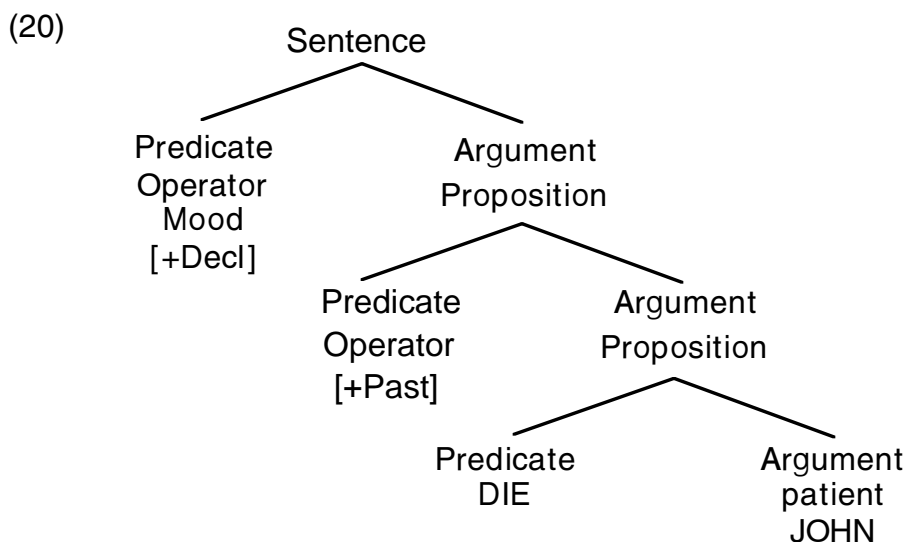


[TENSE] here refers to the logical operator denoting time, and tense to grammatical category of tense. Note that the argument of [TENSE] is enclosed in angular brackets

ets. The upper incomplete proposition takes [TENSE] as its head whose argument is the incomplete proposition DIE <patient: JOHN>.

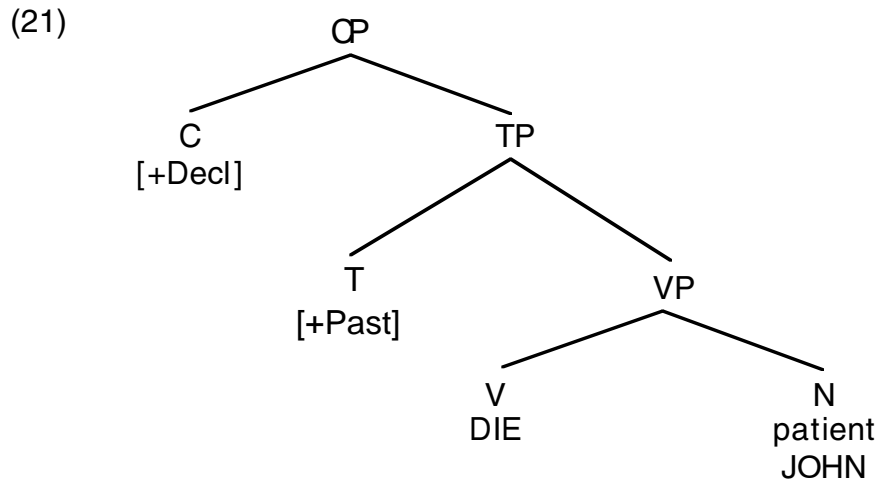
By definition a proposition is that part of a sentence excluding the category of mood (Crystal (1990)). Mood determines such factors as speaker's assumed reality, condition, commands, etc. The mood expressed in the majority of the examples here is the declarative mood, which is a statement the speaker intends to be true, or wants to be accepted as true (as in lying). Other moods will be noted. The term proposition is not really appropriate here, since the logical structure of the sentence must include the mood modifying the sentence plus the propositional content of the sentence. We will use the term Sentence to refer to the logical form of the sentence, and sentence to refer to the syntactic form. We do not anticipate much confusion here. If so, then another term will have to be selected.

The logical structure for the sentence is expanded to the following:



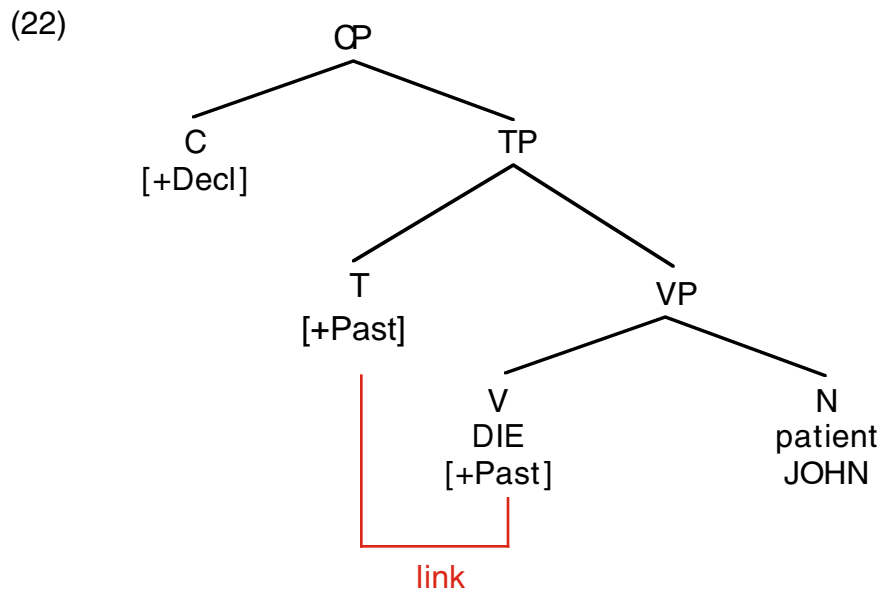
The lexical category is assigned to the above structure starting from the bottom. Each lexical item contains its categorial information. JOHN is assigned N, and DIE V. Since DIE is the head of the proposition that directly contains it (immediately dominates it), the proposition is labelled VP. [+Past] is assigned the label T and the proposition that contains it is labelled TP.⁵ Mood is labelled C (for complementizer rather than M for mood) and the Sentence is labelled CP. In the following structure we not use the labels predicate, operator, argument, proposition for the sake of clarity:

⁵ In a slightly different theory of phrasal labelling, TP would be VP, dominating T (head) and VP (argument). The rationale of this theory is not considered worthy going into at this time.



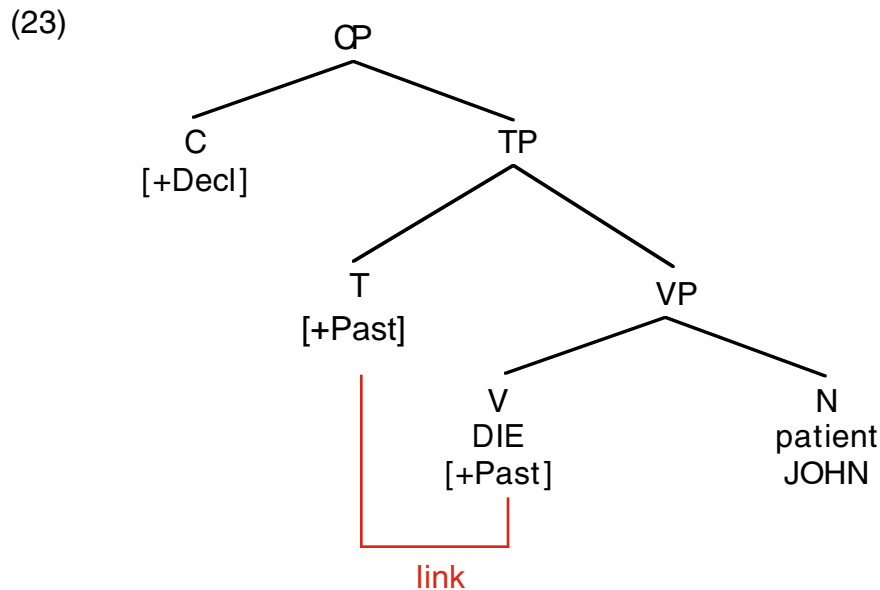
To obtain the phonological form of each predicate we must return to the lexicon to obtain this information. However, before we can do that, certain syntactic processes must be applied.

[+Past] needs a verbal host. It must therefore lower to the closest V. However, rather than claim that T lowers to V, we now say that the features of T are copied onto V. We do this by creating a link between a head and its primary complement. The relevant features of the head are then transferred through the link to the complement:

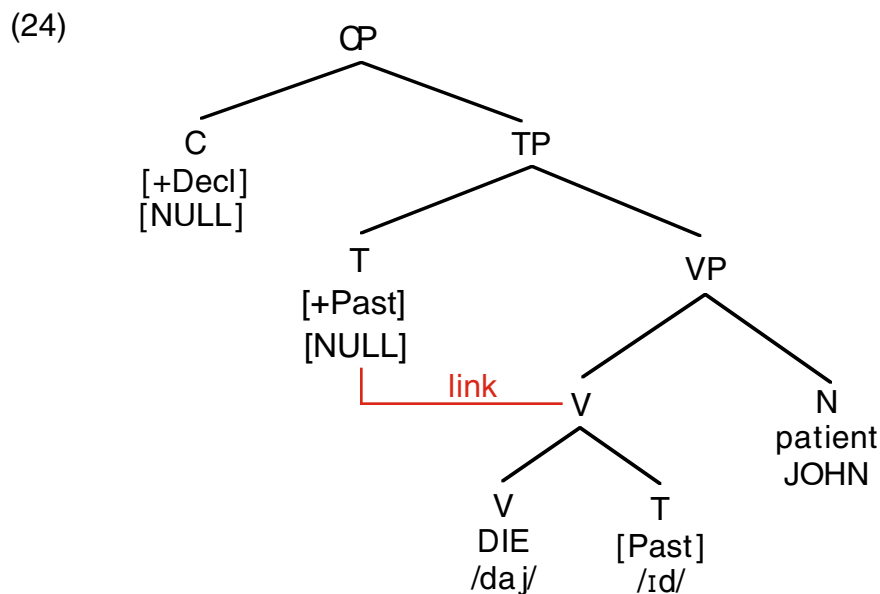


The feature [+Decl] is not spelled out in English if it is in a matrix clause. It is optionally spelled out as that in embedded clauses subject to certain conditions. 'T [+Past]' cannot be spelled out because [+Past] has no host. The task which lies ahead of is to spell out the feature bundle 'V DIE [+Past]'. A default rule splits V

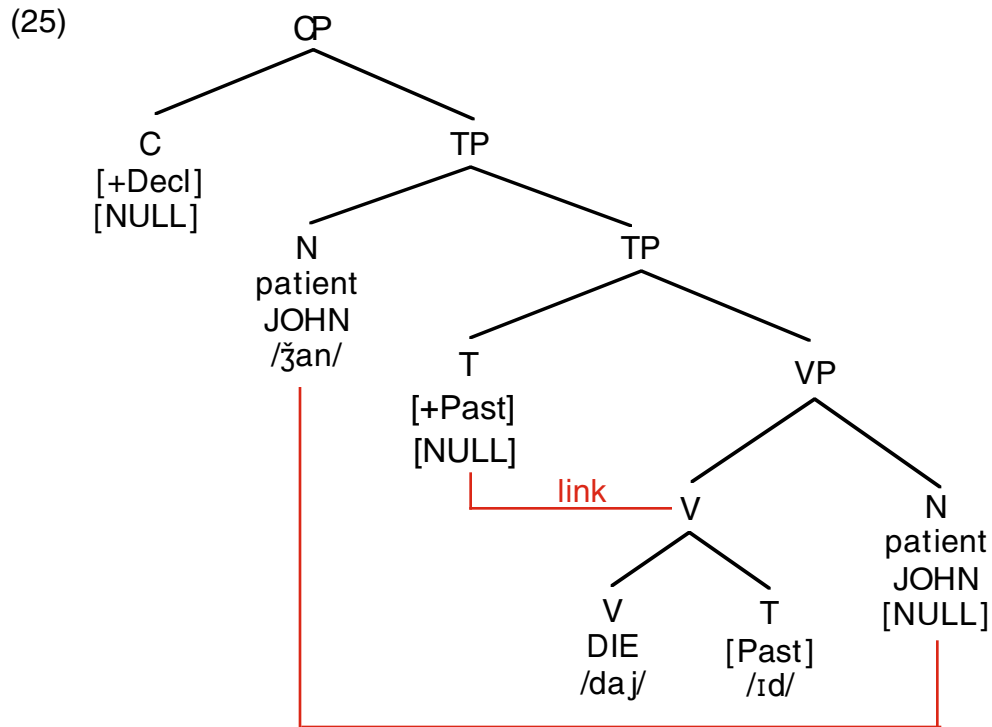
into V + T. The features of V remain under V, while the features of T, here [+Past], go under T:



In exceptional cases V does not split; *v* is spelled out as an irregular verb form; e.g. the past tense of SWIM is spelled out as *swam*. The default for [+Past] is the suffix ‘-ed’:



JOHN is spelled out as John after Case assignment has applied. DIE is intransitive and cannot assign Case to its complement. This plus the Extended Projection Principle (all clauses need a subject to license the verb) causes JOHN to move to TP and be adjoined to it where it receives the nominative Case from T [+Past]:



Once again we may say that the lexical item JOHN is copied to its new position where it is spelled out as John. In its original position, it is not marked for Case and cannot be spelled out.

3 The Form of Lexical Entries

The lexical entry of an item should follow the basic form given in (26):

- (26) Lexical Entry Form
- orthographic form of lexical item
 - /phonological representation/
 - morphological information
 - category ($\pm N$, $\pm V$, $\pm A$, $\pm P$, $\pm T$, $\pm C$)
 - irregularity class (if irregular)
 - exception rules (if any exist for the lexical item)
 - subcategorization frame (syntax) (if unpredictable)
 - host information
 - semantic features <arguments: theta roles>

- (27) $\left[\begin{array}{l} \text{happy} \\ /hæpi/ \\ -V, +N \\ \text{---} \\ \text{HAPPY <theme>} \end{array} \right]$
- (28) $\left[\begin{array}{l} \text{feel} \\ /fi:l/ \\ +V, -N \\ /i/ \rightarrow /ɛ/ / \text{---} [+Past] \\ \text{--- AP} \\ \text{FEEL <experiencer>} \end{array} \right]$
- (29) $\left[\begin{array}{l} \text{top} \\ /tɑ:p/ \\ -V, +N \\ \text{---} [_{PP} \text{ of NP }] \\ \text{SURFACE <theme>} \end{array} \right]$
- (30) $\left[\begin{array}{l} \text{ɪd} \\ /ɪd/ \\ +T \\ \text{(default form)} \\ [-Free] \\ \text{PAST <Pred'>} \end{array} \right]$

The morpheme /ɪd/ is the default ending in that the irregular endings are adjoined to verb stems marked for the irregular endings. The default ending is adjoined to verb stems that are not so-marked.

