XTAG using XMG

A Core Tree-Adjoining Grammar for English

Thesis submitted to

Saarland University

Department of Computational Linguistics

for the Master of Science in

Language and Communication Technologies

Author:

Katya Alahverdzhieva

Advisers: *Co-supervisor:*

Claire Gardent   Hans Uszkoreit

Tania Avgustinova

— 20 September 2008 —

Statement

I hereby confirm that this thesis has been my independent work and that I have not used any other sources except where the resources are stated.

Katya Alahverdzhieva
Saarbrücken
September 20, 2008

ACKNOWLEDGEMENTS
I would like to express my sincere gratitude to ...

... my supervisor from Nancy Claire Gardent for giving me the chance to work on this fascinating project, for the great many creative ideas, for the meticulous scientific discussions, for the patience, support and encouragement.

... my supervisors from Saarbručken – Tania Avgustinova and Hans Uszkoreit for agreeing to work with me on this project and for broadening my scientific knowledge with their great many creative ideas

... Yannick Parmentier for the technical guidance, for responding to my constant enquiries and complaints within few hours, for being actively involved in my thesis.

... Hervé Saint-Amand for providing me with constant technical support, for reading several draft versions of this report and for the inspiration during informal scientific talks.

... all my professors from the University of Saarland and University of Nancy 2, particularly to Patrick Blackburn and Guy Perrier, who raised my knowledge and who made me think critically about scientific issues.

Abstract

This Master's thesis focuses on developing a computational grammar for English using the tree-adjoining syntactic formalism. We proceed from the linguistic assumptions made in an already existing large-scale English grammar and provide a methodology for their implementation. This project is carried out in the context of the XMG framework, which is specially designed for factorising strongly lexicalised tree-based grammars and grammars equipped with a syntax/semantics interface. The description language we use is built upon the observation that fine-grained grammar development can be performed on a higher abstraction level by taking advantage, on one hand, of structure sharing and, on other hand, of alternative structures. This is validated by developing a complete grammar for English verb families, where the two basic operations are composition for specifying shared structures and disjunction for representing possible alternatives.

We begin this report by introducing in a motivated way the increased power and the basic operations of the tree-adjoining grammar formalism. We carry on with a discussion about the language of description, which allows the process of grammar development to be guided by abstracting over grammatical information. The adopted methodology involves four dimensions of abstraction, each of them being aimed at defining different levels of syntactic structures. We apply this methodology to the implementation of a core English grammar and we illustrate its efficiency by building a test suite of sentences covering various syntactic phenomena.

Contents

Introduction

The formal syntactic analysis of natural languages dates back to the 50's when [Chomsky, 1957] attempts to associate language universals with a mathematically precise model of linguistic structures. His ideas collected in 'generative grammars' break away from traditional linguistic theory by defining natural language syntax with a finite set of simple mathematical rules. In the very beginning this may have seemed a rather ambitious goal, but the developments in modern linguistic theory and in particular in computational linguistics have shown this to be an attainable goal.

Further extensions and variations of Chomsky's context-free grammars aim at higher expressive power and stronger lexicalisation. A grammar framework should be on one hand powerful enough to describe all natural language phenomena, and on the other strongly lexicalised to associate a finite set of structures with each lexical item. Lexicalisation has a twofold advantage: first, it makes the number of possible analyses for any given string finite (assuming that the grammar is finite) and second, each lexical item carries information about its subcategorisation frame and the syntactic alternations it undergoes, thereby enhancing performance.

This report addresses the development of a strongly lexicalised tree-based grammar, namely a lexicalised tree-adjoining grammar for English. A major drawback of high lexicalisation is that it comes along with hundreds of redundant elementary tree descriptions, making the maintenance and the further development of the grammar a rather cumbersome task. This is the primary motivation for developing a declarative and minimal language for factorising grammatical information.

The framework used in this Master's topic supports a high level of generalisation, making it possible to significantly reduce the number of tree descriptions. More precisely, the XMG language allows one to identify in a monotonic and non-redundant way tree descriptions yielding a redundant and complete grammar. From a methodological point of view, two types of generalisations are aimed at: shared grammatical structures and alternative realisations of a predicate and its arguments. These two generalisations prove to be sufficient for designing a computational tree-based grammar of English.

Plan

This document is organised in four main chapters: aims and background; presentation of the XMG framework; description of the methodology used to implement the grammar; evaluation. The first chapter introduces the goal of this thesis and details the resources used for its realisation: the grammatical formalism, the linguistic decisions, and the language of description. This chapter, however, does not end here as it has a deeper aim – by detailing some of the drawbacks of a possible implementation of a tree-adjoining grammar, it motivates a minimal description language that allows for specifying grammatical information in an economical and a complete way. The next chapter introduces the XMG framework used for implementing tree-based grammars. We will show how the system we use combines both the linguistic formalism for specifying tree descriptions and the compiler for computing grammar rules out of these descriptions. The XMG framework is put to use in the third chapter where we present the methodology adopted

for developing a grammar for English verbs and illustrate the organisation of the complete grammar by means of examples. The last chapter is dedicated to evaluating the grammar by building a test suite of phrases and sentences that can be parsed by the grammar.

Chapter 1

Aims and Context

This Master's thesis is aimed at specifying a tree-adjoining grammar for English verb families within the context of the *XMG (eXtensible MetaGrammar)* grammar writing environment developed at LORIA in Nancy. More specifically, the work consists in using XMG to re-implement the verbal part of the large-scale English TAG designed by the University of Pennsylvania, and in evaluating its coverage by building a lexicon and a test suite of sentences covering various phenomena that can be parsed by the grammar. The two backbones of this report are thus the XTAG grammar and the XMG tool.

1.1 Tree-Adjoining Grammar

This chapter presents a highly lexicalised formalism which is well-suited for natural language grammars. Tree-adjoining grammar is a mildly context sensitive formalism which has been introduced by [Joshi and Schabes, 1997] to overcome the limitations of context-free grammars, and further augmented with unification-based features (see [Vijay-Shanker, 1987] and [Vijay-Shanker, 1991]) and lexicalisation (see [Schabes, Abeille, and Joshi, 1988]).

Linguistically, an important feature of using tree-adjoining grammars is their extended domain of locality. This means that one can add local constraints to non-local phenomena (e.g.,long distance dependencies, topicalisation) without losing information about the subcategorisation frame, case marking, agreement, etc.

Last but not least, tree-adjoining grammars can be fully lexicalised so that every elementary tree description is associated with a lexical item, called an anchor. The Greibach normal form of context-free grammars where any rule is of the form $A \rightarrow w\ B1\ B2\ \ldots Bn$ is a way for lexicalisation which, unfortunately, does not preserve the tree language (hence it is weakly equivalent to a lexicalised TAG). On other hand, the extended domain of locality and the structured tree objects make it possible to strongly lexicalise a TAG description preserving the same set of structural descriptions.

1.1.1 Formal definition

Definition 1. A tree-adjoining grammar is defined as a quintuple $\langle NT\ T\ S\ I\ A \rangle$ such that

- *NT* is a finite set of non-terminal symbols (corresponding to syntactic categories)

- *T* is a finite set of terminal symbols

- $S \in NT$ is an axiom

- *I* is finite set of *initial trees* with the following properties:

  - interior nodes are marked by non-terminal symbols

  - nodes on the frontier are marked either by non-terminal (indicated as substitution nodes in terms of a downarrow sign ($\downarrow$)) or by terminal symbols

- *A* is a finite set of *auxiliary trees* with the following properties:

  - interior nodes are marked by non-terminal symbols

  - one frontier node is labelled with the same category as the root of the tree, i.e. this node is the foot node and it is marked by an asterisk ($*$)

The union of initial and auxiliary trees $I \cup A$ comprises the set of *elementary trees*.

Elementary trees are combined by means of two rewriting operations: *substitution* and *adjunction*.

Definition 2. The *substitution operation* (Figure 1.1) consists in replacing a leaf node marked with a downarrow ($\downarrow$) labelled by a category Y of a tree α by an elementary tree β rooted in the same category Y.
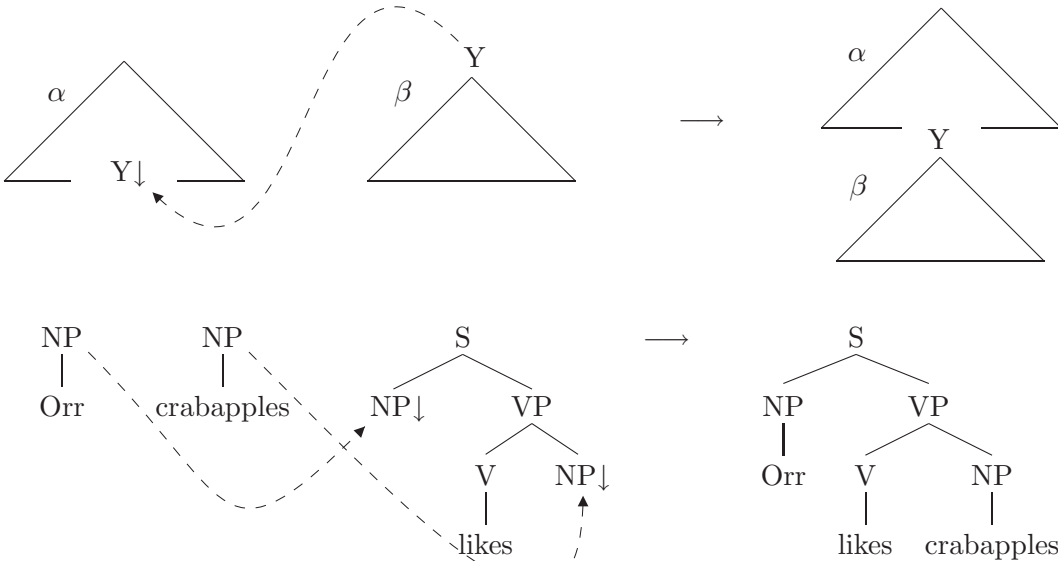
Figure 1.1: Substitution in TAG

Definition 3. The *adjunction operation* (Figure 1.2) consists in splicing an internal node n labelled by a category Y of a tree α to insert an auxiliary tree β. If the foot and the root node of the auxiliary tree are marked by the same category Y, then one can derive a tree γ by means of the adjunction operation where:

- the sub-tree of α dominated by n is split from α;

- the auxiliary tree β is inserted under n so that its root is identified with n;

- the split sub-tree is inserted under the foot node of β so that its root is identified with the

foot node of β

TAG does not allow adjunction to be performed on nodes marked for substitution.
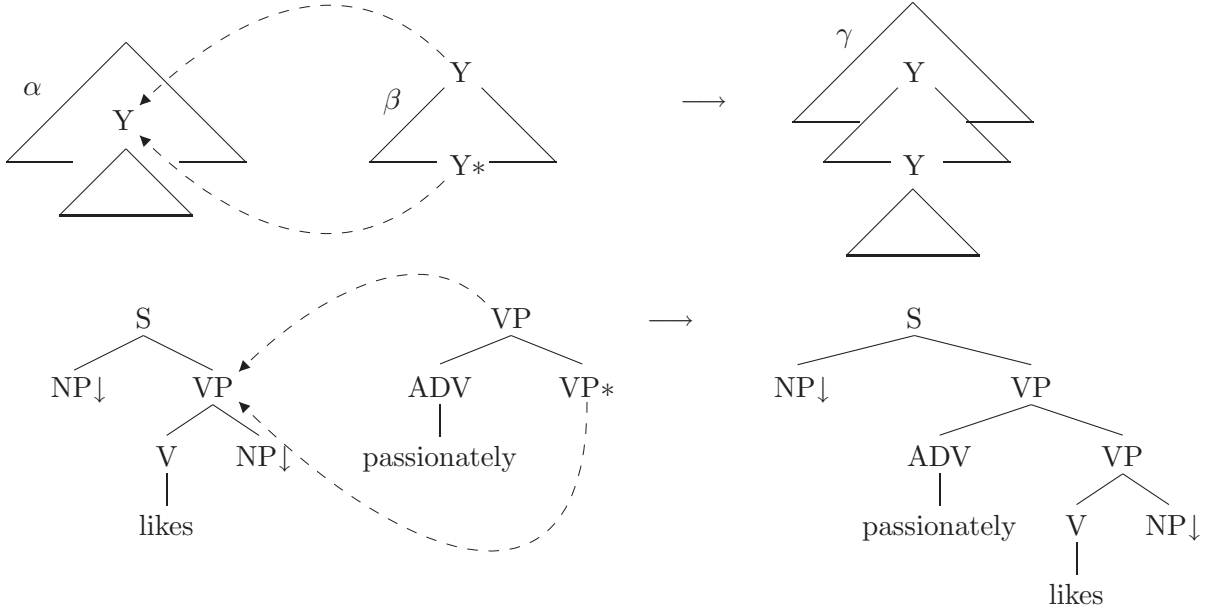
Figure 1.2: Adjunction in TAG

To sum up, when applying the substitution and adjunction operations within a set of elementary trees, one can derive a fully lexicalised tree whose terminal nodes are anchored by lexical items. These lexical items are to be referred to as *anchors* and the resolved trees as *derived trees*. Whenever a tree is anchored by more than one lexical item (e.g., verbs occurring with a specific preposition), the second lexical item is known as a *coanchor*. The anchor and coanchor nodes are graphically represented by the diamond symbol ($\diamond$).

1.1.2   Feature-based Tree Adjoining Grammars

This extension of TAG (proposed by [Vijay-Shanker, 1987]) associates each node in an elementary tree with two feature structures[1] called *top* and *bottom*. The reason for this partition is that we need to account for adjunction. Since by this operation a node is excised in two parts, one should be able to unify its top part with the root node and its bottom part with the foot node of the adjoined tree. The *top* feature of a node n is associated with the features of its dominating node, whereas the *bottom* feature refers to the daughter nodes. The derivation process involves propagating *top* and *bottom* features upwards in the tree. The derived tree is obtained after unifying features as follows: the *top* feature of a node marked for substitution is unified with the *top* feature of the tree to be substituted (Figure 1.3). In the case of adjunction, after splicing the node n for adjunction, its *top* feature structures unify with *top* of the root node of the auxiliary tree, while its *bottom* features unify with the *bottom* features of the foot node of the auxiliary tree. At the end of the derivation, the *top* and *bottom* feature structures of every node should be unified (Figure 1.4).

---

[1] also known as attribute-value matrices
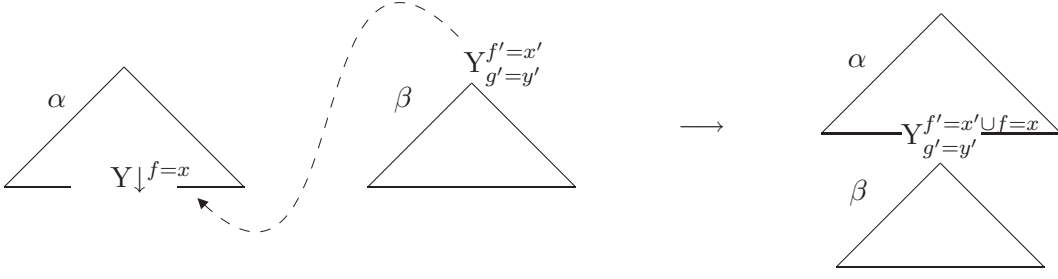
Figure 1.3: Substitution in FTAG
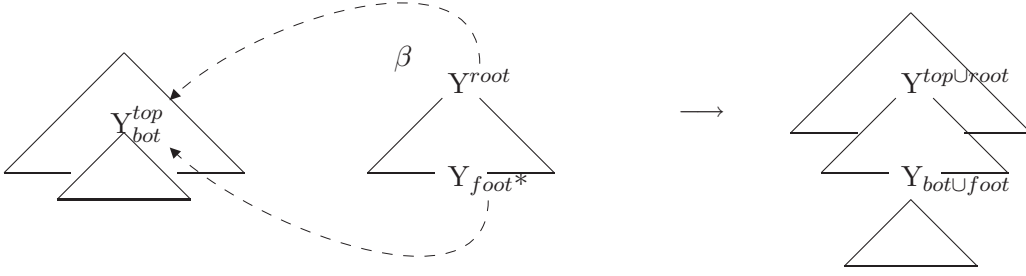
Figure 1.4: Adjunction in FTAG

1.2 XTAG

The XTAG grammar developed for English within the lexicalised tree-adjoining grammar formalism is a linguistically precise and well-motivated grammar, and therefore serves as a starting point for the present thesis. The overall XMG organisation of our work is consistent with the linguistic principles that underlie the XTAG grammar, the major one being organising the grammar into separate *tree families* in accordance to their subcategorisation frame, with a clear-cut distinction between complements and adjuncts. The XMG-based TAG covers the linguistic phenomena in XTAG (e.g., long-distance dependencies, passivisation, relativised complements and adjuncts, pied-piping) differing in the way information is factorised.

We proceed from the assumption that the development of a TAG grammar involves defining a finite set of verb families and a set of syntactic transformations within each family. To put it another way, a lexical anchor subcategorises for complements (which can appear in various syntactic environments) forming a clausal structure. Thus, each verb family selects only those verbs that undergo the same syntactic variations.

Along the lines of traditional generative syntax, the complement/adjunct distinction is drawn on the line of grammaticality: complements are the phrasal categories strongly selected (i.e., subcategorised) by the head of a phrase and adjuncts are optional elements that do not influence the grammaticality of the clause. In TAG, this is handled by including the complements in the elementary trees anchored by the verb and by modelling adjuncts using auxiliary trees.

The driving principles concerning specific phenomena in our grammar are as follows:

- extraction: the XTAG analysis follows the traditional trace-based (or wh-movement) ap-

proach of government and binding (GB, see [Chomsky, 1981]) where the extracted element and the empty category ϵ are coindexed by a $\langle trace \rangle$ feature

- passivisation: the traditional movement analysis assumes that the direct object is moved to a subject position, whereas the syntactic subject is either absent or is headed by a *by*-phrase. In XTAG, the passive constructions appear as separate elementary trees within the corresponding tree family

- gerunds: the XTAG grammar splits the gerunds into two classes: derived nominalisations (where the direct object is instantiated as *of PP*) and sentential gerundives (a.k.a. NP gerunds) which assign case to the direct object and do not require the insertion of a preposition

- relative clauses: following the GB analysis of relative clauses, XTAG represents them as auxiliary trees that adjoin to NP's anchored by a verb in the corresponding tree family. The coindexation between the extracted argument and its position before movement is captured by the $\langle trace \rangle$ feature

1.3   Why XMG?

Hand-written large-scale lexicalised TAGs amount to several thousand trees sharing a big amount of common structures. Redundancy is the major problem when it boils down to writing and maintaining such grammars. A possible solution is then to abstract over information at a higher meta-level, reducing the number of repetitive descriptions. In this report, we resort to the XMG framework for this abstraction for two reasons.

Firstly, XMG provides a linguistic formalism of higher **abstraction** level where tree fragments are seen as reusable chunks of information. Each tree fragment is associated with a unique name and hence, can be used either in an inheritance hierarchy or in a logical operation. Hence, grammatical information is better factorised and the problem of redundant tree descriptions is overcome by factoring out shared pieces of linguistic information.

Secondly, the metagrammar framework provides a monotonic and fully-declarative language that gives an equal status to all trees within a family by defining them as alternatives of each other. This means that XMG does not impose any ordering on the application of syntactic operations (i.e., syntactic variations are not transformations of the canonical form but their alternative constructions). In comparison to that, the XTAG system used for generating grammars from tree descriptions (see [Xia, 2001]) relies on three independent modules corresponding to each other by intermediate interfaces. The first module permits describing canonical subcategorisation frames, the second one is meant to derive additional frames by means of *lexical redistribution rules*, and the third one allows for non-deterministic combination of tree descriptions. The lexical rules incorporate alternative realisations by applying a syntactic transformation t to a subcategorisation frame r. With these means one can generate the passive frame from the active one. This procedural order of application justifies the three-way partitioning of the system. To overcome this non-monotonicity, XMG takes advantage of a single language where the basic operations are **disjunction** for capturing alternatives and **conjunction** for capturing structure

sharing – thus yielding a more compact TAG representation[2].

---

[2]Although the specification of a syntax/semantics interface is not part of the current goals, it is worth mentioning that XMG permits an easy integration of unification-based compositional semantics. The process of interfacing between syntactic functions and semantic roles involves assigning to each elementary tree a flat logical representation (see [Gardent and Kallmeyer, 2003]).

Chapter 2

The Metagrammar Framework

The notion of a *metagrammar* was first introduced by [Candito, 1999], who laid down the guiding principles for semi-automatic generation of TAG. In her PhD thesis, trees are presented as reduced descriptions organised in a modular and hierarchical way. These partial descriptions capture basic linguistic generalisations and their combination produces elementary TAG trees.

XMG builds on the notion of a metagrammar as a finite set of partial tree descriptions aimed at factorising common pieces of grammatical information. At the same time, it was developed as a system for generating large-scale unification-based grammars from meta-descriptions, and particularly grammars equipped with syntax/semantics interface (see [Y. Parmentier and J. Le Roux, 2005]). The framework combines the formalism for describing meta-grammatical information and the compiler for computing the actual grammar from the abstract tree descriptions (see [Duchier, Le Roux, and Parmentier, 2004]).

2.1　XMG as a linguistic formalism

XMG was designed as a framework for grammar writing meant to facilitate the development of strongly lexicalised grammars (see [Crabbé and Duchier, 2004]). The intuition behind its compact representation is that the metagrammar takes advantage of a logic program. The knowledge base consists of `facts`, represented as tree fragments, and of `rules`. A fragment is given a unique name which corresponds to the `rule's head`, and the combination of fragments or the call of a single fragment makes up the `rule's body`. Then, when `queries` are evaluated, the computation of the grammar trees takes place.

Any piece of syntactic information is specified in terms of tree descriptions, called *classes*. The notion of a class is crucial – each one is provided with a unique name, so that it can be further reused when importing a class from a higher hierarchical level or when using it in a disjunctively or conjunctively defined set of tree descriptions. The abstract syntax for specifying a class is given in Figure 2.1.

This rather simple representation language allows the grammar engineer to associate a class name to the items in its description called the `goal`. The goal refers to one of the following

$$Class ::= Name \rightarrow Goal$$
$$Goal ::= Description \mid Name \mid Goal \ \wedge \ Goal \mid Goal \ \vee \ Goal$$
$$Query ::= Name$$

Figure 2.1: Core Language of a Metagrammar

primitives:

1. **description**: tree descriptions are associated with a given name, e.g.:

   CanonicalSubject: WhSubject: ActiveVerb:

   (2.1)
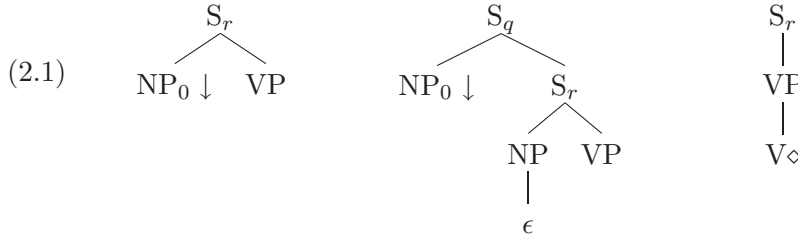
   The left-hand side in example 2.1 states that the canonical realisation of a subject argument NP_0 is rooted in an S node, and the subject node immediately precedes the node of the verbal phrase. The graphical representation in the middle shows that whenever the subject is wh-moved, the canonical subject position is left empty and the subject argument is moved to a certain distance before the verb. The tree description on the right-hand side represents the active verb as forming a skeleton from a root S node to a lexicalised V node.

2. **name:** when calling a tree description from other parts of the grammar specification, one can simply reuse that tree description's name. Thus, the clause

   (2.2) $Subject \rightarrow CanonicalSubject$

   reuses the tree description associated with the class $CanonicalSubject$.

3. **disjunction of tree fragments**: possible realisations of a predicate-argument structure (2.3) and diathesis alternations (2.4) are viewed as equally ranked variants, which are formally defined by means of disjunction:

   (2.3) $Argument \rightarrow ArgumentRealisation_1 \vee ArgumentRealisation_2$

   (2.4) $Family \rightarrow Diathesis_1 \vee Diathesis_2$

   Using the general rule in (2.1) and having already associated tree descriptions to a certain name (2.1), it is straightforward to declare a class $Subject$ as a disjunction of its canonical realisation and its wh-extraction:

   $Subject \rightarrow CanonicalSubject \vee WhSubject$

4. **conjunction of tree fragments**: a class combines tree descriptions into a tree family and thus, non-deterministically provides a complete predicate-argument structure together with its possible realisations:

(2.5) $class\ n0V\ \rightarrow\ Subject\ \wedge\ ActiveVerb$

The above example states that the complete predicate-argument schema of the intransitive verb family $n0V$ is built by combining the tree descriptions of a subject with the description of an active verb form.

2.2   XMG as a metagrammar compiler

Once the metagrammar specification is completed, the basic tree fragments can be combined by the XMG compiler. It takes as input the concrete metagrammatical rules and outputs TAG trees in three steps:

- translating the concrete syntactic objects of the metagrammar into instructions for a virtual machine designed in the spirit of the Warren Abstract Machine (see [Aït-Kaci, 1991]). This step involves resolving the relations that hold between the nodes (precedence and dominance relations, transitive reflexive and non-reflexive closure, node equality);

- executing the instructions. This includes unification of structures referring to relevant information within fragments;

- solving tree descriptions

For instance, considering the snippets of a grammar definition given as examples above, when given the axiom $n0V$ the interpreter generates two possible solutions (after the rightarrow sign below) capturing the alternative subject realisations:

CanonicalSubject: ActiveVerb:
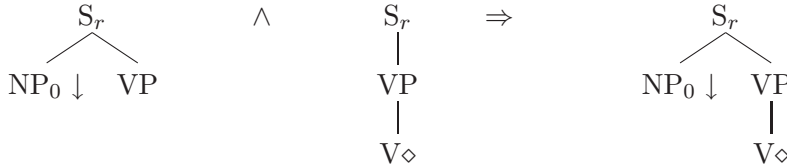
WhSubject: ActiveVerb:
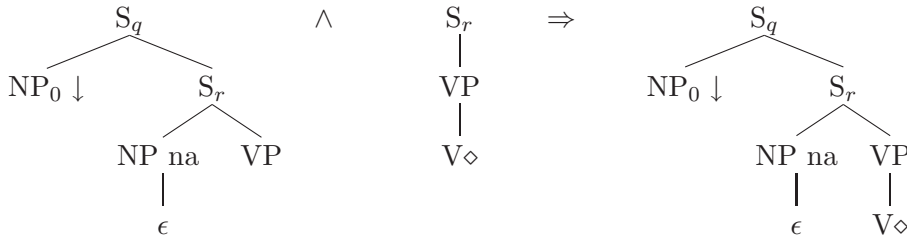
2.3   Constraining the language: colouring principle

To ensure the well-formedness of the resolved trees, the metagrammar system allows for constraining the process of solving permissible tree fragments. [Crabbé and Duchier, 2004] propose

decorating each node with colour-based polarities in order to restrict the merging of nodes. This principle postulates that each node in a final tree description should be painted either black, red or white where:

- a black node can be combined with any number (incl. 0) of white nodes, producing a black node

- a red node cannot be merged with any other node

- white indicates a non-saturated node which needs to be merged with a nonzero number of black nodes

This can be graphically represented as follows:

	B	**W**	**R**
B	$\perp$	B	$\perp$
W	B	W	$\perp$
R	$\perp$	$\perp$	$\perp$

Figure 2.2: Colouring principle

This model is similar to the polarity-based tree descriptions in Interaction Grammars described by [Perrier, 2007]. According to this principle, tree descriptions are viewed as consumable resources, which need to be fully saturated to produce a valid syntactic description. Negative polarity on a node ($\alpha \leftarrow \beta$) expresses the need of providing a resource for consumption. A positive polarity ($\alpha \rightarrow \beta$) indicates a resource available for consumption. Finally, equal polarity ($\alpha = \beta$) marks fully-saturated non-consumable resources.

Summing up   This chapter introduced the basic ideas of the metagrammar framework we use. It was shown that a simple representation language of higher abstraction level suffices for specifying tree descriptions and for constraining their combination. We saw that this monotonic language does not involve any transformation for deriving a tree description from another but rather lets us define fragments of equal status without going through intermediate steps.

Chapter 3

XMG-based XTAG

This chapter is structured as follows: the first section begins by discussing the methodology adopted for specifying the grammar in the presented formalism. We shall see that grammatical information is factorised in four dimensions, i.e., tree families, diathesis alternations, syntactic functions and partial tree descriptions. In the four sections that follow we put this methodology into practice: sections (3.2), (3.3), (3.4) and (3.5) detail the specification of each of these description levels by introducing theoretical background and empirical data.

3.1 Methodology

The three-dimensional methodology introduced by [Candito, 1999] serves as a main inspiration for the development of a core French tree-adjoining grammar (see [**?**]) and it is the keystone for the organisation of an XMG-based TAG. The proposed three-dimensional system addresses the problem of redundant descriptions by representing lexico-syntactic structures in a hierarchy that preserves the well-formedness of TAG trees[1]. Non-redundant grammatical descriptions can be obtained by generalisations, which specify structure sharing or alternations. To be more specific, these generalisations can, on one hand, factor out common syntactic structures described as a set of independent tree fragments. These fragments represent either syntactic realisations of a predicate argument or realisations of the predicate[2] itself. On the other hand, generalisations can be expressed as alternative surface-realisations of a diathesis (active/passive), of an argument structure, or of a verb form. These two types of grammatical generalisations are formally defined either by means of conjunction or by means of disjunction.

Since we aim at generating verb families with their possible subcategorisation realisation, the overall work abstracts over generalisations in four dimensions:

1. Tree families. At this level, a class defines the set of trees associated with verbs of a given syntactic type. For instance, the *n0Vn1* class defines the set of trees associated with transitive verbs such as *love*.

---

[1]The first dimension handles canonical subcategorisation frames, the second dimension handles redistribution of syntactic functions, and the third is concerned with possible realisations of syntactic functions (e.g., in a canonical, extracted or clitic position)

[2]By a predicate, we mean full-meaning (i.e., non-auxiliary) verbs and nouns subcategorising for syntactic arguments

A tree family class is defined as a disjunction of diathesis classes.

2. Diathesis. A diathesis class defines the set of trees associated with a given (tree family, diathesis) pair. Here we aim at capturing active/passive variations withing a verb family. For instance, the *n0Vn1DiathesisActive* class defines the set of trees associated with the active form of a transitive verb, whereas the *n0Vn1DiathesisPassive* class is associated with its passive form.
   A diathesis class is defined as a conjunction of a verb form and syntactic function classes.

3. Syntactic function classes. These classes describe the set of tree fragments associated with each possible syntactic function. For instance, the *Subject* class defines the set of tree fragments describing the possible syntactic realisation of a subject (canonical subject, questioned subject, relativised subject, etc.)
   Verb forms and syntactic function classes are defined as disjunctions of tree fragments classes.

4. Tree fragments. At this level, a class describes the basic tree fragments used to build the grammar.
   Tree fragments are defined as tree descriptions and are organised in an inheritance hierarchy.

3.1.1 Defining tree families

In the TAG literature, verbs are organised in distinct families on the basis of the subcategorisation frame a lexical item selects. More specifically, a verb family comprises the set of alternative surface realisations of the same predicate-argument structure. In contrast to XTAG, where the different contexts of a predicate and its arguments are handled as syntactic transformations, we regard them as variations of the prototypical (i.e., realisation of an argument in its canonical position) case: extraposition, relativisation, active gerund, etc. Therefore, a tree family can be seen as a union of the syntactic diatheses. In XMG, we define this union of syntactic configurations as a disjunction of diatheses. For instance, the transitive *n0Vn1* class (see the definition below) is defined as a disjunction of its active form, of its passive and of its agentless passive. This simple rule lets us capture the 39 syntactic configurations of the transitive family, the canonical contexts of which are exemplified in (3.1)[3].

$n0Vn1 \rightarrow n0Vn1DiathesisActive \lor n0Vn1DiathesisPassive \lor n0Vn1DiathesisShortPassive$[4]

(3.1) a. Yossarian took another pill.
       b. Another pill was taken by Yossarian.
       c. Another pill was taken.

(3.1a) instantiates the canonical argument realisation with an active verb, (3.1b) is an example of passive diathesis with canonical argument and (3.1c) illustrates a passive verb form with a canonical subject without a *by*-agent.

---

[3]For the full list of syntactic alternations within each of the 57 implemented verb families, along with examples, the reader is referred to appendix A

[4]We use the notions of *short passive, agentless passive* and *passive without by-phrase* interchangeably.

3.1.2 Defining diathesis alternations

This level of abstraction is concerned with alternative ways of redistributing predicate-arguments among syntactic functions. This means that here we have to specify what the syntactic configurations of semantic roles are when combined with an active or a passive verb. The examples in 3.2 illustrate that the semantic roles in $see(yossarian, chaplain)$ correspond to two different surface realisations: in (3.2a) the agent is encoded as a subject and the patient as a direct object, whereas in (3.2b) the patient takes the position of a subject and the agent of an object of a *by*-phrase.

(3.2) a. Yossarian saw the chaplain. (active)

b. The chaplain was seen by Yossarian. (passive with *by*-phrase)

Therefore, the diathesis variant is formally defined as a conjunction of the relevant syntactic functions classes and a verb form as follows:

$n0Vn1DiathesisActive \rightarrow Subject \wedge Object \wedge ActiveVerb$

$n0Vn1DiathesisPassive \rightarrow Subject \wedge PassiveVerb \wedge ByAgent$

$n0Vn1DiathesisShortPassive \rightarrow Subject \wedge PassiveVerb$

It is worth mentioning that in our approach, the so called 'erasing' (i.e., agentless passive) is treated as an alternative of the passive with *by*-phrase and of the active form, all of them being of equal rank. Hence, we do not need to derive the agentless passive from the passive with *by*-agent by 'erasing' the *by*-phrase. By using a disjunction of alternatives, the sequencing problems in the derivation are thus easily overcome.

3.1.3 Defining syntactic functions

Now that the relevant predicate arguments are identified, their alternative positions at the phrase structure level must be specified. The syntactic functions are regarded as generalisations that characterise the set of variant linear orders of verbal dependents – they appear either in their canonical position or extracted to the left of the clause. The analysis adopted in XTAG is a GB movement analysis where the moved element is related to the empty position ϵ by means of the coindexing $\langle trace \rangle$ feature. Since XMG does not impose any priority of operations, a syntactic function is defined as a disjunction of alternative predicate arguments.

$Subject \rightarrow CanonicalSubject \vee WhSubject \vee RelativeSubject$

$Object \rightarrow CanonicalObject \vee WhObject \vee RelativeObject$

$ByAgent \rightarrow CanonicalByAgent \vee WhByAgent \vee WhAgent\ of\ by\text{-}phrase$

Following this definition, the abstraction of a syntactic function expresses whether it is in its prototypical case or has moved to an unbounded distance before the verb. For instance, the *Object* class encapsulates the following realisations:

(3.3) a. Milo served him **broiled Maine lobster**. (CanonicalObject)

b. **What** did Milo serve him? (WhObject)

c. **the broiled Main lobster** which Milo served him (RelativeObject)

3.1.4 Defining tree fragments

The last abstraction level is concerned with capturing shared partial grammatical structures. As explained above, the elementary building blocks of a grammar are fragmentary tree descriptions associated with a unique name and defined in classes. The XMG language allows us to factor out tree fragments by organising them in an inheritance hierarchy. The inheritance is controlled by the `export` and `import` declarations. For instance, in (3.4) the class *SubjectArgument* exports three node identifiers: xS, $xSubj$ and xVP, thus making them available for further access. Then, when the *SubjectArgument* class is imported into the *CanonicalSubject* class, the daughter inherits the information from its mother class as if locally defined. Then, this class adds further description on the $xSubj$ by constraining its category to an NP. The tree on the right-hand side illustrates the final tree description after the inheritance had taken place.

(3.4)
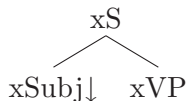
class SubjectArgument	class CanonicalSubject	*Resolved Tree*
declare xS xSubj xVP	import SubjectArgument	
export xS xSubj xVP		
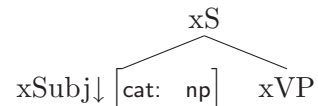
It should be mentioned that the number of required tree descriptions grows proportionally to the size of the grammar. Using an inheritance hierarchy we avoid specifying redundant information and thus, reduce the time and space for grammar implementation.

The table below illustrates the number of classes that were implemented in the XMG-based grammar. With the total amount of 485 classes we managed to describe 57 families and approximately 990 syntactic configurations.

Abstraction level	Nb of classes
Tree families	146
Diathesis alternation	106
Syntactic functions	32
Tree fragments	201
TOTAL	485

3.2 Tree Families

Up to this point we introduced a possible methodology for grammar implementation which takes advantage of generalised descriptions by abstracting in four dimensions. It is time now to put this into practice and to discuss the actual grammar development for each of the abstraction levels.

The notion of a tree family was introduced by [Abeillé, 1991] to group together the set of elementary trees selected by verbs of a given subcategorisation frame. This set encapsulates all possible syntactic transformations of a predicate and its arguments.

In this section, we will present the tree families that were implemented in the metagrammar

system. It should be reminded that the overall organisation of verb classes followed the linguistic decisions made by [XTAG-Research-Group, 2001].

Notational conventions The trees are named according to the abbreviation principle nX and pnX, where n stands for a nominal argument, X designates the conventional number of a semantic argument, and pn refers to a nominal object introduced by a preposition. In the same manner, *byn* and *ton* indicate an object of a *by*-phrase and *to*-phrase, respectively. The anchor of a tree is marked by an uppercase letter: V for verbal, A for adjective, etc.

3.2.1 Intransitive verb families

The intransitive verb families are selected by full-meaning verbs whose subcategorisation frame has one position for a subject argument. Adjuncts can freely adjoin in without influencing the grammaticality of a sentence. Strictly speaking, these verbs do not undergo any passivisation. [XTAG-Research-Group, 2001] groups intransitive verbs with nominal arguments in four distinct families to account for additional selectional restrictions of the lexical items. They can handle the contexts listed in below.

(3.5) a. Yossarian was waiting   (n0V)

      b. They stood up slowly   (n0Vpl)

      c. Bill ventured into the cave   (n0Vpn1)

      d. Milo grew flustered   (n0Va1)

(3.5b) represents verbs coanchored by a particle. The meaning of a verb-particle pair can not be determined by decomposing it into two stand-alone lexical items. Therefore, in the lexicon they should be specified within one lexical entry. The class in (3.5c) is selected by verbs that are not strictly intransitive in the sense that they accept passivisation. Verbs belonging to class *n0Va1* can appear either as bare intransitives or can take an adjective complement. These differences are formalised as follows:

$n0V \rightarrow dian0Vactive$

$n0Vpl \rightarrow dian0Vplactive$

$n0Vpn1 \rightarrow dian0Vpn1active \lor dian0Vpn1passive \lor dian0Vpn1shortpassive$

$n0Va1 \rightarrow dian0Va1active$

3.2.2 Transitive verb families

Verbs whose subcategorisation frame has two open slots, one for a subject and one for an object complement, belong to the transitive verb family group. There are no restrictions on the transitive argument if it is rooted in an NP – it can vary from a simple noun phrase to a complex relative structure with long distance dependencies, gerunds or NP's with embedded clauses. This group handles the four contexts as follows:

(3.6) a. Orr **likes** apple cheeks   (n0Vn1)

      b. Major Major **gave up** the whole thing

Major Major **gave** the whole thing **up**   (n0Vpln1)

   c. Sergent Towser **thinks of** the dead man in Yossarian's tent   (n0VPn1)

   d. The ice **melted**   (En1V)

(3.6b) is an illustration of verb-particle pairs whose meaning interpretation is non-compositional. These verbs are thus anchored both by the verb and by the particle. The position of the particle is not restricted – it can appear either before the direct object or after it, and this property of the verb-particle pairs distinguishes them from verbs selecting a prepositional phrase. The class exemplified in (3.6c) is selected by verb-preposition pairs with non-compositional interpretation and thus, the family is defined as being anchored by two lexical items. The ergative verbs (3.6d) are analysed as a subset of the transitive family and therefore they select both *n0Vn1* and *En1V* families. They possess the peculiarity of undergoing an ergative transformation known as a causative/inchoative transformation. This process consists in raising the object into a subject position as the examples below illustrate:

- The sun melted the butter   (n0Vn1)

- The butter melted   (En1V)

We saw that the intransitive families are defined by means of their active diathesis. Since transitive verbs have the general property of undergoing passivisation, their formal definition adds to the active diathesis two additional descriptions to account for the passive alternation with and without a *by*-agent:

n0Vn1 → dian0Vn1active ∨ dian0Vn1passive ∨ dian0Vn1shortpassive

n0Vpln1 → dian0Vpln1active ∨ dian0Vpln1passive ∨ dian0Vpln1shortpassive

n0VPn1 → dian0VPn1active ∨ dian0VPn1passive ∨ dian0VPn1shortpassive

En0V → diaEn1Vactive

3.2.3 Ditransitive verb families

The group of ditransitives includes those verbs which subcategorise for a subject argument and two complement arguments. From a semantic point of view, the complements correspond to the theme (which in the monotransitive family is realised as an object) and to the goal/beneficient. On one hand, the goal can take the syntactic function of an indirect object without a preposition, in which case it is located before the direct object. On other hand, it can be realised as an prepositional phrase positioned after the direct object. The ditransitive families that were implemented cover the four contexts as follows:

(3.7) a. Bill **baked** Lisa a cake   (n0Vn2n1)

    b. Bill **baked** a cake **for** Lisa   (n0Vn1Pn2)

    c. Bill **puts** the cake on a table   (n0Vn1pn2)

    d. Bill **opens up** Lisa a bank account

      Bill **opens** Lisa **up** a bank account

      Bill **opens** Lisa a bank account **up**   (n0Vpln2n1)

In order to account for the different subcategorisation structure of the so called dative shifts (3.7a and 3.7b), the verbs undergoing that alternation select both the *n0Vn2n1* and *n0Vn1Pn2* fam-

ilies. The former is anchored by the verbal lexical item, whereas the latter imposes a constraint on the preposition by making it its coanchor. Verbs which do not undergo that alternation select only for a noun phrase followed by a prepositional phrase without any restrictions on the preposition (3.7c) (hence, it is defined as a substitution node). The prepositional phrase contributes to the grammaticality of the sentence and therefore, we can not model it as an adjunct but as an obligatory complement in the subcategorisation frame. This is exemplified in the following sentences:

- Bill put a cake on the table vs.

- *Bill put a cake

Verb-particle pairs subcategorising for two noun phrase complements select the family instantiated in (3.7d). The tree is anchored by both the verbal item and the particle, and the particle movement is the property that distinguishes between *n0Vn1Pn2* and *n0Vpln2n1*.

The ditransitive verb families are formalised as a disjunction of the alternations they undergo: active – passive – agentless passive. The complexity of ditransitive verb-particle pairs might be the reason why they are not attested in passive:

n0Vn2n1 → dian0Vn2n1active ∨ dian0Vn2n1shortpassive ∨ dian0Vn2n1passive

n0Vn1Pn2 → dian0Vn1Pn2active ∨ dian0Vn1Pn2shortpassive ∨ dian0Vn1Pn2passive

n0Vn1pn2 → dian0Vn1pn2active ∨ dian0Vn1pn2shortpassive ∨ dian0Vn1pn2passive

n0Vpln2n1 → dian0Vpln2n1active

3.2.4 Verb families with a sentential argument

Lexical items selecting for a sentential argument include it as an S node in their elementary tree. It is treated the same way as a phrasal complement – it can be marked for adjunction and for substitution, and it can form either an initial or an auxiliary tree. The type of the sentential argument (indicative, infinitive, subjunctive, gerundive or base) and the distribution of complementisers are correlated since the complementiser itself restricts the type of its embedded clause. See for instance the following examples:

(3.8)   a. Appleby wondered whether to remove the flies in his eyes.
        b. Appleby wondered whether he had removed the flies in his eyes
        c. Appleby said that he removed the flies in his eyes.
        d. *Appleby said that to remove the flies in his eyes.

The complementisers in (3.8a) and (3.8b) do not impose any restrictions on the type of the clause: no matter if it appears as an indicative or an infinitive, the sentence is still grammatical. This is not the case in (3.8c) and (3.8d), which show that *that* is bound to an indicative clause and its coocurrence with an infinitive makes the whole construction ungrammatical. The distribution of the complementiser with respect to the type of the clause is handled by a fine-grained feature system. The type of the clause is defined for each verb in the lexicon.

Our analysis handles the following six cases:

(3.9)   a. Yossarian does not **want** to fly many missions   (n0Vs1)

       b. The colonel forces Yossarian to fly many missions   (n0Vn1s2)

       c. That Milo sold the cotton **impressed** Yossarian   (s0Vn1)

       d. That Milo buys Polish sausage **matters** to the syndicate   (s0V)

       e. That he is in love with the chaplain **occurred to** Yossarian   (s0Vton1)

       f. That Bill drives a Jaguar makes Lisa angry   (s0Vs1)

(3.9a) exemplifies the case of a verb subcategorising for a subject argument and for an S complement whereas the verb in (3.9b) selects a subject argument, an object argument and an S complement. Some verbs allow for both indicative and infinitive embedded clauses whereas others restrict the type of the clause. The S node of the prototypical tree is a foot node so that it attaches to an elementary tree rooted in S of a permissible clause. Its type is constrained within the lexicon and the sentential complement of the declarative tree is marked for adjunction. In so doing, one can handle long-distance dependencies where the extracted element of the sentential complement appears at an unbounded distance before the verb. The sentence in Figure 3.2 is derived from the auxiliary tree and the initial tree in Figure 3.1, splicing the latter in the S_r node to insert the auxiliary tree, as shown in the graphical representation below:
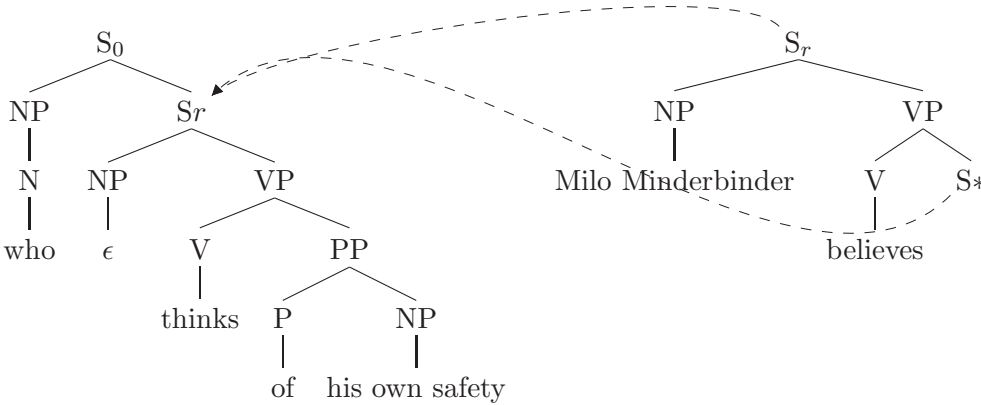
Figure 3.1: Trees *Who thinks of his own safety* and *Milo Minderbinder believes*
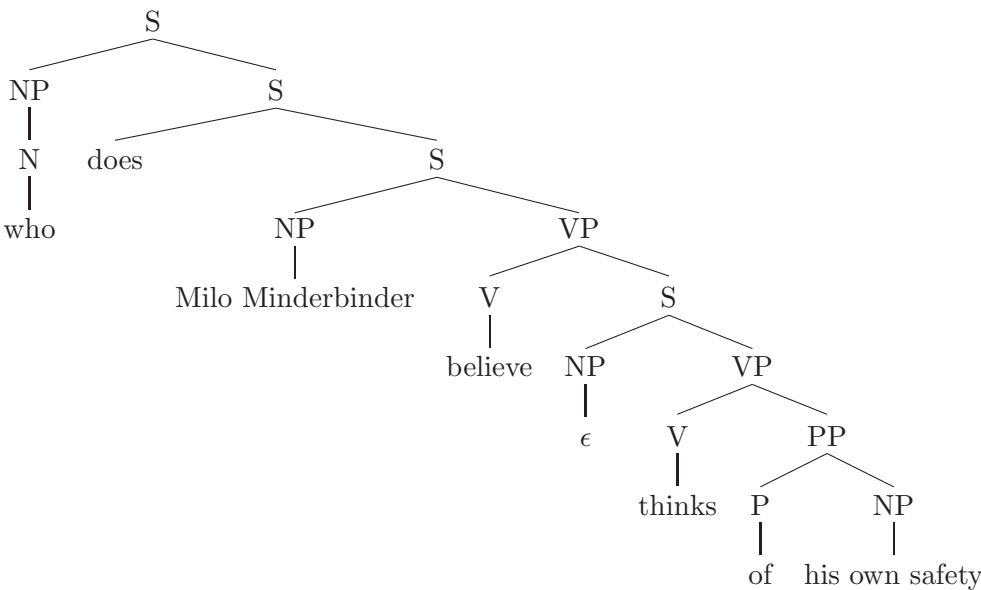
Figure 3.2: Tree for *Who does Milo believe thinks of his own safety*

This process illustrates the increased power of TAG – one can add recursively as many auxiliary trees as wanted so that a predicate argument appears at a significant distance before the verb it

is selected by. Adjunction operation permits this to happen without interfering with the local predicate-argument information. This means that in the above example *who* keeps the properties of a nominal subject argument of *think* although it was positioned far from its governor.

Sentential subjects appear in three contexts (3.9c, 3.9d, 3.9e and 3.9f) where *s0V* is selected only by the verb *matter*; *s0Vn1* requires an additional NP complement; *s0Vton1* includes an argument headed by a *to*-preposition and *s0Vs1* is selected only by the verb *make* and is should cooccur with a small clause sentence. The type of the sentential subject clause can be indicative (with an obligatory complementiser) or infinitive (the complementiser is not required).

The formal definition accounts for their active diathesis. Only one family appears with a passive verb form:

$n0Vs1 \rightarrow dian0Vs1active$

$n0Vn1s2 \rightarrow dian0Vn1s2active \lor dian0Vn1s2shortpassive \lor dian0Vn1s2passive$

$s0Vn1 \rightarrow dias0Vn1active$

$s0V \rightarrow dias0Vactive$

$s0Vton1 \rightarrow dias0Vton1active$

$s0Vs1 \rightarrow dias0Vs1active$

3.2.5 Families with cleft arguments

English cleft sentences are defined as initial trees with four compulsory elements:

- the dummy *it* subject

- the main verb *be*

- the cleft component

- the sentential clause

The declarative trees are anchored by both the dummy subject and by the main verb. The category of the cleft element is either a noun phrase (3.10a), a prepositional phrase (3.10b) or an adverb (3.10c). The following sentences exemplify the different categories of the cleft element:

(3.10) a. It is Colonel Cathcart that never hesitated to volunteer his men for any target
   available (ItVn1s2)
   b. It is on Pianosa island that Milo runs M. & M. enterprise (ItVpn1s2)
   c. It was reluctantly that Yossarian tasted the chocolate-covered cotton (ItVad1s2)

Cleft sentences are formally defined by invoking the classes of their active diathesis:

$ItVn1s2 \rightarrow diaItVn1s2active$

$ItVpn1s2 \rightarrow diaItVpn1s2active$

$ItVad1s2 \rightarrow diaItVad1s2active$

3.2.6 Light verb families

Verb-noun pairs whose interpretation can not be computed from the meaning of their elements belong to the group of light verbs. Structurally, they resemble the tree descriptions for transitive and ditransitive verbs, with the only difference that the noun phrase can not be extracted and that a verbal item should cooccur with a specific noun. Thus, we differentiate between constructions as follows:

- Yossarian **took a bite** vs. Yossarian **took** a blanket

- Milo **gave** Yossarian **a promise** vs. Milo **gave** Yossarian a chocolate-covered cotton

In the left-hand-side examples the verb and the noun have their own semantic content contributing to a joint predication and therefore both nodes appear in the lexicon (hence, the tree is anchored by the verb and coanchored by the noun), whereas the complement in the right-hand side is marked for substitution. It should also be noted that the morphological alternations of the noun in the light verb families are restricted, in contrast to the noun complements in mono-transitive and ditransitive classes, e.g.:

(3.11)   a. The colonel took his leave
         b. The colonel took a leave
         c. *The colonel took the leave
         d. *The colonel took the leaves

In the proposed formalism, each morphological variant is defined as a coanchor node of the light verb it cooccurs with. The ditransitive light verbs participate in the dative shift alternation where the beneficient is realised either before the noun anchor as an indirect object (3.12b) or after it as an object headed by the *to*-preposition (3.12c).

The light verb families that were implemented cover the following contexts:

(3.12)   a. Lisa **took** a short **nap**   (n0lVN1)
         b. Bill **gave** Lisa a suspicious **look**   (n0lVN1Pn2 non-shifted)
         c. Bill **gave** a suspicious look **to** Lisa   (n0lVN1Pn2 with PP shift)

The light verb families are formalised as a disjunction of active – passive alternations (where applicable). The distribution of agentless passive, however, is severely restricted to trees anchored by a gerund verb. The formal definition is thus as follows:

$$n0lVN1 \rightarrow dian0lVN1active$$

$$n0lVN1Pn2 \rightarrow dian0lVn2N1active \lor dian0lVn2N1passive \lor dian0lVn2N1NPgerundshortpassive$$
$$\lor \; dian0lVN1Pn2active \lor dian0lVN1Pn2passive \lor dian0lVN1Pn2NPgerundshortpassive$$

3.2.7 Verb family with equative BE

To this group belong the verb *be* which subcategorises for two NP arguments. Unlike predicative verbs (e.g., *seem, look, etc*) and copula, the verb *be* has a non-predicative usage by which the two NP arguments are equated and hence, can be used interchangeably. See for instance the following examples:

(3.13) a. The bombardier **is** this man over there   (n0BEn1)

      b. This man over there **is** the bombardier   (n0BEn1)

      c. *My uncle seems this man over there

The subject and the complement in (a and b) can exchange positions without this to influence the grammaticality of the sentence. This is, however, not the case in (c) where swapping the places of the NP arguments of a predicative verb leads to ungrammaticality. It should also be mentioned that the equative *be* can be inverted which makes it behave as an auxiliary verb.

The equative *be* family appears in active and thus, its formal definition calls the class of its active diathesis as follows:

$n0BEn1 \rightarrow dian0BEn1active$

3.2.8 Small clause families

To the small clause group belong syntactic structures anchored not by a verb but by another part of speech: an adjective, a noun or a preposition. These constructions are known as 'small clauses' since they lack an overt verb but at the same time an implicit verb (the copula *be*) can be added. For instance, in the sentence *Most people consider [JK Rowling a prolific **author**]* the minimal fragment (*[JK Rowling a prolific **author**]*) forms a stand-alone predicate clause with its own semantics being *JK Rowling is a prolific author* and with its own predicate-argument structure (in this case the predicate in bold selects an NP object). Despite that one might be tempted to call a small clause an S complement, a major property for their differentiation is the lack of a complementiser in small clauses. Adding a COMP makes the sentence ungrammatical as in the following example: *Most people consider that/if/whether JK Rowling a prolific author.*

Overall the grammar contains 20 small clause families. The distinction between them is drawn upon the anchor category and the selected predicate-arguments. A showcase of small clauses anchored by a noun are the examples that follow:

(3.14) a. Most people consider [JK Rowling a prolific **author**]   (n0N1)

      b. To fly many missions is **insanity**   (s0N1)

      c. These were **accusations** that Orr ate the chestnuts   (n0N1s1)

      d. The innovation was **thanks to** the chaplain   (n0NPn1)

      e. That he ran a marathon was **thanks to** his spirit   (s0NPn1)

      f. The chaplain was **in charge of** Yossarian   (n0PNaPn1)

      g. That Bill ran a marathon was **on account of** his spirit   (s0PNaPn1)

The families illustrated in (a) and (b) selects a subject argument which can be realised either as an NP or as an S complement in infinitive (without complementiser) or in indicative (with an obligatory complementiser). There exists also predicate nouns which should cooccur with an NP subject argument and an S complement (again in indicative or in infinitive mode) – they are exemplified in (c). Noun-preposition pairs with a decompositional meaning are differentiated in separate families: the anchor node should appear either with an NP subject (d) or with an indicative or infinitive sentential subject (e). The anchor node can be a more complex phrase of a noun and two prepositions selecting an NP subject (e) or a sentential subject (f).

Since small clauses lack the basic properties of verb families, like passivisation, their formal

definition takes into account the union of the participating syntactic function. More formally, the definition of a small clause tree family calls the class of their one verbless 'diathesis' as follows:

n0N1 → dian0N1

s0N1 → dias0N1

n0N1s1 → dian0N1s1

n0NPn1 → dian0NPn1

s0NPn1 → dias0NPn1

n0PNaPn1 → dian0PNaPn1

s0PNaPn1 → dias0PNaPn1

3.2.9   Verb families with exceptional case marking

There exists a certain class of verbs which are structurally the same as the *n0Vs1* family, i.e. they subcategorise for an NP subject argument and an S complement. However, these verbs are known as exceptional case marking (ECM) in the sense that they assign an accusative case to the subject of the S complement. In the Latin grammar, this construction is known as an 'accusativus-cum-infinitivo' and it consists in assigning accusative case on the subject of clause. This can be illustrated with the following example:

(3.15) *Julius   Claudium   ridere   vult.*
       Julius   Claudius.ACC   to-laugh.INF   wants.3sg
       Julius wants Claudius to laugh.

The subject of the embedded predicate clause appears in accusative assigned by the main verb. This phenomenon of a subject with object marking has left traces in modern English too and this is the syntactic motivation for differentiating a separate verb family of verbs which assign an exceptional case to the subject of the S complement. In most cases, these are verbs of perception and of wishing (e.g., *believe, expect, consider, want* etc). For instance, in the sentence that follows:

(3.16) a. Lisa **wants** him to sell the house   (Xn0Vs1)
       b. I **saw** her fall   (Xn0Vs1)

the subject of the embedded predicate clause (*him*) appears in accusative. The sentential complement should be either infinitive form (a) or a 'bare' infinitive (b). The formal definition of the ECM family is the same as the one for *n0Vs1* verbs: reference to their active diathesis:

Xn0Vs1 → diaXn0Vs1active

As one can expect, the declarative trees are again auxiliary trees to permit adjoining in the S node of an initial tree and thus, maintain long-distance dependencies. The operation is the same as the one in shown in Figure 3.2.

3.2.10 Idioms

Frozen expressions which are multiply anchored by a phrase of more than one lexical item belong to the class of idioms. These expressions might have the same structure as transitive verbs (e.g., *bury the hatchet*) or verbs with a PP (e.g., *make a short work of*). Their interpretation, however, cannot be extracted from the meaning of each lexical item and this makes them non-compositional. Depending on the category of the anchor nodes, the grammar differentiates eight idiom verb classes. These are as follows:

(3.17)   a. Bill **bit the dust**   (n0VDN1)
        b. Lisa **sang a different tune**   (n0VDAN1)
        c. Lisa **cries wolf**   (n0VN1)
        d. Lisa **cries bloody murder**   (n0VAN1)
        e. Lisa **makes a big deal of** the accusations   (n0VDAN1Pn2)
        f. Lisa **makes short work of** the cake   (n0VAN1Pn2)
        g. Bill **looks daggers at** the detective   (n0VN1Pn2)
        h. Bill **broke the news to** his wife   (n0VDN1Pn2)

The idiom family instantiated in (3.17a) is selected by complex phrases anchored by a verb, determiner and a noun which should cooccur with an NP subject argument. Adding an adjective anchor to this description creates another idiomatic family as illustrated in (3.17b). Idioms anchored by verb-noun pairs are shown in (3.17c) and idioms anchored by verb-adjective-noun phrases are exemplified in (3.17d). Further on, each of this syntactic configurations is possible also in the context of a PP where the preposition is also an anchor node. Next to the NP subject, these complex phrases require also an object of the PP. They are illustrated in (3.17e [anchored by a verb, determiner, noun and a preposition], (3.17f)[anchored by a verb, adjective, determiner, noun and a preposition], (3.17g) [anchored by a verb, noun and a preposition] and (3.17h [anchored by a verb, determiner, noun and a preposition])).

Idioms are formally defined as a disjunction of their active, passive and agentless passive diathesis. Verb families with a $Prep$ anchor node undergo also an outer passive and an outer passive without a *by*-agent.

$n0VDN1 \rightarrow dian0VDN1active \lor dian0VDN1passive \lor dian0VDN1passive$

$n0VDAN1 \rightarrow dian0VDAN1active \lor dian0VDAN1passive \lor dian0VDAN1passive$

$n0VN1 \rightarrow dian0VN1active \lor dian0VN1passive \lor dian0VN1passive$

$n0VAN1 \rightarrow dian0VAN1active \lor dian0VAN1passive \lor dian0VAN1passive$

$n0VDAN1Pn2 \rightarrow dian0VDAN1Pn2active \lor dian0VDAN1Pn2passive \lor dian0VDAN1Pn2passive$

$n0VAN1Pn2 \rightarrow dian0VAN1Pn2active \lor dian0VAN1Pn2passive \lor dian0VAN1Pn2passive$

$n0VN1Pn2 \rightarrow dian0VN1Pn2active \lor dian0VN1Pn2passive \lor dian0VN1Pn2passive$

$n0VDN1Pn2 \rightarrow dian0VDN1Pn2active \lor dian0VDN1Pn2passive \lor dian0VDN1Pn2passive$

3.2.11   Resultative tree families

Complex verb-adjective or verb-preposition predicates belong to the resultative tree families. They are again non-compositional and thus, these trees are multiply anchored by a verb plus adjective or by a verb plus a preposition. They form a complex predicate with one semantics in the sense that the secondary predicate signifies the 'result' of the verb. Syntactically, the verb can be transitive or ergative followed either by an adjective or by a PP predicate. This is the criterium that differentiates the following four resultative families:

(3.18)   a. The detective **beat** Bill **dead**   (Rn0Vn1A2)

        b. Bill **ran** his shoes **into** pieces   (Rn0Vn1Pn2)

        c. The snow **froze solid**   (REn1VA2)

        d. The snow **is melting into** a puddle   (REn1VPn2)

(3.18a) exemplifies a resultative tree where the complex predicate is a verb-adjective pair and the verb is transitive. The transitive verb can be coanchored by a preposition heading an NP – this is illustrated in (3.18b). Ergative verb-adjective pairs are presented in (3.18c) and ergative verb-prepositional pairs are shown in (3.18d). The formal definition of resultative families depends on the type of the verb: transitive verbs undergho passivisation and thus, they are formalised as a disjunction of their active, passive and agentless diathesis, whereas ergative verbs have only active forms and the resultative family with an ergative verb are defined in terms of their active diathesis. More formally, this was specified as follows:

$Rn0Vn1A2 \rightarrow diaRn0Vn1A2active \lor diaRn0Vn1A2passive \lor diaRn0Vn1A2shortpassive$

$Rn0Vn1Pn2 \rightarrow diaRn0Vn1Pn2active \lor diaRn0Vn1Pn2passive \lor diaRn0Vn1Pn2shortpassive$

$REn1VA2 \rightarrow diaREn1VA2active$

$REn1VPn2 \rightarrow diaREn1VPn2active$

Summing up   In this section we presented the groups of families that were implemented in the TAG grammar. We saw that each family can be formalised either as a disjunction of its active – passive – agentless passive diatheses or as a reference to a tree description. The general organisation and distinction between families is inspired by the [XTAG-Research-Group, 2001] project. The full grammar of verb classes amounts to 57 families.

3.3   Diathesis Alternations

Along the lines of lexical-functional grammar [Bresnan and Kaplan, 1982], where active/passive alternatives are not a matter of transformation but a functional relation between two verbs, and where this relation is defined in terms of lexical rules, for the purpose of the current grammar active/passive alternatives are considered different realisations of the same predicate-argument structure. The distinct argument realisations are captured by means of disjunction instead of elaborate linguistic rules. This approach is consistent with the approach of [Levin, 1993] for getting one semantic interpretation and n-number of possible surface alternatives.

This level of abstraction is concerned with the distribution of semantic arguments among syntac-

tic functions and, more specifically, the possible realisations of a predicate-argument structure. In the previous chapter we outlined what diathesis alternations a verb undergoes. We now define the tree descriptions encapsulated in an active/passive diathesis. Passive constructions are treated as a syntactic alternative of the declarative subcategorisation frame and they are accompanied by differently marked morphological form. Within this analysis, passive trees are restricted by their $\langle mode = ppart \rangle$ value, which permits only auxiliary trees with $\langle rmode = ppart \rangle$[5] and $\langle passive = + \rangle$ to adjoin in. In the proposed framework, the active and passive structures are formally defined as a conjunction of argument realisations and a verbal skeleton.

3.3.1   Intransitive diathesis

Strictly speaking, intransitive families do not undergo any passivisation, and therefore their formal definition specifies the mapping between the predicate arguments and surface realisation of grammatical functions. More precisely, in the intransitive family the NP_0 argument is encoded as a subject, and the NP_1 (*n0Vpn1* family) can be realised as an object of a preposition or as a subject. The diathesis of the intransitive families can be unfolded as follows:

n0VDiathesisActive → *active* ∧ *Subject*

n0Vpn1DiathesisActive → *active* ∧ *Subject* ∧ *PPObject*

n0Vpn1DiathesisPassive → *passive* ∧ *Subject* ∧ *PPObject* ∧ *ByAgent*

n0Vpn1DiathesisShortPassive → *passive* ∧ *Subject* ∧ *PPObject*

This formalisation accounts for the following contexts:

(3.19) a. Lisa sleeps   (active)
      b. Bill slogged through traffic for almost 2 hours   (active)
      c. The traffic was slogged by Bill for almost two hours   (passive)
      d. The traffic was slogged for almost two hours   (agentless passive)

3.3.2   Monotransitive and ditransitive diathesis

The transitive family is without any doubt the one that is most manifestly subject to active – passive alternations. In the last section we defined the transitive class as a disjunction of active, passive and agentless passive. Now, in order to identify the distribution of predicate arguments in each of the diatheses cases, we include the following rules:

n0Vn1DiathesisActive → *active* ∧ *Subject* ∧ *Object*

n0Vn1DiathesisPassive → *passive* ∧ *Subject* ∧ *ByAgent*

n0Vn1DiathesisShortPassive → *passive* ∧ *Subject*

These rules suffice to express the fact that the mapping between thematic roles and grammatical functions is not one-to-one. To be more precise, an NP_0 agent can be assigned the grammatical function of a subject or that of an object of a *by*-phrase. In the same way, the NP_1 argument can be encoded as a direct object, or it can be raised to a subject position. In the perspective

---

[5]inspired by Laura Kallmeyer (pc)

of the monotonicity of the description language, passive variants are not generated by means of transformational derivation but are handled as an alternative, which is ranked equally as the distribution of the declarative subcategorisation frame. In the same perspective, the treatment of the agentless passive gets away with the 'erasing' and, thus, allows for the derivation of a short passive without an intermediate transformation level.

Following the above rules, the grammar can generate the sentences in (3.20):

(3.20) a. Doc Daneeka accepted the explanation   (active)

       b. The explanation was accepted by Doc Daneeka   (passive)

       c. The explanation was accepted   (agentless passive)

The assignment of grammatical functions in the ditransitive family resembles the assignment in the monotransitive family, with one main difference – ditransitive verbs select for an additional NP_2 argument with the semantic role of a beneficient. In the surface structure this argument can be realised as in indirect object (as in 3.21a) or can be fronted to the subject position (as in examples 3.21b and 3.21c):

(3.21) a. Bill asked Lisa a question   (active)

       b. Lisa was asked a question by Bill   (passive)

       c. Lisa was asked a question   (agentless passive)

These contexts are formalised as follows:

$n0Vn2n1DiathesisActive \rightarrow active \wedge Subject \wedge Object \wedge IObject$

$n0Vn2n1DiathesisPassive \rightarrow passive \wedge Subject \wedge Object \wedge ByAgent$

$n0Vn2n1DiathesisShortPassive \rightarrow shortpassive \wedge Subject \wedge Object$

Summing up   In this section we illustrated the mapping between predicate arguments and syntactic grammatical functions in the surface structure. It has been shown that this mapping is not one-to-one, and that it is handled by means of equally ranked production rules. The transformational derivations are replaced by alternation of active – passive – agentless passive diatheses, avoiding any sequencing problems. The description language allows for abstracting over each diathesis by defining it as a set of argument realisations and an active/passive verbal morphology.

3.4   Syntactic Functions

In the perspective of meaning-text theory (see [Iordanskaja and Mel'cuk, to appear]) a syntactic function is defined as a set of binary dependency relations between two lexical items: a governor g and its dependents d (represented as $g \rightarrow r \rightarrow d$), where the dependents correspond to clause arguments. A syntactically dependent tree has a unique top node, which does not depend on the word precedence. In most European languages the top node is a finite verb governing a set of surface syntactic functions. For instance, in the sentence *'Lisa sleeps'*, *sleeps* is the governor and *Lisa* is the dependent. The syntactic relation determines the type of the dependent, i.e., the subjectival relation presupposes a **subject** dependent, a direct-objectival relation presupposes a **direct object** dependent. In contrast to deep syntactic relations, surface functions are language

dependent and they can be established in an empirical manner.

However, if we define a syntactic function as the mere set of functional relations between two clausal elements, we are ignoring important information, namely the fact that a clausal component encapsulates syntactic constructions varying at the phrase-structure level. For the sake of a precise analysis, we have incorporated the different ways of surface realisation of the same grammatical function.

The elements that we operate with in this chapter are tree descriptions of a higher level, abstracting over various way of expressing a syntactic function. In the perspective of our description language, a syntactic function is defined as a disjunction of alternatives of equal status:

$$SyntFunction \rightarrow CanonicalSF \vee WhSF \vee RelativeSF$$

This abstract notation allows us to account for linear realisations of grammatical functions without imposing any order of application for their derivation[6].

More specifically, a syntactic function is declared in a separate class providing an exhaustive list of its variants. This includes a canonical realisation, wh-extraction, relative clause and gerund argument (i.e., an argument rooted in a noun phrase). The subject class can be realised as a *PRO subject* and as an *Imperative*, and in this case it preserves its subject agreement. The sentential subject can be realised either in its canonical position before the verb or wh-extracted to the front (in the latter case its structure equals the one of a wh-moved noun subject). The grammar does not differentiate between finite and non-finite sentential subject[7]: complementisers are defined as adjoining trees with features imposing restriction on the selected complementiser and clause type.

The subject definition below allows us to account for the different contexts in (3.22).

$$Subject \rightarrow CanonicalSubj \vee WhSubj \vee RelativeOvertSubj \vee RelativeCovertSubj \vee$$
$$ImperativeSubj \vee ProSubject \vee NPGerundSubject \vee DeterminerGerundSubject$$

$$Object \rightarrow CanonicalObj \vee WhObj \vee RelativeOvertObj \vee RelativeCovertObj \vee$$
$$DeterminerGerundObject$$

This description covers the following contexts:

(3.22) a. Orr took the chestnuts out   (canonical subject and canonical object)
       b. Who took the chestnuts out   (wh-moved subject and canonical object)
       c. What did Orr take out   (wh-moved object and canonical subject)
       d. the bombardier who took the chestnuts out   (relative overt subject)
       e. the bombardier that took the chestnuts out   (relative covert subject)
       f. Take the chestnuts out!   (imperative subject)
       g. Orr does not want [PRO to take the chestnuts out]   (PRO subject)
       h. Orr's taking out the chestnuts   (NP gerund subject)
       i. His taking out the chestnuts is bothering   (determiner gerund subject)

---

[6]This can be compared to the XTAG approach, which uses metarules. For instance, a wh-moved subject is produced by applying a wh-extraction rule to the canonical subject tree.

[7]in contrast to the French XMG grammar of [Crabbé, 2005]

Summing up   At the third level of abstraction, we defined syntactic functions as a disjunction of possible realisations in the surface structure. These generalisations account for the varying linear order of a function in the phrase structure tree – in a canonical, extracted or empty position. The precise precedence and dominance relations in a tree description are to be specified in the next section, where we discuss the building blocks of the grammar.

3.5   Tree Fragments

So far we have been working at a higher level abstracting over family alternations and syntactic functions. It is time to move to the gist of our implementation where we will define the building blocks needed for a core grammar of English verbs. The objects we will operate with are shared structures of grammatical information: features, verbal arguments and verbal spines. We begin by outlining briefly the way we represent tree fragments and the mechanisms for their combination. Then, we will concentrate on feature declaration, and more specifically we will see how to factore out common feature structures. At the end of this chapter we will illustrate an economical way of factoring out common information by representing a possible hierarchical organisation of verbal arguments and verb forms.

3.5.1   General ideas

A tree description consists of a set of tree node variables and of relations between these variables specifying precedence, dominance and equality. A tree fragment is specified in terms of node variables, precedence-dominance relations between the nodes, feature information and node colours, whereby the colouring principle (cf 2.3) imposes constraints on merging nodes. Figure (3.3) exemplifies a set of three independent tree fragments (a fragment for a canonical subject, adjective complement and an active verb form). The combination of these partial tree descriptions renders a fully saturated predicate argument, which is given on the right of the equality sign.
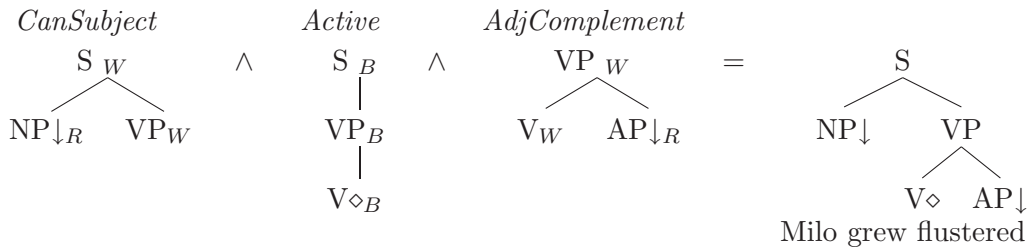
Figure 3.3: Combination of Tree Fragments

3.5.2   Factorising feature information

To ensure the well-formedness of the grammar, XMG allows for decorating tree descriptions with feature information. One way to do this is to specify a feature-value pair for each node in a tree fragment. This, however, leads to redundant and repetitive information. In order to better factorise feature structures, we define separate classes which carry information about the

distribution of a feature-value pair among tree nodes. The classes of relevant feature information are then inherited by using the `import` statement. Another way to factor out common pieces of feature structures is to declare them in a mother node within the hierarchy of tree fragments. The next paragraphs detail the features currently we use in the XMG-based grammar.

Subject agreement   This feature is provided to ensure agreement between the verb and the subject. The agreement rule is a complex feature which stipulates that the subject should agree in person and number[8] with the verb. The features of the lexical items are specified in the lexicon and they are propagated up the top feature of the verbal head. Agreement is performed in two steps so that we can insert an auxiliary tree in the VP node. In that case the root node of the adjoining tree equals the *top* feature of the spliced node and the foot node information is shared with the *bottom* feature of the split node. Graphically, this is represented as follows:
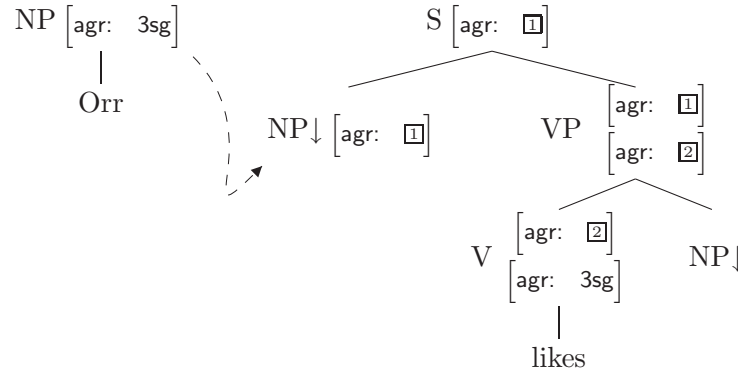
Figure 3.4: Subject Agreement

Movement agreement   Movement agreement is imported into classes where an element has been extracted from its canonical position to appear in verb-fronted position, in particular in the case of wh-extraction and relativisation. After movement, it is assumed that the extracted component leaves a trace and the pre-existing agreement between the noun phrase and the verb should be preserved. Compare for instance the following examples:

- $[Which\ word]_i$ does Milo reckon $[is]_i$ anathema to him? vs.
- $*[Which\ word]_i$ does Milo reckon $[are]_i$ anathema to him?

This is provided by equating the case and agreement values between the extracted argument and its trace:

$WH.top : \langle case \rangle = NP.top : \langle case \rangle$

$WH.top : \langle agr \rangle = NP.top : \langle agr \rangle$

Case assignment   The case analysis is drawn from the government and binding theory (GB) and the minimalistic program in the following two perspectives:

- As in GB, an abstract case is assigned by a number of case assigners like verbs and prepositions;

---

[8]We exclude gender since it is not applicable in an English grammar

- As in the minimalistic framework, a lexical item carries its own feature, which is checked in the process of unification;

Noun phrases are assigned an unambiguous case by their governing heads, that is verbs and prepositions. At the same time the process of combining elementary trees involves unification of feature values between nouns (which come with a case value predefined in the lexicon) and the case assigning category. The possible values of the $\langle case \rangle$ feature are **nom, acc, gen** and **none**.

1. Prepositional case assigner

   All prepositions assign an accusative case to the NP they govern and this is specified in the lexicon: $\langle assign - case \rangle = acc$. This value is propagated to the $\langle assign - case \rangle$ feature of the mother node of the prepositional phrase and it is directly linked by equation to the *top* $\langle case \rangle$ feature of the NP (3.5).
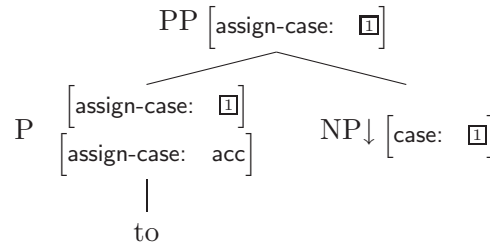
Figure 3.5: Prepositional Case Assigner

2. Verbal case assigner

   The case assigned to a subject is provided by linking the $\langle case \rangle$ feature of the NP substitution node and the *top* $\langle assign - case \rangle$ feature of the VP. Further on, this agreement is passed up the root node of the tree. For the sake of keeping the description as compact as possible, this information is added to the class, ensuring agreement between the subject and the verb. The case assigned to NP complements is factored out in a separate class ensuring that a verb selecting for one or two arguments assigns an accusative case to their sister nodes (3.6).
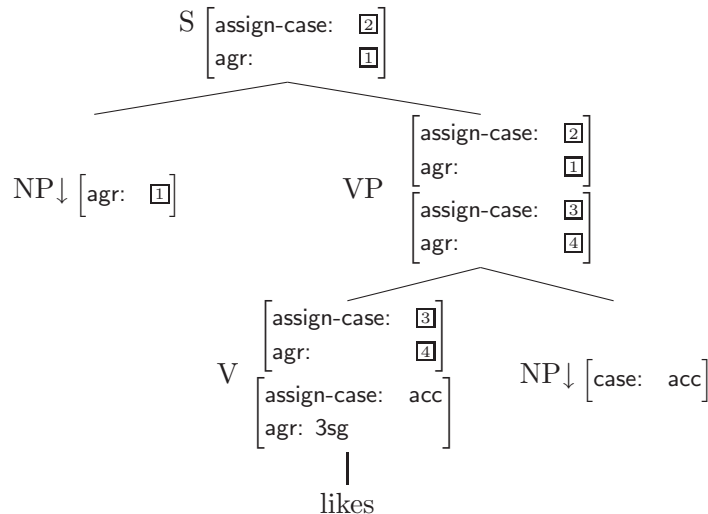
Figure 3.6: Verbal Case Assigner

3.5.3 Verb form

Within this section we will detail the tree fragments used as a basis for a factorised representation of verbal forms. Previously, we defined a diathesis alternation as a conjunction of syntactic functions and an active or a passive verbal form. The choice of the latter marks the different surface realisations in (3.23)

(3.23) a. Dunbar **dropped** his bombs hundreds of yards past the village.

b. The bombs **were dropped** hundreds of yards past the village.

[Bresnan and Kaplan, 1982] describe the process of passivisation as a lexical rule accompanied by changing the morphological form of the verb and redistributing the surface syntactic functions. In that perspective, the passive form is built by rules specifying the transformation of a flexed verb form into a participle and the movement of a subject into an object position. The formation of a *by*-object out of a subject is non-obligatory.

Since the description language we use does not permit any sequence order of operations, active and passive forms are modelled as alternatives of equal right whose feature-values differ in the terminal nodes (Figure 3.7). Auxiliary verbs are defined in separate elementary trees which adjoin to the VP node of the passive tree. Well-formedness of the derived passive tree is ensured by propagating the $\langle agr \rangle$ and $\langle mode \rangle$ values of the inflected auxiliary verb up the tree. Since the main passive verb does not select for any auxiliary, we add an additional $\langle rmode \rangle$ feature to the anchor node of the auxiliary tree, coindexed with the $\langle mode \rangle$ of the main verb.
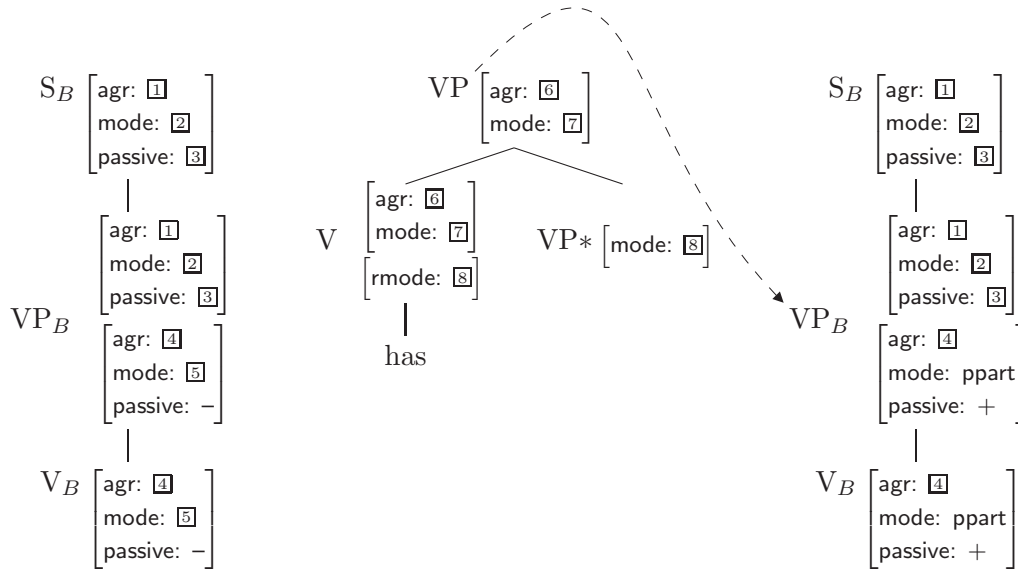
Figure 3.7: Active and Passive Verb Forms

To better factorise information, verb forms are organised in an inheritance hierarchy with one mother node and two daughters:
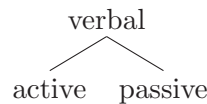
Figure 3.8: Verb Form Hierarchy

3.5.4   Verbal arguments

Having shown the organisation of the verbal forms, we move to the bigger group of verbal dependents. Here, the elementary trees are factored out in a hierarchical network, reducing the number of repetitive and redundant tree descriptions. The partitioning of our hierarchy is based on the linear order of predicate arguments (Figure 3.9): the *Subject Argument* class dominates the subject realised in front of the verb; noun complements placed at their canonical position on the right-hand side of the verb are grouped under the *Canonical Complement* node; arguments that have been moved from their canonical position are split into a separate *Extraction* class; arguments rooted in NP are differentiated in an additional *Gerund Argument* class; finally, cleft arguments are organised in a *Cleft Argument* group.

$$VerbalArgument$$

SubjectArg CanCompl ExtractedArg GerundArg CleftArg
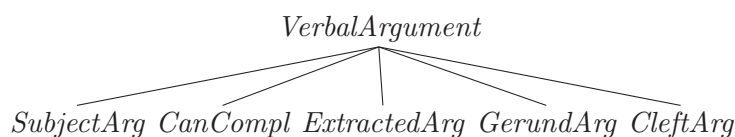
Figure 3.9: Major Tree Fragments

Subject arguments   Non-extracted subject arguments realised in their canonical position on the left hand side of the verb cover the contexts as follows:

(3.24)   a. **Major Major** never sees anyone in his office while he is in his office.

(Canonical nominal subject)

    b. **To force his men to fly more missions than everyone else** impressed Colonel Cathcart as the most tangible achievement for him.

(Canonical sentential subject)

    c. **A promise** was made to Milo by the colonel.   (Subject of a light verb)

    d. Aarfy was always trying [**PRO** to help Nately].   (PRO subject)

    e. [ϵ Come] to see me tomorrow.   (Imperative subject)

When the subject argument is of category **np** and is marked for substitution, it can be replaced by a nominal phrase of any complexity, including embedded relative clauses. Tree families selecting a sentential subject do not permit any extraction out of it, and therefore the S subject appears as a substitution node. The S subject licenses almost the same clause types as a sentential complement: indicative and infinitive with an extra constraint: an indicative clause must cooccur with a complementiser. Compare these examples:

(3.25)   a. That the lives of the enlisted men were only slightly more interesting than the lives of the officers disappointed Yossarian. vs.

    b. *The lives of the enlisted men were only slightly more interesting than the lives of the officers disappointed Yossarian.

Another difference between sentential subjects and sentential complements is that an S subject does not license gerundive, since subjects in gerund mode are analysed as noun phrases.

(3.24c) exemplifies a subject of a light verb family which behaves in the same way as a canonical nominal subject, with the difference that the NP subject node is anchored by a specific lexical item. The declarative subcategorisation frame of light verbs is of the form *[S NP$_1$ PP$_2$]*, and

in passive diathesis NP_1 is realised in a subject position. This motivates the specification of a subject of a light verb family.

(3.24d) and (3.24e) are surface-realised as empty elements with a distinct behaviour. The null *PRO* element appears in tenseless clauses (as an adjunct or infinitive) and it is assigned the feature $\langle case = none \rangle$. On the other hand, imperative subjects (filled by an ϵ) are treated as non-overt subjects with morphological behaviour: the subject position should agree with the verb $\langle agr = 2sg \rangle$ and the mode of the S node is set to imperative. Overt and negative imperatives are not handled. The complete hierarchy of subject arguments can be seen in Figure 3.10

$$SubjectArgument$$

$$CanSubj \quad CanSentSubj \quad LightVerbSubj \quad ProSubj \quad ImperativeSubj$$
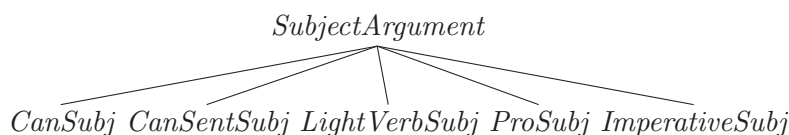
Figure 3.10: Hierarchy of Subject Arguments

Canonical complements The group of canonical complements comprises non-extracted sentential and nominal arguments whose canonical position is after the verb. They cover the following contexts:

(3.26) a. They wrapped **a batch of bandages** around Yossarian.

(Canonical nominal object)

b. The bombardiers did not give **a promise** to the colonel.

(Object of a light verb)

c. The doctor showed **Yossarian two fingers**.

(Canonical indirect object & direct obj)

d. The doctor showed **two fingers to Yossarian**.

(Canonical nominal direct object & PP-object)

e. Yossarian suggested **that they send him flowers.**

(Canonical sentential complement)

The NP node of the canonical complement is represented as a substitution node without any constraints on the complexity of the NP. (3.26b) exemplifies the active alternation of a family anchored by a light verb. In comparison to (3.24c), the noun cooccuring with the light verb is assigned an object function, and therefore inherits the features specific to arguments following the verb (e.g., verbal case assigner). As (3.26c) and (3.26d) indicate, the position of the direct object depends on the surrounding context – when used with an indirect object, it is placed after it and when used with a *PPobject*, it is given priority in the linear order. In our description language this was implemented by enforcing unique ranking on the nodes of interest. More specifically, the indirect object is given the highest rank, followed by the direct object and the the PP complements are declared of lower rank.

Verbs selecting a sentential complement impose constraints on their $\langle mode \rangle$ and $\langle comp \rangle$ values, which are declared in the lexical entry of the verb. For instance, the verb *hope* selects both an infinitive and an indicative clause types. In the first case it does not license any complementiser (hence, $\langle comp = null \rangle$), whereas the indicative mode must cooccur with an obligatory *that* complementiser. The selectional restrictions on the $\langle mode \rangle$ and $\langle comp \rangle$ are a peculiarity specific to verbs, and therefore these feature values are specified in the lexicon. In the declarative

tree, the S complement is marked as a foot node, to allow for the derivation of sentences with wh-extraction. (see the graphical representation of the derivation process in Figures 3.2 and 3.1)

Extracted arguments   The extraction phenomenon affects nominal complements and is characterised by the complements being "moved" at an unbounded distance to the left-hand side of the verb. In the tradition of generative grammars, extracted arguments are viewed as placed in a *Spec* position of a *CP*, i.e., at the beginning of a clause. The moved element can be surface-realised as a wh-word (3.27a), topicalised-NP (3.27b) or a relative clause (3.27c). Our analysis considers $\langle wh = + \rangle$ extraction and relative clauses, leaving aside topicalisation cases.

(3.27) a. Who declared war on all modifiers?
    b. The bombardier who declared war on all modifiers.
    c. War on all modifiers, the bombardier declared.

Wh-extraction   The extraction of an argument to the left-hand side of the verb is represented by adding above the root S_r node two additional nodes, marked by $\langle extracted = + \rangle$ at the topmost node:
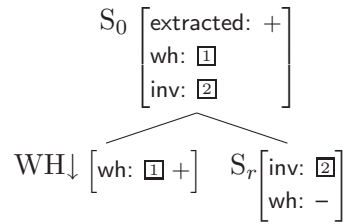
Figure 3.11: *Wh*-Extaction

The WH node is coindexed by a trace feature with the empty position, and therefore it can be substituted by the same phrase as the extracted one, including wh-NP (*who, what*), wh-phrase headed by a preposition (*by whom*), adjectival question (*how*) and a determiner phrase (*which island*).

Relative clauses   Relative clauses involve the extraction of an argument or an adjunct forming a complex nominal phrase which appears as an auxiliary tree within the relevant tree families. They are implemented by adding four additional nodes having an NP as a root node
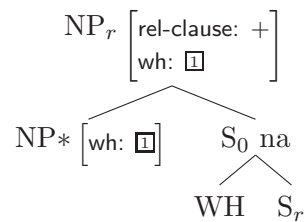
Figure 3.12: Relative Clause

The implementation of relative clauses differentiates between overt and covert relative clauses, i.e. clauses with an overt wh-NP (and without a complementiser, (3.28a)) and clauses with a covert complementiser (and without a surface realised wh-NP, (3.28b)).

(3.28) a. the girl [which$_i$ [Nately had fallen so deeply in love with ϵ_i]

b. the girl [[$whNP\ \epsilon$]$_j$ that Nately had fallen so deeply in love with ϵ_j]

Relative clauses with an overt extracted wh-argument involve the substitution of wh-NP by a lexical item with a positively marked wh-feature $\langle wh = + \rangle$. The adjunction of a complementiser is blocked by the $\langle comp = nil \rangle$ value on the S$_r$ node. The WH node in covert extracted relative clauses is anchored by an empty category ϵ, disallowing substitution of a wh-NP. The insertion of an auxiliary tree anchored by *that*-COMP is permitted by the unification of $\langle assign - comp = that \rangle$ on the foot node of the auxiliary tree and the $\langle assign - comp = that/ind\_nil/inf\_nil \rangle$ on the S$_r$ node of the relative tree.

Gerund arguments The implementation of subject and object arguments coocurring with a gerund clause takes into consideration the following distributions:

(3.29) a. Milo believes that [$_{NP}$ the making of arrangements with a four-year-old pimp] is useful.

b. Yossarian does not approve of [$_{NP}$ Milo making arrangements with a four-year-old pimp].

The examples above mark the different character of gerund clauses: (3.29a) exemplifies a determiner gerund of a transitive verb with a direct object, instantiated as an object of a prepositional phrase. The way determiner gerunds appear with determiners and can be modified by adjectives make them behave as basic nominals. Apart from *of*-phrases, the implemented determiner gerund allow for indirect object instantiated as a *for*-object (e.g., *the baking for Lisa of a cake*). The example in (3.29b) illustrates NP gerunds which are claimed to have an internal sentential structure. Therefore, an NP gerund tree has the same active/passive alternations and functional distributions as the declarative case in the corresponding family, with the only difference being the NP_r category.

Cleft arguments Cleft arguments are characterised by the following main properties:

1. coanchoring by a dummy subject and the main verb *be*; and

2. a cleft element of an *NP, Adv* or *PP* category

The current implementation covers the declarative cases in (3.30a), as well as wh-movement of the cleft component without inversion and with obligatory adjunction of an auxiliary verb (3.30b):

(3.30) a. It was colonel Korn who has mapped out this way of life for the chaplain.

b. Who might it be who has mapped out this way of life for the chaplain.

Summing up In this chapter we introduced the methodology adopted for developing a core tree-adjoining grammar for English verbs. The methodological principles were validated in the context of a core grammar implementation. The overall organisation was laid out over four levels of abstraction, each of them capturing generalised descriptions. The generalisations we studied concern alternative surface realisations at phrase-structure level and partial structures shared among several elementary trees. We saw that the grammar information can be factorised only by means of inheritance hierarchy and a few logical operations: conjunction and disjunction.

This justifies the use of a monotonic and minimal meta-language that yields a compact tree representation and a complete grammar.

Chapter 4

Evaluation and Results

The aim of the evaluation is to test the developed grammar and, more specifically, to verify whether it returns a correct analysis for each clause of a specially designed test suite. The evaluation method proceeded in three steps:

1. **Building a test corpus**. The clauses were built so that they cover the possible syntactic configurations for each subcategorisation frame. Every phenomenon was exemplified by at least one grammatical clause. In order to check the propagation of feature information, we included 11 ungrammatical clauses. In this way, we tested whether the grammar could (incorrectly) return a positive analysis for ungrammatical sentences. It should be taken into consideration that the test corpus was built so that it illustrates every grammatically possible syntactic configuration without focusing on the ungrammatical verification. At this stage, the test suite contains 1025 clauses, covering 990 syntactic variations among 57 tree families.

2. **Building a lexicon**. The lexical information is spread in two lexicons: one containing the main forms of the lexical items (i.e., the lemmas) from the test suite, and another with their morphological variants. The former gives information about the tree family a lexical item belongs to. In the case of verbs which cooccur with a specific particle, preposition or a noun complement, this lexical item is specified within the verbal entry as its coanchor. Currently, this lexicon comprises 276 entries. The other lexicon gives the inflected forms and the feature structure of each lexical entry: the lemma of a verb is linked to its present indicative, past indicative, base form, gerund and past participle. The lemma of a noun corresponds to its singular and plural variants; determiners, adjectives, prepositions and complementisers do not undergo any inflection and therefore, they are specified only in terms of features. The grammar was then further extended with 28 tree descriptions to cover non-verbal trees and auxiliary verbs. This number includes nouns, noun modifiers, adjectives, adverbs (modifying a verb on the left and a verb on the right), prepositions, prepositional phrases, determiners, complementisers, conjunctions and auxiliary verbs. The implementation of auxiliary verbs accounts for verbs adjoining to a V node (when there is no inversion) and for verbs adjoining to an S node (when there is subject - verb inversion).

3. **Parsing the test suite of sentences**. At this step all the previously specified grammar

and lexical information come into play. The evaluation was performed using the TuLiPA parsing architecture (see [Kallmeyer et al., 2008]) which takes as an input a three-level lexicon (the grammar specification file, the list of lexical entries and the lexicon with inflected forms) and a string from the test suite corpus. If the syntactic configuration is specified properly in the grammar and if all the lexical items appear in the lexicon, the parser outputs the solution in the form of a tree anchored by the lexical items from the parsed clause. Currently, the parser returns the intended parse trees for 990 grammatical clauses.

Covered phenomena   To illustrate the covered phenomena we grouped the 990 syntactic configurations among the 57 verb families into 7 general classes of syntactic configurations(with active/passive distinction if available). These classes are graphically represented in Figures (4.1) and in (4.2). These figures show that the canonical argument realisation in active and passive altogether comprises 154 test sentences. The imperative bar shows that imperative forms occur in almost 43 cases. The same number holds for clauses with a PRO-subject in active and slightly less for determiner gerunds. The NP gerund phenomenon is distributed among 102 of the test clauses, which include gerunds with canonical and PRO-subject in active and passive (without overlapping between gerund clauses and indicative clauses). Our grammar covers mostly relativisation phenomena (443 clauses). Finally, the wh-extraction of a subject, direct and indirect object is present in 186 sentences.
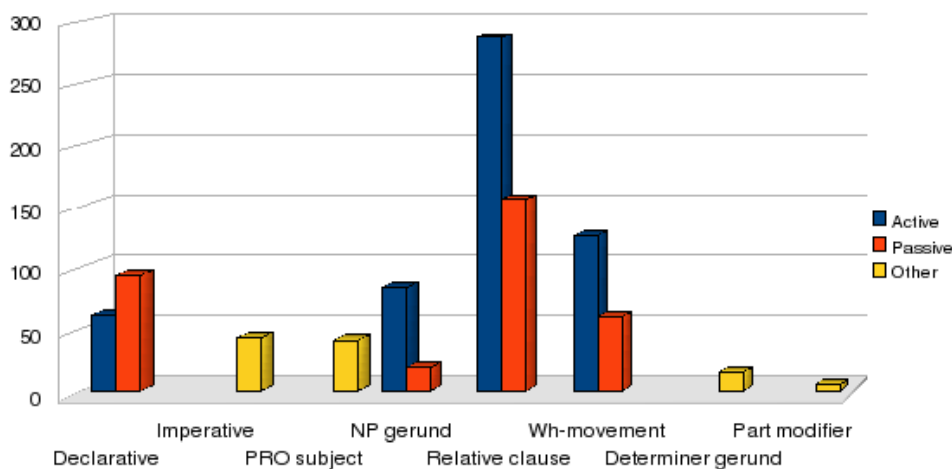
Figure 4.1: Graph of Covered Phenomena

Observed errors   The most common error consists in assigning more than one parse structure to a sentence. For instance, the NP-complement in *Yossarian approves of him taking out the chestnuts* is interpreted on one hand as a gerund rooted in a noun phrase, and on the other as a reduced relative clause. In that case both structures are permissible but there is only one which is preferred. A possible solution to this explosion of readings would be to assign cost values to tree descriptions in the process of grammar development, and adapt the parser accordingly so that in the end only the least costly (or highest ranked structure) will be licensed.

Another overgeneration error is closely related to extreme factorisation. It was observed that one tree fragment may require different feature structures depending on the surrounding context.

Declarative active	61	6.16%
Declarative passive	93	9.39%
Determiner gerund	15	1.52%
Imperative	43	4.34%
Inverted question	1	0.10%
NP Gerund active	83	8.38%
NP Gerund passive	19	1.92%
PRO Subject	41	4.14%
Participial modifier	5	0.51%
Relative active	287	28.99%
Relative passive	156	15.76%
Wh-moved active	126	12.73%
Wh-moved passive	60	6.06%
Total	990	100%

Figure 4.2: Distribution of Covered Phenomena

For instance, the tree description defining a canonical subject should be accompanied by an indicative mode when used in a declarative context and by $\langle mode = base \rangle$ when used with wh-extraction. Solving this problem comes at the cost of 'loading' the grammar with more tree descriptions.

Comparison to XTAG   This report began by saying that the current grammar takes as a starting point the linguistic decision of the XTAG project. Let is now briefly discuss to what extent we managed to cover the verb families and their syntactic transformations as defined by [XTAG-Research-Group, 2001]. XTAG distributes the verbs into 57 subcategorisation frames generating around 1000 trees. Our grammar covered all verb families which amount to 57 and was able to generate successfully 1017 trees. Future work might involve automation of the comparison between the coverage of the XTAG implementation and the currently developed grammar.

Conclusion and Perspectives

In this report we discussed a possible implementation of a core tree-adjoining grammar for English. We proposed a novel framework for representing strongly lexicalised tree-based formalisms which takes advantage of a high abstraction level and minimal descriptions. The proposed language was built with the intuition that the two methodological axes which suffice for designing a wide-coverage grammar are shared structures and alternatives. These can be easily handled by means of conjunction and disjunction operations. Further on, we showed that the number of tree descriptions can be reduced by using an inheritance hierarchy which lets the grammar writer invoke on-the-fly the fragments of interest.

The proposed language and methodology were validated by building a realistic English grammar based on the linguistic decisions of [XTAG-Research-Group, 2001]. The implementation process consisted in abstracting in four description levels: we started by defining a tree family as a disjunction of alternatives of each other, then we moved to the next step of declaring what the tree descriptions in each of these alternation is, then we described each of these elements as equally ranked alternatives and, finally, we designed the building blocks of the grammar, i.e., elementary trees. The main practical gains are, on one hand the monotonic language, which does not impose any order for deriving a grammatical structure from another one, but instead regards possible surface-realisations as equally ranked alternatives, and, on other, the minimal language, which allows for high level of generalisation only by means of conjunction, disjunction and inheritance. The very systematic way of implementation facilitates thus the maintenance and the possible extension of the grammar. The grammar writer manipulates a very ordered list of classes and the relations between them without having to interfere with the overall logic behind the grammar implementation.

Further investigations may proceed in several directions. The grammar and the lexicon can be further extended with the purpose of scaling up to a large-coverage grammar. As it was observed, the grammar can be easily extended by adding more lexical classes and by reusing the existing elementary tree descriptions. Another possible perspective would be to add underspecified semantic representation for the purposes of natural language generation. This would add a great value to the available resources considering the fact that at this stage there is no large-scale implemented TAG integrating compositional semantics.

References

[Abeillé1991] Abeillé, Anne. 1991. *Une grammaire lexicalisée d'arbres adjoints pour le français.* Université Paris 7. PhD thesis.

[Aït-Kaci1991] Aït-Kaci, Hassan. 1991. *Warren's abstract machine: a tutorial reconstruction.* MIT Press, Cambridge, MA, USA.

[Bresnan and Kaplan1982] Bresnan, Joan and Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In J. Bresnan, editor, *The Mental Representation of Grammatical Relations.* MIT Press, Cambridge, MA, pages xvii–lii.

[Candito1999] Candito, Marie-Hélène. 1999. *Représentation modulaire et paramétrable de grammaires électroniques lexicaliseés : application au français et à l'italien.* Université Paris 7. PhD thesis.

[Chomsky1957] Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

[Chomsky1981] Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures.* Foris Publications.

[Crabbé and Duchier2004] Crabbé, B. and D. Duchier. 2004. Metagrammar redux. In *International Workshop on Constraint Solving and Language Processing - CSLP 2004, Copenhagen.*

[Crabbé2005] Crabbé, Benoit. 2005. *Représentation informatique de grammaires fortement lexicalisées: Application à la grammaire d'arbres adjoints.* Ph.D. thesis. Ph.D. thesis.

[Duchier, Le Roux, and Parmentier2004] Duchier, D., J. Le Roux, and Y. Parmentier. 2004. The Metagrammar Compiler: An NLP Application with a Multi-paradigm Architecture. In *Second International Mozart/Oz Conference (MOZ'2004)*, Charleroi.

[Gardent and Kallmeyer2003] Gardent, C. and L. Kallmeyer. 2003. Semantic construction in ftag. In *Proceedings of the 10th EACL*, Budapest, Hungary.

[Iordanskaja and Mel'cukto appear] Iordanskaja and Mel'cuk. to appear. Towards establishing an inventory of surface-syntactic relations: Valency-controlled surface-syntactic dependents of the verb in french.

[Joshi and Schabes1997] Joshi, Aravind K. and Yves Schabes. 1997. Tree-adjoining grammars. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3. Springer, Berlin, New York, pages 69–124.

[Kallmeyer et al.2008] Kallmeyer, Laura, Timm Lichte, Wolfgang Maier, Yannick Parmentier, and Johannes Dellert. 2008. Developing an mctag for german with an rcg-based parser. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC-2008)*, page 1240. to appear.

[Levin1993] Levin, Beth. 1993. *English Verb Classes and Alternations: a preliminary investigation.* University of Chicago Press, Chicago and London.

[Perrier2007] Perrier, Guy. 2007. A french interaction grammar. In *International Conference on Recent Advances in Natural Language Processing – (RANLP-2007)*, pages 463–467.

[Schabes, Abeille, and Joshi1988] Schabes, Yves, Anne Abeille, and Aravind K. Joshi. 1988. Parsing strategies with 'lexicalized' grammars: application to tree adjoining grammars. In *Proceedings of the 12th conference on Computational linguistics*, pages 578–583, Morristown, NJ, USA. Association for Computational Linguistics.

[Vijay-Shanker1991] Vijay-Shanker, Joshi A. K. 1991. Unification based tree adjoining grammars. In G. Rozenberg and A. Salomaa, editors, *Unification-based Grammars*. MIT Press, Cambridge, Massachusetts.

[Vijay-Shanker1987] Vijay-Shanker, K. 1987. A study of tree adjoining grammar. University of Pennsylvania, Philadelphia. PhD thesis.

[Xia2001] Xia, Fei. 2001. *Automatic grammar generation from two different perspectives*. Ph.D. thesis. Supervisor-Martha Palmer and Supervisor-Aravind Joshi.

[XTAG-Research-Group2001] XTAG-Research-Group. 2001. A lexicalized tree adjoining grammar for english. Technical Report IRCS-01-03, IRCS, University of Pennsylvania.

[Y. Parmentier and J. Le Roux2005] Y. Parmentier and J. Le Roux. 2005. XMG: an Extensible Metagrammatical Framework. In *Student session of the 17th European Summer School in Logic, Language and Information,*, Aug.

Appendix A

Available syntactic configurations

A.1 n0V

1. Declarative
   Bill sleeps; a man sleeps; the man sleeps

2. Wh-moved subject
   who sleeps

3. Subject relative clause with overt extracted wh-NP
   the man who sleeps, the man whose wife sleeps
   * the man who sleep, the man who to sleep
   * the man who sleeping, the man whose wife sleeping

4. Subject relative clause with covert extracted wh-NP
   the man that sleeps,the man to sleep
   * the man who that sleeps
   * the man whose wife to sleep

5. Imperative
   Sleep!

6. Determiner gerund
   Bill disapproves of [her dancing]

7. NP gerund
   Bill disapproved of [me dancing]

8. PRO subject
   Bill wants [PRO to sleep]

9. NP gerund with PRO subject
   [PRO dancing] fascinates Lisa.

10. Adjunct (gap-less) relative clause with PP pied-piping
    the house in which Bill sleeps

11. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day Bill sleeps

12. Pre-nominal participial modifier
   the sleeping man

A.2   En1v

1. Declarative
   the snow melts

2. Wh-moved subject
   what melted

3. Imperative
   melt

4. Subject relative clause with covert extracted wh-NP
   the snow that melts
   the snow to melt

5. Subject relative clause with overt extracted wh-NP
   the snow which melts

6. Determiner gerund
   Bill disapproves of [the melting]

7. NP gerund
   [it melting]

8. PRO subject
   the snow needs [PRO to melt]

9. NP gerund with PRO subject
   [PRO melting] is necessary

10. Adjunct (gap-less) relative clause with PP pied-piping
   the island on which the snow melts

11. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day the snow melts

12. Pre-nominal participial modifier
   Lisa likes [the melting snow]

A.3   n0Vn1

1. Declarative
   Bill baked a cake
   Bill loves Lisa

2. Wh-moved object
    what did Bill bake
    who does Bill like

3. Wh-moved subject
    who baked a cake

4. Imperative
    bake the cake

5. Passive with by-phrase
    a cake is baked by Bill
    * a cake is bake by Bill

6. Passive without by-phrase
    a cake was baked

7. Passive with wh-moved subject and by-phrase
    what was baked by Bill

8. Passive with wh-moved subject and no by-phrase
    what was baked

9. Passive with wh-moved by-phrase
    by whom was the cake baked

10. Passive wit wh-moved object out of the by-phrase
    whom was the cake baked by
    who was the cake baked by

11. Object relative clause with overt extracted wh-NP
    the cake whose cherries Bill likes
    the cake which Bill baked

12. Object relative clause with covert extracted wh-NP
    the cake that Bill bakes

13. Subject relative clause with overt extracted wh-NP
    the man who ate the cake

14. Subject relative clause with covert extracted wh-NP
    the man that ate the cake

15. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day Bill ate the cake

16. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
    the day the cake was baked by Bill

17. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive without by-phrase
    the day the cake was baked

18. Adjunct (gap-less) relative clause with PP pied-piping

the house in which Bill baked a cake

19. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the house in which the tree was planted by Bill

20. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase
    the house in which the tree was planted

21. Passive with relative clause on subject and no by-phrase with overt extracted wh-NP
    the cake whose cherries were eaten

22. Passive with relative clause on subject and no by-phrase with covert extracted wh-NP
    the cake that was eaten

23. Passive with relative clause on subject and by-phrase with overt extracted wh-NP
    the cake whose cherries were eaten by Bill

24. Passive with relative clause on subject and by-phrase with covert extracted wh-NP
    the cake that was eaten by Bill

25. Passive with relative clause on object on the by-phrase with overt extracted wh-NP
    the man who the tree was planted by

26. Passive with relative clause on object on the by-phrase with covert extracted wh-NP
    the man that the tree was planted by

27. Passive with relative clause on object on the by phrase with with PP pied-piping
    the man by whom the cake was baked
    Lisa likes [the man by whom the tree was planted]
    [the man by whom the tree was planted] likes Lisa

28. Determiner gerund
    Lisa approves of [his baking of the cake]

29. NP gerund
    Bill approved of [Lisa baking a cake]

30. PRO subject
    Bill wants [PRO to bake a cake]

31. NP gerund with PRO subject
    Bill approved of [PRO baking a cake]

32. Gerund passive with by-phrase with PRO subject
    Lisa likes [PRO being loved by Bill].

33. Gerund passive without by-phrase with PRO subject
    Lisa likes [PRO being loved]

34. Passive with by-phrase with PRO subject
    Bill wants [PRO to be loved by Lisa].

35. Passive without by-phrase with PRO subject
    Bill wants [PRO to be loved].

36. Gerund passive with by-phrase
    Bill being loved by Lisa

37. Gerund passive without by-phrase
    Bill being loved

38. Pre-nominal participial modifier
    the baked cake

39. Derived adjective
    totally baked

A.4   n0Vn2n1

1. Declarative
   Bill asks Lisa a question

2. Wh-moved subject
   who asked Lisa a question

3. Wh-moved direct object
   what did Bill ask Lisa

4. Wh-moved indirect object
   whom did Bill ask a question

5. Subject relative clause with overt extracted wh-NP
   the man who asked Lisa a question

6. Subject relative clause with covert extracted wh-NP
   the man that asked Lisa a question
   the man to ask Lisa a question

7. Direct object relative clause with overt extracted wh-NP
   the question which Bill asks Lisa

8. Indirect object relative clause with overt extracted wh-NP
   the woman whom Bill asks a question

9. Direct object relative clause with covert extracted wh-NP
   the question that Bill asked Lisa

10. Indirect object relative clause with covert extracted wh-NP
    the woman that Bill asks a question

11. Imperative
    ask Lisa a question

12. Determiner gerund
    the asking for Lisa of a question

13. NP gerund

Bill approves of [him asking Lisa a question]

14. Passive with by-phrase
    Lisa was asked a question by Bill

15. Passive without by-phrase
    Lisa was asked a question

16. Passive with wh-moved subject and indirect object and by-phrase
    what was Lisa asked by Bill

17. Passive with wh-moved subject and no by-phrase
    what was Lisa asked

18. Passive with wh-moved subject and direct object and by-phrase
    who was asked a question by Bill

19. Passive with wh-moved subject and direct object without by-phrase
    who was asked a question

20. Passive with wh-moved by-phrase
    by whom was Lisa asked a question

21. Passive with wh-moved object out of the by-phrase
    who was Lisa asked a question by

22. Passive with relative clause on subject (on NP1) and by-phrase with overt extracted wh-NP
    the question which Lisa was asked by Bill

23. Passive with relative clause on subject (on NP1) and by-phrase with covert extracted wh-NP
    the question that Lisa was asked by Bill

24. Passive with relative clause on subject (on NP1) and no by-phrase with overt extracted wh-NP
    the question which Lisa was asked

25. Passive with relative clause on subject (on NP1) and no by-phrase with covert extracted wh-NP
    the question that Lisa was asked

26. Passive with relative clause on subject (on NP2) and by-phrase with overt extracted wh-NP
    the man whose wife was asked a question by Bill

27. Passive with relative clause on subject (on NP2) and by-phrase with covert extracted wh-NP
    the woman that was asked a question by Bill

28. Passive with relative clause on subject (on NP2) without by-phrase with overt extracted wh-NP
    the man whose wife was asked a question

29. Passive with relative clause on subject (on NP2) without by-phrase with covert extracted wh-NP

> *the woman that was asked a question*

30. Passive with relative clause on object of by-phrase with overt extracted wh-NP
    > *the man who Lisa was asked a question by*

31. Passive with relative clause on object of by-phrase with covert extracted wh-NP
    > *the man that Lisa was asked a question by*

32. Passive with relative clause on by-phrase with PP pied-piping
    > *the man by whom Lisa was asked a question*

33. Adjunct (gap-less) relative clause with covert extracted wh-NP
    > *the day Bill asked Lisa a question*

34. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
    > *the day Lisa was asked a question by Bill*

35. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive without by-phrase
    > *the day Lisa was asked a question*

36. Adjunct (gap-less) relative clause with PP pied-piping
    > *the house in which Bill asked Lisa a question*
    > *the house where Bill asked Lisa a question*

37. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    > *the house in which Lisa was asked a question by Bill*

38. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase
    > *the house in which Lisa was asked a question*

39. PRO subject
    > *Bill wants [PRO to ask Lisa a question]*

40. Passive without by-phrase with PRO subject
    > *Lisa wants [PRO to be asked a question]*

41. Passive with by-phrase with PRO subject
    > *Lisa wants [PRO to be asked a question by Bill]*

42. NP gerund passive without by-phrase and PRO subject
    > *Lisa likes [PRO being asked a question]*

43. NP gerund passive with by-phrase and PRO subject
    > *Lisa likes [PRO being asked a question by Bill]*

44. Gerund passive with by-phrase
    > *[him being asked a question by Bill]*

45. Gerund passive without by-phrase
    > *[him being asked a question]*

46. NP gerund with PRO subject
    > *Bill approves of [PRO asking Lisa a question]*

A.5   n0Vn1pn2

1. Declarative

    Bill puts the cake on a table

2. Wh-moved PP

    on what does Bill put the cake
    where does Bill put the cake

3. Wh-moved object of PP

    what does Bill put the cake on

4. Wh-moved direct object

    what does Bill put on the table

5. Wh-moved subject

    who puts the cake on the table

6. Imperative

    put the cake on the table

7. PRO subject

    Bill wants [to put a cake on the table]

8. Passive without by-phrase

    the cake was put on the table

9. Passive with by-phrase

    a cake was put on the table by Bill

10. Passive with wh-moved PP and no by phrase:

    where was the cake put

11. Passive with wh-moved by phrase

    by whom was the cake put on the table

12. Passive with wh-moved object out of the by-phrase

    who was the cake put on the table by

13. Passive with wh-moved object out of the PP and no by phrase

    what was the cake put on

14. Passive with wh-moved object out of the PP and by phrase

    what was the cake put on by Bill

15. Passive with wh-moved subject and no by phrase

    what was put on the table

16. Passive with wh-moved subject and by phrase

    what was put on the table by Bill

17. Passive with wh-moved PP and by phrase

    where was the cake put by Bill

18. Passive with wh-moved PP and no by phrase
    where was the cake put

19. Passive without by-phrase with PRO subject
    [PRO to be put on the table]

20. Passive with by-phrase with PRO subject
    [PRO to be put on the table by Bill]

21. Determiner gerund
    [the putting of the cake on the table]

22. NP gerund
    Lisa likes [Bill putting the cake on the table]
    [it being put on the table]

23. Gerund passive with by-phrase
    [it being put on the table by Bill].

24. NP gerund with PRO subject
    [Bill approves of [PRO putting the cake on the table]

25. NP gerund passive without by-phrase and with PRO subject
    [PRO being put on the table]

26. NP gerund passive with by-phrase and PRO subject
    the cake likes [PRO being put on the table by Bill]

27. Direct object relative clause with overt extracted wh-NP
    the cake which Bill puts on the table

28. Direct object relative clause with covert extracted wh-NP
    a cake that Bill puts on the table

29. Subject relative clause with overt extracted wh-NP
    the man who put the cake on the table

30. Object of PP relative clause with overt extracted wh-NP
    the table which Bill put the cake on

31. Object of PP relative clause with covert extracted wh-NP
    the table that Bill put the cake on

32. PP relative clause with PP pied-piping
    the table on which Bill puts the cake

33. Subject relative clause with covert extracted wh-NP
    the man that puts the cake on the table
    the man to put the cake on the table

34. Passive with relative clause on subject and no by phrase with overt extracted wh-NP
    the cake which was put on the table

35. Passive with relative clause on subject and no by phrase with covert extracted wh-NP

the cake that was put on the table; the cake to be put on the table

36. Passive with relative clause on subject and by phrase with covert extracted wh-NP
    the cake that was put on the table by Bill

37. Passive with relative clause on subject and by phrase with overt extracted wh-NP
    the cake which was put on the table by Bill

38. Passive with relative clause on object of PP and by phrase with overt extracted wh-NP
    the table which the cake was put on by Bill

39. Passive with relative clause on object of PP and by phrase with covert extracted wh-NP
    the table that the cake was put on by Bill

40. Passive with relative clause on object of PP and no by phrase with covert extracted wh-NP
    the table that the cake was put on

41. Passive with relative clause on PP indirect object and by phrase with PP pied-piping
    the table on which the cake was put by Bill

42. Passive with relative clause on PP indirect object and no by phrase with PP pied-piping
    the table on which the cake was put

43. Passive with relative clause on object of PP and no by phrase with overt extracted wh-NP
    the table which the cake was put on

44. Passive with relative clause on object of by-phrase with overt extracted wh-NP
    the man who the cake was put on the table by

45. Passive with relative clause on object of by-phrase with covert extracted wh-NP
    the man that the cake was put on the table by

46. Passive with relative clause on by-phrase with PP pied-piping
    the man by whom the cake was put on the table

47. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day Bill put the cake on the table

48. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
    the day the cake was put on the table by Bill

49. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive without by-phrase
    the day the cake was put on the table

50. Adjunct (gap-less) relative clause with PP pied-piping
    the house where Bill puts the cake on the table

51. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the house where the cake was put on the table by Bill

52. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase
    the house where the cake was put on the table

A.6   n0Vn1Pn2

1. Declarative

    the fish reminds Yossarian of other fish

2. Wh-moved subject

    what reminds Yossarian of other fish

3. Wh-moved direct object

    whom does the fish remind of other fish

4. Wh-moved object of PP

    what does the fish remind Yossarian of

5. Wh-moved PP

    of what does the fish remind Yossarian

6. Imperative

    remind Yossarian of other fish

7. PRO subject

    the colonel wants [to remind Yossarian of other fish]

8. Passive without by-phrase

    Yossarian was reminded of other fish

9. Passive with by-phrase

    Yossarian was reminded of other fish by the fish

10. Passive with wh-moved by-phrase

    by what was Yossarian reminded of other fish

11. Passive with wh-moved object out of the PP and no by phrase

    what was Yossarian reminded of

12. Passive with wh-moved object out of the PP and by phrase

    what was Yossarian reminded of by the fish

13. Passive with wh-moved subject and no by phrase

    who was reminded of other fish

14. Passive with wh-moved subject and by phrase

    who was reminded of other fish by the fish

15. Passive with wh-moved PP and by phrase

    of what was Yossarian reminded by the fish

16. Passive with wh-moved PP and no by phrase

    of what was Yossarian reminded

17. Passive without by-phrase with PRO subject

    Yossarian wants [PRO to be reminded of other fish]

18. Passive with by-phrase with PRO subject

      Yossarian wants [PRO to be reminded of other fish by the fish]

19. Passive with wh-moved object out of the by-phrase
      what was Yossarian reminded of other fish by

20. Determiner gerund
      [the reminding of Yossarian of other fish]

21. NP gerund
      Orr approves of [me reminding Yossarian of other fish]

22. Gerund passive without by-phrase
      [him being reminded of other fish]

23. Gerund passive with by-phrase
      [him being reminded of other fish by the fish]

24. NP gerund with PRO subject
      Orr approves of [reminding Yossarian of other fish]

25. NP gerund passive without by-phrase and with PRO subject
      Yossarian approves of [PRO being reminded of other fish]

26. NP gerund passive with by-phrase and PRO subject
      Yossarian approves of [PRO being reminded of other fish by the fish]

27. Direct object relative clause with overt extracted wh-NP
      the bombardier whom the fish reminds of other fish

28. Direct object relative clause with covert extracted wh-NP
      the bombardier that the fish reminds of other fish

29. Subject relative clause with overt extracted wh-NP
      the fish which reminds the bombardier of other fish

30. Subject relative clause with covert extracted wh-NP
      the fish that reminds the bombardier of other fish
      the fish to remind the bombardier of other fish

31. Object of PP relative clause with overt extracted wh-NP
      other fish which the fish reminds Yossarian of

32. Object of PP relative clause with covert extracted wh-NP
      other fish that the fish reminds Yossarian of

33. Passive with relative clause on subject and no by phrase with overt extracted wh-NP
      the bombardier who was reminded of other fish

34. Passive with relative clause on subject and no by phrase with covert extracted wh-NP
      the bombardier that was reminded of other fish
      the bombardier to be reminded of other fish

35. Passive with relative clause on subject and by phrase with covert extracted wh-NP
      the bombardier that was reminded of other fish by the fish

36. Passive with relative clause on subject and by phrase with overt extracted wh-NP

    the bombardier who was reminded of other fish by the fish

37. Passive with relative clause on object of PP and by phrase with overt extracted wh-NP

    other fish which Yossarian was reminded of by the fish

38. Passive with relative clause on object of PP and by phrase with covert extracted wh-NP

    other fish that Yossarian was reminded of by the fish

39. Passive with relative clause on object of PP and no by phrase with overt extracted wh-NP

    other fish which Yossarian was reminded of

40. Passive with relative clause on object of PP and no by phrase with covert extracted wh-NP

    other fish that Yossarian was reminded of

41. Passive with relative clause on object of by-phrase with overt extracted wh-NP

    the fish which Yossarian was reminded of other fish by

42. Passive with relative clause on object of by-phrase with covert extracted wh-NP

    the fish that Yossarian was reminded of other fish by

43. Passive with relative clause on by-phrase with PP pied-piping

    the fish by which Yossarian was reminded of other fish

44. Adjunct (gap-less) relative clause with PP pied-piping

    the island on which the fish reminds Yossarian of other fish

    the island where the fish reminds Yossarian of other fish

45. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day the fish reminds Yossarian of other fish

46. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase

    the day Yossarian was reminded of other fish by the fish

47. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive without by-phrase

    the day Yossarian was reminded of other fish

48. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase

    the island where Yossarian was reminded of other fish by the fish

49. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase

    the island where Yossarian was reminded of other fish

50. Passive with relative clause on PP indirect object and by phrase with PP pied-piping

    other fish of which Yossarian was reminded by the fish

51. Passive with relative clause on PP indirect object and no by phrase with PP pied-piping

    other fish of which Yossarian was reminded

52. PP relative clause with PP pied-piping

    other fish of which the fish reminds Yossarian

A.7   n0Vn1s2

1. Declarative

   the colonel forces Yossarian to fly many missions

2. Wh-moved subject

   who forced Yossarian to fly many missions

3. Wh-moved sentential complement

   what does the colonel force Yossarian

4. Imperative

   force the bombardiers to fly many missions

5. Passive with PRO subject with by-phrase after the sentential complement

   Yossarian wanted [PRO to be forced to fly many missions by the colonel]

6. Passive without by phrase

   the bombardiers were forced to fly many missions

7. Passive with wh-moved subject and no by phrase

   who was forced to fly many missions

8. Passive with PRO subject without by-phrase

   [PRO to be forced to fly many missions].

9. Passive with by-phrase after the sentential complement

   the bombardiers were forced to fly many missions by the colonel

10. Passive with wh-moved subject and by-phrase after the sentential complement

    who was forced to fly many missions by the colonel

11. Determiner gerund

    [his forcing of the bombardiers to fly many missions] was unnecessary

12. NP Gerund

    [me betting everyone that the bombardiers flew many missions]

13. NP gerund passive with by-phrase after the sentential complement

    Yossarian being forced to fly many missions by the colonel

14. NP gerund passive without by-phrase

    Yossarian being forced to fly many missions

15. NP gerund with PRO subject

    [PRO forcing the bombardiers to fly many missions] was extreme.

16. NP gerund passive with PRO subject without by-phrase

    [PRO being forced to fly many missions]

17. NP gerund passive with PRO subject with by-phrase after the sentential complement

    [PRO being forced to fly many missions by the colonel]

18. Subject relative clause with overt extracted wh-NP

the colonel who forces Yossarian to fly many missions

19. Subject relative clause with covert extracted wh-NP
    the colonel that forces Yossarian to fly many missions

20. Passive with relative clause on subject and no by-phrase with overt extracted wh-NP
    the bombardier who was forced to fly many missions

21. Passive with relative clause on subject and no by-phrase with covert extracted wh-NP
    the bombardier that was forced to fly many missions

22. Passive with relative clause on subject and by-phrase after the sent complement with overt extracted wh-NP
    the bombardiers who were forced to fly many missions by the colonel

23. Passive with relative clause on subject and by-phrase after the sent complement with covert extracted wh-NP
    the bombardiers that were forced to fly many missions by the colonel

A.8   n0Vpl

1. Declarative
   Yossarian woke up

2. Wh-moved subject
   who woke up

3. Imperative
   wake up

4. Subject relative clause with overt extracted wh-NP
   the bombardier who woke up

5. Subject relative clause with covert extracted wh-NP
   the bombardier that woke up

6. Adjunct (gap-less) relative clause with PP pied-piping
   the island where the bombardier woke up
   the island on which the bombardier woke up

7. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day Yossarian woke up

8. Determiner gerund
   the colonel disapproves of [the waking up]

9. NP gerund
   the colonel approves of [him waking up]

10. PRO subject
    Yossarian wants [PRO to wake up]

11. NP gerund with PRO subject

    Yossarian dreads [PRO waking up]

A.9   n0Vpln1

1. Declarative with a particle after the NP

        Orr took the chestnuts out

2. Declarative with a particle before the NP

        Orr took out the chestnuts

3. Wh-moved object

        what did Orr take out

4. Wh-moved subject with particle before the NP

        who took out the chestnuts

5. Wh-moved subject with particle after the NP

        who took the chestnuts out

6. Imperative with particle after the NP

        take the chestnuts out

7. Imperative with particle before the NP

        take out the chestnuts

8. Passive with by-phrase with verb particle before by-phrase

        the chestnuts were taken out by Orr

9. Passive without by-phrase

        the chestnuts were taken out

10. Passive with wh-moved subject and by-phrase with verb particle before by-phrase

        what was taken out by Orr

11. Passive with wh-moved subject and no by-phrase

        what was taken out

12. Passive with wh-moved by-phrase

        by whom were the chestnuts taken out

13. Passive with wh-moved object out of the by-phrase

        who were the chestnuts taken out by

14. Object relative clause with overt extracted wh-NP

        the chestnuts which Orr took out

15. Object relative clause with covert extracted wh-NP

        the chestnuts that Orr took out

16. Subject relative clause with particle before the NP with overt extracted wh-NP

        the bombardier who took out the chestnuts

17. Subject relative clause with particle after the NP with overt extracted wh-NP
    the bombardier who took the chestnuts out

18. Subject relative clause with particle before the NP with covert extracted wh-NP
    the bombardier that took out the chestnuts

19. Subject relative clause with particle after the NP with covert extracted wh-NP
    the bombardier that took the chestnuts out

20. Adjunct (gap-less) relative clause with particle after the NP with PP pied-piping
    the island on which Orr took the chestnuts out

21. Adjunct (gap-less) relative clause with particle before the NP with PP pied-piping
    the island on which Orr took out the chestnuts

22. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the island on which the chestnuts were taken out by Orr

23. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase
    the island on which the chestnuts were taken out

24. Adjunct (gap-less) relative clause with particle after the NP with covert extracted wh-NP
    the day Orr took the chestnuts out

25. Adjunct (gap-less) relative clause with particle before the NP with covert extracted wh-NP
    the day Orr took out the chestnuts

26. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
    the day the chestnuts were taken out by Orr

27. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive without by-phrase
    the day the chestnuts were taken out

28. Passive with relative clause on subject and no by-phrase with overt extracted wh-NP
    the chestnuts which were taken out

29. Passive with relative clause on subject and no by-phrase with covert extracted wh-NP
    the chestnuts that were taken out

30. Passive with relative clause on subject and by-phrase with overt extracted wh-NP
    the chestnuts which were taken out by Orr

31. Passive with relative clause on subject and by-phrase with covert extracted wh-NP
    the chestnuts that were taken out by Orr

32. Passive with relative clause on the object of by-phrase with overt extracted wh-NP
    the bombardier who the chestnuts were taken out by

33. Passive with relative clause on the object of by-phrase with covert extracted wh-NP
    the bombardier that the chestnuts were taken out by

34. Determiner gerund with particle before the NP
    Orr approves of [the taking out of the chestnuts]

35. NP gerund with particle after the NP
    Yossarian likes [him taking the chestnuts out]

36. NP gerund with particle before the NP
    Yossarian likes [him taking out the chestnuts]

37. Gerund passive with by-phrase
    [Orr being beaten up by Yossarian]

38. Gerund passive with by-phrase
    [Orr being beaten up]

39. PRO subject with verb particle after the NP
    Orr wants [PRO to take the chestnuts out]

40. PRO subject with verb particle before the NP
    Orr wants [PRO to take out the chestnuts]

41. NP gerund with PRO subject with verb particle before the NP
    Yossarian approves of [PRO taking out the chestnuts]

42. NP gerund with PRO subject with verb particle after the NP
    Yossarian approves of [PRO taking the chestnuts out]

43. Gerund passive with by-phrase with PRO subject
    Orr likes [PRO being beaten up by Yossarian]

44. Gerund passive without by-phrase with PRO subject
    Orr likes [PRO being beaten up]

45. Passive with PRO subject with by-phrase
    Orr wants [PRO to be beaten up by Yossarian]

46. Passive with PRO subject and without by-phrase
    Orr wants [PRO to be beaten up]

A.10   n0Vpln2n1

1. Declarative with particle after the NPs
   Bill opens Lisa a bank account up

2. Declarative with particle between the NPs
   Bill opens Lisa up a bank account

3. Declarative with particle before the NPs
   Bill opens up Lisa a bank account

4. Wh-moved subject with particle after the NPs
   who opens Lisa a bank account up

5. Wh-moved subject with particle between the NPs
   who opens Lisa up a bank account

6. Wh-moved subject with particle before the NPs
   who opens up Lisa a bank account

7. Wh-moved indirect object with particle after the NPs
   whom did Bill open a bank account up

8. Wh-moved indirect object with particle before the NPs
   whom did Bill open up a bank account

9. Wh-moved direct object with particle between the NPs
   what did Bill open Lisa up

10. Wh-moved direct object with particle before the NPs
    what did Bill open up Lisa

11. Imperative with particle after the NPs
    open Lisa a bank account up

12. Imperative with particle between the NPs
    open Lisa up a bank account

13. Imperative with particle before the NPs
    open up Lisa a bank account

14. PRO subject with particles after the NPs
    Bill wants [PRO to open Lisa a bank account up]

15. PRO subject with particles between the NPs
    Bill wants [PRO to open Lisa up a bank account]

16. PRO subject with particles before the NPs
    Bill wants [PRO to open up Lisa a bank account]

17. Determiner gerund with particle before the NPs
    the opening up for Lisa of a bank account

18. NP gerund with particle before the NPs
    [me opening up Lisa a bank account]

19. NP gerund with particle between the NPs
    [me opening Lisa up a bank account]

20. NP gerund with particle after the NPs
    [me opening Lisa a bank account up]

21. NP gerund with PRO subject with particles before the NPs
    Bill approves of [PRO opening up Lisa a bank account]

22. NP gerund with PRO subject with particles between the NPs
    Bill approves of [PRO opening Lisa up a bank account]

23. NP gerund with PRO subject with particles after the NPs
    Bill approves of [PRO opening Lisa a bank account up]

24. Direct object relative clause with particle between the NPs with overt extracted wh-object

a bank account which Bill opens Lisa up

25. Direct object relative clause with particle before the NPs with overt extracted wh-object
    a bank account which Bill opens up Lisa

26. Direct object relative clause with particle between the NPs with covert extracted wh-object
    a bank account that Bill opens up Lisa

27. Direct object relative clause with particle before the NPs with covert extracted wh-object
    a bank account that Bill opens Lisa up

28. Subject relative clause with particle after the NPs with overt extracted wh-NP
    the man who opens Lisa a bank account up

29. Subject relative clause with particle between the NPs with overt extracted wh-NP
    the man who opens Lisa up a bank account

30. Subject relative clause with particle before the NPs with overt extracted wh-NPs
    the man who opened up Lisa a bank account

31. Subject relative clause with particle after the NPs with covert extracted wh-NP
    the man that opens Lisa a bank account up

32. Subject relative clause with particle between the NPs with covert extracted wh-NP
    the man that opens Lisa up a bank account

33. Subject relative clause with particle before the NPs with covert extracted wh-NPs
    the man that opens up Lisa a bank account

34. Indirect object relative clause with particle before the NPs with overt extracted wh-NP
    the woman who Bill opens up a bank account

35. Indirect object relative clause with particle before the NPs with covert extracted wh-NP
    the woman that Bill opens up a bank account

36. Indirect object relative clause with particle after the NPs with overt extracted wh-NP
    the woman who Bill opens a bank account up

37. Indirect object relative clause with particle after the NPs with covert extracted wh-NP
    the woman that Bill opens a bank account up

38. Adjunct (gap-less) relative clause with particle before the NPs with PP pied-piping
    the country in which Bill opens up Lisa a bank account

39. Adjunct (gap-less) relative clause with particle between the NPs with PP pied-piping
    the country in which Bill opens Lisa up a bank account

40. Adjunct (gap-less) relative clause with particle after the NPs with PP pied-piping
    the country in which Bill opens Lisa a bank account up

41. Adjunct (gap-less) relative clause with particle before the NPs with covert extracted wh-NP
    the day Bill opens up Lisa a bank account

42. Adjunct (gap-less) relative clause with particle between the NPs with covert extracted wh-NP

   the day Bill opens Lisa up a bank account

43. Adjunct (gap-less) relative clause with particle after the NPs with covert extracted wh-NP
   the day Bill opens Lisa a bank account up

A.11 n0Vs1

1. Canonical subject
   Yossarian wants to fly many missions

2. Wh-moved subject
   who wants to fly many missions

3. Wh-moved sentential complement
   what does Yossarian want

4. Imperative
   insist that Yossarian fly many missions

5. PRO subject
   Yossarian wanted [PRO to think that he flew many missions]

6. NP gerund
   [Yossarian wanting to fly many missions]

7. NP gerund with PRO subject
   [PRO wanting to fly many missions] fascinates the colonel

8. Determiner gerund
   Orr approves of [the insisting that Yossarian fly many missions]

9. Subject relative clause with overt extracted wh-NP
   the bombardier who wants to fly many missions

10. Subject relative clause with covert extracted wh-NP
   the bombardier that wants to fly many missions

A.12 n0Va1

1. Canonical subject
   the flowers smelled enthralling

2. Wh-moved adjective
   how did the flowers smell

3. Wh-moved subject
   what smelled enthralling

4. PRO subject
   Lisa wants [to look happy]

5. Imperative

    look happy

6. NP gerund

    Lisa approves of [it smelling enthralling]

7. NP gerund with PRO subject

    Lisa likes [smelling enthralling]

8. Subject relative clause with overt extracted wh-NP

    the flowers which smelled enthralling

9. Subject relative clause with covert extracted wh-NP

    the flowers that smelled enthralling

10. Adjunct (gap-less) relative clause with PP pied-piping

    the island on which the flowers smelled enthralling

11. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day the flowers smelled enthralling

A.13   n0Vpn1

1. Canonical subject

    Lisa ventures into the cave

2. Wh-moved subject

    who ventures into the cave

3. Wh-moved object of the PP

    what did Lisa venture into

4. Wh-moved PP

    into what does Lisa venture

5. Imperative

    venture into the cave

6. PRO subject

    Lisa wants [to venture into the cave]

7. Passive with by-phrase

    the cave is ventured into by Lisa

8. Passive without by-phrase

    the cave is ventured into

9. Passive with wh-moved by-phrase

    by whom is the cave ventured into

10. Passive with wh-moved object out of the by-phrase

    who was the cave ventured into by

11. Passive with wh-moved subject and by-phrase
    what is ventured into by Lisa

12. Passive with wh-moved subject and without by-phrase
    what is ventured into

13. Passive with PRO subject with by-phrase
    the cave likes [being ventured into by Lisa]

14. Passive with PRO subject without by-phrase
    the cave likes [being ventured into]

15. Subject relative clause with overt extracted NP's
    the man who ventures into the cave

16. Subject relative clause with covert extracted NP's
    the man that ventures into the cave

17. Object of the PP relative clause with overt extracted wh-NP
    the cave which Lisa ventures into

18. Object of the PP relative clause with covert extracted wh-NP
    the cave that Lisa ventures into

19. Object of the PP relative clause with PP pied-piping
    the cave into which Lisa ventured

20. Passive with relative clause on the object of by-phrase with overt extracted wh-NP
    the man who the cave was ventured into by

21. Passive with relative clause on the object of by-phrase with covert extracted wh-NP
    the man that the cave was ventured into by

22. Passive with relative clause on the object of by-phrase with PP pied-piping
    the man by whom the cave was ventured into

23. Passive with relative clause on subject and by-phrase with overt extracted wh-NP
    the cave which was ventured into by Lisa

24. Passive with relative clause on subject and by-phrase with covert extracted wh-NP
    the cave that was ventured into by Lisa

25. Passive with relative clause on subject without by-phrase with overt extracted wh-NP
    the cave which was ventured into

26. Passive with relative clause on subject without by-phrase with covert extracted wh-NP
    the cave that was ventured into

27. Adjunct (gap-less) relative clause with PP pied-piping
    the island on which Lisa ventures into the cave

28. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the island on which the cave was ventured into by Lisa

29. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase

    the island on which the cave was ventured into

30. Adjunct relative clause with covert extracted wh-NP
    the day Lisa ventured into the cave

31. Adjunct relative clause with covert extracted wh-NP with passive with by-phrase
    the day the cave was ventured into by Lisa

32. Adjunct relative clause with covert extracted wh-NP with passive without by-phrase
    the day the cave was ventured into

33. Determiner gerund
    [the venturing into the cave]

34. NP gerund
    Lisa wants [me venturing into the cave]

35. Gerund passive with by-phrase
    Bill approves of [it being ventured into by Lisa]

36. Gerund passive without by-phrase
    Bill approves of [it being ventured into]

37. NP gerund with PRO subject
    Lisa fears [PRO venturing into the cave]

38. NP gerund passive with PRO subject with by-phrase
    the cave approves of [PRO being ventured into by Lisa]

39. NP gerund passive with PRO subject without by-phrase
    the cave approves of [PRO being ventured into]

A.14   n0VPn1

1. Declarative
    Yossarian thinks of the chaplain

2. Wh-moved subject
    who thinks of the chaplain

3. Wh-moved object of the PP
    who does Yossarian think of

4. Wh-moved PP
    of whom does Yossarian think

5. Imperative
    think of the chaplain

6. PRO subject
    Yossarian wants [PRO to think of the chaplain]

7. Passive with by-phrase

    the chaplain is thought of by Yossarian

8. Passive without by-phrase

    the chaplain is thought of

9. Passive with wh-moved by-phrase

    by whom is the chaplain thought of

10. Passive with wh-moved object out of the by-phrase

    who was the chaplain thought of by

11. Passive with wh-moved subject and by-phrase

    who is thought of by Yossarian

12. Passive with wh-moved subject and without by-phrase

    who is thought of

13. Passive with PRO subject with by-phrase

    the chaplain approves of [being thought of by Yossarian]

14. Passive with PRO subject without by-phrase

    the chaplain approves of [being thought of]

15. Subject relative clause with overt extracted NP's

    the bombardier who thinks of the chaplain

16. Subject relative clause with covert extracted NP's

    the bombardier that thinks of the chaplain

17. Object of the PP relative clause with overt extracted wh-NP

    the chaplain who Yossarian thinks of

18. Object of the PP relative clause with covert extracted wh-NP

    the chaplain that Yossarian thinks of

19. Object of the PP relative clause with PP pied-piping

    the chaplain of whom Yossarian thinks

20. Passive with relative clause on the object of by-phrase with overt extracted wh-NP

    the bombardier who the chaplain was thought of by

21. Passive with relative clause on the object of by-phrase with covert extracted wh-NP

    the bombardier that the chaplain was thought of by

22. Passive with relative clause on the object of by-phrase with PP pied-piping

    the bombardier by whom the chaplain was thought of

23. Passive with relative clause on subject and by-phrase with overt extracted wh-NP

    the chaplain who was thought of by Yossarian

24. Passive with relative clause on subject and by-phrase with covert extracted wh-NP

    the chaplain that was thought of by Yossarian

25. Passive with relative clause on subject without by-phrase with overt extracted wh-NP

    the chaplain who was thought of

26. Passive with relative clause on subject without by-phrase with covert extracted wh-NP
    the chaplain that was thought of

27. Adjunct (gap-less) relative clause with PP pied-piping
    the island on which Yossarian thinks of the chaplain
    the island where Yossarian thinks of the chaplain

28. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the island on which the chaplain was thought of by Yossarian

29. Adjunct (gap-less) relative clause with PP pied-piping with passive without by-phrase
    the island on which the chaplain was thought of

30. Adjunct relative clause with covert extracted wh-NP
    the day Yossarian thinks of the chaplain

31. Adjunct relative clause with covert extracted wh-NP with passive with by-phrase
    the day the chaplain was thought of by Yossarian

32. Adjunct relative clause with covert extracted wh-NP with passive without by-phrase
    the day the chaplain was thought of

33. Determiner gerund
    [his thinking of the chaplain]

34. NP gerund
    the colonel wants [him thinking of the chaplain]

35. Gerund passive with by-phrase
    the colonel wants [him being thought of by Yossarian]

36. Gerund passive without by-phrase
    the colonel wants [him being thought of]

37. NP gerund with PRO subject
    Yossarian likes [PRO thinking of the chaplain]

38. NP gerund passive with PRO subject with by-phrase
    the chaplain likes [PRO being thought of by Yossarian]

39. NP gerund passive with PRO subject without by-phrase
    the chaplain likes [PRO being thought of]

A.15   s0V

1. Declarative
   to love the chaplain matters considerably
   that Yossarian loves the chaplain matters considerably

2. Wh-moved subject
   what matters considerably

3. Adjunct (gap-less) relative clause with PP pied-piping
    the island where to love the chaplain matters considerably

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day to love the chaplain matters considerably

A.16   s0Vn1

1. Declarative
    that Milo sold the cotton impressed Yossarian
    to even sell the cotton impresses Yossarian

2. Wh-moved subject
    what impresses Yossarian (that Milo sold the cotton)

3. Adjunct (gap-less) relative clause with PP pied-piping
    the island where to sell the cotton impresses Yossarian

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day (that) Milo sold the cotton impressed Yossarian

A.17   s0Vton1

1. Declarative
    to love the chaplain occurred to Yossarian
    that he loves the chaplain occurs to Yossarian
    * that he loves the chaplain occur to Yossarian

2. Wh-moved subject
    what occurred to Yossarian (that he loves the chaplain)

3. Adjunct (gap-less) relative clause with PP pied-piping
    the island where to love the chaplain occurred to Yossarian

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day (that) he loves the chaplain occurred to Yossarian

A.18   n0lVN1

1. Declarative
    Bill took notice

2. Wh-moved subject
    who took notice

3. Subject relative clause with overt extracted wh-NP
    the man who took notice

4. Subject relative clause with covert extracted wh-NP

    the man that took notice
    the man to take notice

5. Adjunct (gap-less) relative clause with PP pied-piping

    the room where Bill took notice
    the room in which Bill took notice

6. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day that Bill took notice

7. Imperative

    take notice

8. Determiner gerund

    Bill approves of [the taking of notice]

9. NP gerund

    Lisa approves of [him taking notice]

10. PRO subject

    Bill wants to take notice

11. NP gerund with PRO subject

    Lisa approves of taking notice

A.19   n0lVN1Pn2

1. Declarative

    Bill makes Lisa a promise

2. Wh-moved subject

    who makes Lisa a promise

3. Wh-moved indirect object

    whom does Bill make a promise
    * who does Bill makes a promise

4. Subject relative clause with overt extracted wh-NP

    the man who made Lisa a promise

5. Subject relative clause with covert extracted wh-NP

    the man that made Lisa a promise

6. Adjunct (gap-less) relative clause with PP pied-piping

    the room where Bill makes Lisa a promise

7. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase

    the room where Lisa was made a promise by Bill

8. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day Bill made Lisa a promise

9. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
   the day Lisa was made a promise by Bill

10. Indirect object relative clause with overt extracted wh-NP
    the woman whom Bill made a promise

11. Indirect object relative clause with covert extracted wh-NP
    the woman that Bill made a promise

12. Imperative
    make Lisa a promise

13. NP gerund
    Bill approves of him making Lisa a promise

14. Passive with by-phrase
    Lisa was made a promise by Bill

15. Gerund passive with by-phrase
    Bill likes [him being made a promise by Lisa]

16. Gerund passive without by-phrase
    Bill likes [him being made a promise]

17. PRO subject
    Bill wants [to make Lisa a promise]

18. NP gerund with PRO subject
    Bill approves of making Lisa a promise

19. Passive with PRO subject with by-phrase
    Lisa wants to be made a promise by Bill

20. NP gerund passive with by-phrase with PRO subject
    Lisa likes [PRO being made a promise by Bill]

21. NP gerund passive without by-phrase with PRO subject
    Lisa likes [PRO being made a promise]

22. Declarative
    Bill makes a promise to Lisa

23. Passive with by-phrase
    a promise was made to Lisa by Bill

24. Wh-moved subject
    who makes a promise to Lisa

25. Wh-moved object of PP
    who does Bill make a promise to

26. Wh-moved PP
    to whom does Bill make a promise

27. Subject relative clause with overt extracted wh-NP

the man who made a promise to Lisa

28. Subject relative clause with covert extracted wh-NP
    the man that made a promise to Lisa

29. Object of PP relative clause with overt extracted wh-NP
    the woman who Bill made a promise to

30. Object of PP relative clause with covert extracted wh-NP
    the woman that Bill made a promise to

31. Object of PP relative clause with PP pied-piping
    the woman to whom Bill made a promise

32. Adjunct (gap-less) relative clause with PP pied-piping
    the room where Bill makes a promise to Lisa

33. Adjunct (gap-less) relative clause with PP pied-piping with passive with by-phrase
    the room where a promise was made to Lisa by Bill

34. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day Bill made a promise to Lisa

35. Adjunct (gap-less) relative clause with covert extracted wh-NP with passive with by-phrase
    the day a promise was made to Lisa by Bill

36. Imperative
    make a promise to Lisa

37. Determiner gerund
    the making of a promise to Lisa

38. NP gerund
    Bill approves of him making a promise to Lisa

39. Gerund passive with by-phrase
    a promise being made to Lisa by Bill

40. Gerund passive without by-phrase
    a promise being made to Lisa

41. PRO subject
    Bill wants to make a promise to Lisa

42. NP gerund with PRO subject
    Bill approves of [PRO making a promise to Lisa]

A.20   ItVn1s2

1. Declarative
   it is Milo that runs the Syndica

2. Wh-moved object with be not inverted

who might it be that likes Lisa

3. Adjunct (gap-less) relative clause with PP pied-piping
    the island where it is Milo that runs the Syndicate

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day it is Milo that runs the Syndicate

A.21 ItVpn1s2

1. Declarative
    it is on Pianosa island that Milo runs the Syndicate

2. Wh-moved prepositional phrase with be not inverted
    where might it be that Milo runs the Syndicate

3. Adjunct (gap-less) relative clause with PP pied-piping
    the country where it is on Pianosa island that Milo runs the Syndicate

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day it is on Pianosa island that Milo runs the Syndicate

A.22 ItVad1s2

1. Declarative
    it is reluctantly that Milo runs the Syndicate

2. Wh-moved prepositional phrase with be not inverted
    how might it be that Milo runs the Syndicate

3. Adjunct (gap-less) relative clause with PP pied-piping
    the island where it is reluctantly that Milo runs the Syndicate

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day it is reluctantly that Milo runs the Syndicate

A.23 n0A1

1. Declarative
    cannabis is forbidden

2. Wh-moved subject
    what is forbidden

3. Wh-moved adjective how
    how is cannabis

4. Imperative

    be forbidden

5. Relative clause on subject with overt extracted wh-NP

       the cannabis which is forbidden

6. Relative clause on subject with covert extracted wh-NP

       the cannabis that is forbidden

7. Adjunct (gap-less) relative clause with with PP pied-piping

       the country where cannabis is forbidden

8. Adjunct (gap-less) relative clause with covert extracted wh-NP

       the day cannabis was forbidden

9. NP gerund

       the cannabis being forbidden

10. PRO subject

        the cannabis wants [to be forbidden]

11. NP gerund with PRO subject

        the cannabis approves of [being forbidden]

A.24   n0A1s1

1. Declarative

       Milo was able to sell the cotton

2. Wh-moved subject

       who was able to sell the cotton

3. Imperative

       be able to sell the cotton

4. Relative clause on subject with overt extracted wh-NP

       the bombardier who was able to sell the cotton

5. Relative clause on subject with covert extracted wh-NP

       the bombardier that was able to sell the cotton

6. Adjunct (gap-less) relative clause with with PP pied-piping

       the island where Milo was able to sell the cotton

7. Adjunct (gap-less) relative clause with covert extracted wh-NP

       the day Milo was able to sell the cotton

8. NP gerund

       Milo being able to sell the cotton

9. PRO subject

       Milo wants to be able to sell the cotton

10. NP gerund with PRO subject

Milo likes being able to sell the cotton

A.25   s0A1

1. Declarative

   to drive a Jaguar is blissful

2. Wh-moved subject

   what is blissful

3. Adjunct (gap-less) relative clause with with PP pied-piping

   the country where to drive a Jaguar is blissful

4. Adjunct (gap-less) relative clause with covert extracted wh-NP

   the day to drive a Jaguar is blissful

A.26   n0BEn1

1. Declarative

   the bombardier is this man

2. Inverted question

   is the bomardier this man

A.27   n0N1

1. Declarative

   Raveesh is Hindi
   everyone considers Raveesh Hindi

2. Wh-moved subject

   who is Hindi

3. Wh-moved object

   what is Raveesh

4. Imperative

   be Hindi

5. Relative clause on subject with overt extracted wh-NP

   the man who is Hindi

6. Relative clause on subject with covert extracted wh-NP

   the man that is Hindi

7. Adjunct (gap-less) relative clause with with PP pied-piping

   the country where everyone considers Raveesh Hindi

8. Adjunct (gap-less) relative clause with covert extracted wh-NP

   the day everyone considered Raveesh Hindi

9. NP gerund

   Raveesh being Hindi

10. PRO subject

    Raveesh wants to be Hindi

11. NP gerund with PRO subject

    Ravesh likes [being Hindi]

A.28   n0N1s1

1. Declarative

   these were accusations that Orr ate the chestnuts

2. Wh-moved subject

   which documents were accusations that Orr ate the chestnuts

3. Imperative

   be accusations that Orr ate the chestnuts

4. Relative clause on subject with overt extracted wh-NP

   the documents which were accusations that Orr ate the chestnuts

5. Relative clause on subject with covert extracted wh-NP

   the documents that were accusations that Orr ate the chestnuts

6. Adjunct (gap-less) relative clause with with PP pied-piping

   the island where these were accusations that Orr ate the chestnuts

7. Adjunct (gap-less) relative clause with covert extracted wh-NP

   the day these were accusations that Orr ate the chestnuts

8. NP gerund

   these being accusations that Orr ate the chestnuts

9. PRO subject

   these want [to be accusations that Orr ate the chestnuts]

10. NP gerund with PRO subject

    being accusations that Orr ate the chestnuts

A.29   s0N1

1. Declarative

   to fly many missions is insanity

2. Wh-moved subject

what is insanity

3. Adjunct (gap-less) relative clause with with PP pied-piping
   the island where to fly many missions was insanity

4. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day that to fly many missions was insanity

A.30   n0Pn1

1. Declarative
   the key is in the chest

2. Wh-moved subject
   what is in the chest

3. Wh-moved PP
   in what were the keys

4. Wh-moved object of PP
   what were the keys in

5. Imperative
   be in the chest

6. Relative clause on subject with overt extracted wh-NP
   the keys which were in the chest

7. Relative clause on object of PP with overt extracted wh-NP
   the chest which the keys were in

8. Relative clause on object of PP with PP pied-piping
   the chest in which the keys were

9. Relative clause on subject with covert extracted wh-NP
   the keys that were in the chest

10. Relative clause on object of PP with covert extracted wh-NP
    the chest that the keys were in

11. Adjunct (gap-less) relative clause with with PP pied-piping
    the island where the keys were in the chest

12. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day the keys were in the chest

13. NP gerund
    the keys being in the chest

14. PRO subject
    Orr wants [to be in the chest]

15. NP gerund with PRO subject

    Orr approves of [being in the chest]

A.31   n0P1

1. Declarative
   Lisa is abroad

2. Wh-moved subject
   who is abroad

3. Wh-moved PP
   where is Lisa

4. Imperative
   be abroad

5. Relative clause on subject with overt extracted wh-NP
   the woman who is abroad

6. Relative clause on subject with covert extracted wh-NP
   the woman that is abroad

7. Adjunct (gap-less) relative clause with with PP pied-piping
   the island where Lisa is abroad

8. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day Lisa was abroad

9. NP gerund
   Lisa being abroad

10. PRO subject
    Lisa wants [to be abroad]

11. NP gerund with PRO subject
    Lisa approves of [being abroad]

A.32   s0Pn1

1. Declarative
   to sell the cotton is beyond belief

2. Wh-moved subject
   what is beyond belief

3. Relative clause on object of the PP with overt extracted wh-NP
   the belief which to sell the cotton is beyond

4. Relative clause on object of the PP with covert extracted wh-NP
   the belief that to sell the cotton is beyond

5. NO !!! Relative clause on object of the PP with PP pied-piping
   the belief beyond which to sell the cotton is

6. Adjunct (gap-less) relative clause with PP pied-piping
   the island where to sell the cotton is beyond belief

7. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day to sell the cotton is beyond belief

A.33   n0ARBPn1

1. Declarative
   Bill is ahead of Lisa

2. Wh-moved subject
   who is ahead of Lisa

3. Wh-moved PP
   ahead of whom is Bill

4. Wh-moved object of PP
   who is Bill ahead of

5. Imperative
   be ahead of Lisa

6. Relative clause on subject with overt extracted wh-NP
   the man who is ahead of Lisa

7. Relative clause on subject with overt extracted wh-NP
   the man that is ahead of Lisa

8. Relative clause on object of PP with overt extracted wh-NP
   the woman whom Bill is ahead of

9. Relative clause on object of PP with overt extracted wh-NP
   the woman that Bill is ahead of

10. Adjunct (gap-less) relative clause with PP pied-piping
    the island where Bill is ahead of Lisa

11. Adjunct (gap-less) relative clause with PP covert extracted wh-NP
    the day Bill was ahead of Lisa

12. NP gerund
    Bill being ahead of Lisa

13. PRO subject
    Bill wanted to be ahead of Lisa

14. NP gerund with PRO subject
    Bill approves of [being ahead of Lisa]

A.34   n0APn1

1. Declarative
   Bill is void of hope

2. Wh-moved subject
   who is void of hope

3. Wh-moved object of PP
   what is Bill void of

4. Relative clause on subject with overt extracted wh-NP
   the man who is void of hope

5. Relative clause on subject with covert extracted wh-NP
   the man that is void of hope

6. Relative clause on object of PP with overt extracted wh-NP
   the hope which Bill is void of

7. Relative clause on object of PP with overt extracted wh-NP
   the hope that Bill is void of

8. Adjunct (gap-less) relative clause with PP pied-piping
   the island where Bill is void of hope

9. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day Bill was void of hope

10. PRO subject
    Bill wants to be void of hope

A.35   n0NPn1

1. Declarative
   the innovation was thanks to the chaplain

2. Wh-moved subject
   what is thanks to the chaplain

3. Wh-moved object of PP
   whom was the innovation thanks to

4. Relative clause on subject with overt extracted wh-NP
   the innovation which was thanks to the chaplain

5. Relative clause on subject with covert extracted wh-NP
   the innovation that was thanks to the chaplain

6. Relative clause on object of PP with overt extracted wh-NP
   the chaplain whom the innovation is thanks to

7. Relative clause on object of PP with overt extracted wh-NP
   the chaplain that the innovation was thanks to

8. Adjunct (gap-less) relative clause with PP pied-piping
   the island where the innovation was thanks to the chaplain

9. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day the innovation was thanks to the chaplain

10. PRO subject
    the innovation wants [to be thanks to the chaplain]

A.36   n0PPn1

1. Declarative
   the detective is on to Bill

2. Wh-moved subject
   who is on to Bill

3. Imperative
   be on to Bill

4. Wh-moved PP
   on to whom is the detective

5. Wh-moved object of PP
   whom was the detective on to

6. Relative clause on subject with overt extracted wh-NP
   the detective who is on to Bill

7. Relative clause on subject with covert extracted wh-NP
   the detective that was on to Bill

8. Relative clause on object of PP with overt extracted wh-NP
   the man whom the detective is on to

9. Relative clause on object of PP with covert extracted wh-NP
   the man that the detective is on to

10. Relative clause on object of PP with PP pied-piping
    the man on to whom the detective is

11. Adjunct (gap-less) relative clause with PP pied-piping
    the island where the detective was on to Bill

12. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day the detective was on to Bill

13. PRO subject
    the detective wants [to be on to Bill]

14. NP gerund

    the detective being on to Bill

15. NP gerund with PRO subject

    [being on to Bill]

A.37 n0PNaPn1

1. Declarative

   the chaplain was in charge of Yossarian

2. Wh-moved subject

   who was in charge of Yossarian

3. Imperative

   be in charge of Yossarian

4. Wh-moved PP

   in charge of whom was the chaplain

5. Wh-moved object of PP

   whom was the chaplain in charge of

6. Relative clause on subject with overt extracted wh-NP

   the chaplain who was in charge of Yossarian

7. Relative clause on subject with covert extracted wh-NP

   the chaplain that was in charge of Yossarian

8. Relative clause on object of PP with overt extracted wh-NP

   the bombardier whom the chaplain was in charge of

9. Relative clause on object of PP with covert extracted wh-NP

   the bombardier that the chaplain was in charge of

10. Relative clause on object of PP with PP pied-piping

    the bombardier in charge of whom the chaplain was

11. Adjunct (gap-less) relative clause with PP pied-piping

    the island where the chaplain was in charge of Yossarian

12. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day the chaplain was in charge of Yossarian

13. PRO subject

    the chaplain wants [to be in charge of Yossarian]

14. NP gerund

    the colonel approves of [the chaplain being in charge of Yossarian]

15. NP gerund with PRO subject

    [being in charge of Yossarian]

A.38   s0ARBPn1

1. Declarative
   to bake a cake was contrary to the expectations

2. Wh-moved subject
   what was contrary to the expectations

3. Relative clause on object of PP with overt extracted wh-NP
   the expectations which to bake a cake was contrary to

4. Relative clause on object of PP with covert extracted wh-NP
   the expectations that to bake a cake was contrary to

5. Adjunct (gap-less) relative clause with PP pied-piping
   the island where to bake a cake was contrary to the expectations

6. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day to bake a cake was contrary to the expectations

A.39   s0APn1

1. Declarative
   to kill a mockingbird is devoid of sense

2. Wh-moved subject
   what is devoid of sense

3. Relative clause on object of PP with overt extracted wh-NP
   the sense which to kill a mockingbird is devoid of

4. Relative clause on object of PP with covert extracted wh-NP
   the sense that to kill a mockingbird is devoid of

5. Adjunct (gap-less) relative clause with PP pied-piping
   the country where to kill a mockingbird is devoid of sense

6. Adjunct (gap-less) relative clause with covert extracted wh-NP
   the day to kill a mockingbird was devoid of sense

A.40   s0NPn1

1. Declarative
   to run a marathon was thanks to his spirit

2. Wh-moved subject
   what is thanks to his spirit

3. Relative clause on object of PP with overt extracted wh-NP
   his spirit which to run a marathon was thanks to

4. Relative clause on object of PP with covert extracted wh-NP

    his spirit that to run a marathon is thanks to

5. Adjunct (gap-less) relative clause with PP pied-piping

    the island where to run a marathon was thanks to his spirit

6. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day to run a marathon was thanks to his spirit

A.41   s0PPn1

1. Declarative

    to kill a mockingbird was outside of his intentions

2. Wh-moved subject

    what was outside of his intentions

3. Relative clause on object of PP with overt extracted wh-NP

    his intentions which to kill a mockingbird was outside of

4. Relative clause on object of PP with covert extracted wh-NP

    his intentions that to kill a mockingbird was outside of

5. Adjunct (gap-less) relative clause with PP pied-piping

    the island where to kill a mockingbird was outside of his intentions

6. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day to kill a mockingbird was outside of his intentions

A.42   s0PNaPn1

1. Declarative

    that Bill ran a marathon was on account of his spirit

2. Wh-moved subject

    what was on account of his spirit

3. Relative clause on object of PP with overt extracted wh-NP

    his spirit which to run a marathon was on account of

4. Relative clause on object of PP with covert extracted wh-NP

    his spirit that to run a marathon was on account of

5. Adjunct (gap-less) relative clause with PP pied-piping

    the island where to run a marathon was on account of his spirit

6. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day that Bill ran a marathon was on account of his spirit

A.43 s0Vs1

1. Declarative

    that Bill drives a Jaguar makes Lisa angry

2. Wh-moved subject

    what makes Lisa angry

A.44 n0n1ARB

1. Declarative

    Lisa is many feet away

2. Wh-moved subject

    who is many feet away

3. Relative clause on subject with overt extracted wh-NP

    the woman who is many feet away

4. Relative clause on subject with overt extracted wh-NP

    the woman that is many feet away

5. Wh-moved locative

    how many feet away is Lisa

6. Imperative

    be many feet away

7. Wh-moved locative

    what is Lisa away

8. NP gerund

    Lisa being many feet away

9. PRO subject

    Lisa wants to be many feet away

10. NP gerund with PRO subject

    being many feet away

A.45 Xn0Vs1

1. Declarative

    Lisa expects Bill to sleep

2. Relative clause on subject with overt extracted wh-NP

    the woman who expects Bill to sleep

3. Relative clause on subject with covert extracted wh-NP

    the woman that expects Bill to sleep

4. Wh-moved subject

    who expects Bill to sleep

5. Imperative

    expect Bill to sleep

6. Adjunct (gap-less) relative clause with PP pied-piping

    the room where Lisa expects Bill to sleep

7. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day Lisa expects Bill to sleep

8. NP gerund

    Lisa expecting Bill to sleep

9. PRO subject

    Lisa wants to expect Bill to sleep

A.46   n0VDN1

1. Declarative

    Bill bit the dust

2. Wh-moved subject

    who bit the dust

3. Imperative

    bite the dust

4. Subject relative clause with overt extracted wh-NP

    the man who bit the dust

5. Subject relative clause with covert extracted wh-NP

    the man that bit the dust

6. Passive without by-phrase

    the dust was bitten

7. Passive without by-phrase

    the dust was bitten by Bill

8. Wh-moved object of by-phrase

    whom was bitten the dust by

9. Wh-moved by-phrase

    by whom was the dust bitten

10. Relative with overt extracted wh-NP on subject of passive

    the man whom the dust was bitten by

11. Relative with covert extracted wh-NP on subject of passive

    the man that the dust was bitten by

12. Adjunct (gapless) relative with covert extracted wh-NP
    the day Bill bit the dust

13. Adjunct (gapless) relative with PP pied-piping
    the island where Bill bit the dust

14. NP gerund
    Bill biting the dust

15. PRO subject
    Bill wanted [to bite the dust]

16. NP gerund with PRO subject
    Bill approves of [biting the dust]

A.47   n0VDAN1

1. Declarative
   Lisa sang a different tune

2. Wh-moved subject
   who sang a different tune

3. Imperative
   sing a different tune

4. Subject relative clause with overt extracted wh-NP
   the woman who sang a different tune

5. Subject relative clause with covert extracted wh-NP
   the woman that sang a different tune

6. Passive without by-phrase
   a different tune was sung

7. Passive without by-phrase
   a different tune was sung by Lisa

8. Wh-moved object of by-phrase
   whom was a different tune sung by

9. Wh-moved by-phrase
   by whom was a different tune sung

10. Relative with overt extracted wh-NP on subject of passive
    the woman whom a different tune was sung by

11. Relative with covert extracted wh-NP on subject of passive
    the woman that a different tune was sung by

12. Adjunct (gapless) relative with covert extracted wh-NP
    the day Lisa sang a different tune

13. Adjunct (gapless) relative with PP pied-piping
    the island where Lisa sang a different tune

14. NP gerund
    Lisa singing a different tune

15. PRO subject
    Lisa wanted [to sing a different tune]

16. NP gerund with PRO subject
    Lisa approves of [singing a different tune]

A.48   n0VN1

1. Declarative
   Lisa cries wolf

2. Wh-moved subject
   who cries wolf

3. Imperative
   cry wolf

4. Subject relative clause with overt extracted wh-NP
   the woman who cried wolf

5. Subject relative clause with covert extracted wh-NP
   the woman that cried wolf

6. Passive without by-phrase
   wolf was cried

7. Passive without by-phrase
   wolf was cried by Lisa

8. Wh-moved object of by-phrase
   whom was wolf cried by

9. Wh-moved by-phrase
   by whom was wolf cried

10. Relative with overt extracted wh-NP on subject of passive
    the woman whom wolf was cried by

11. Relative with covert extracted wh-NP on subject of passive
    the woman that wolf was cried by

12. Adjunct (gapless) relative with covert extracted wh-NP
    the day Lisa cried wolf

13. Adjunct (gapless) relative with PP pied-piping
    the island where Lisa cried wolf

14. NP gerund

    Lisa crying wolf

15. PRO subject

    Lisa wanted [to cry wolf]

16. NP gerund with PRO subject

    Lisa approves of [crying wolf]

A.49   n0VAN1

1. Declarative

    Lisa cries bloody murder

2. Wh-moved subject

    who cries bloody murder

3. Imperative

    cry bloody murder

4. Subject relative clause with overt extracted wh-NP

    the woman who cried bloody murder

5. Subject relative clause with covert extracted wh-NP

    the woman that cried bloody murder

6. Passive without by-phrase

    bloody murder was cried

7. Passive without by-phrase

    bloody murder was cried by Lisa

8. Wh-moved object of by-phrase

    whom was bloody murder cried by

9. Wh-moved by-phrase

    by whom was bloody murder cried

10. Relative with overt extracted wh-NP on subject of passive

    the woman whom bloody murder was cried by

11. Relative with covert extracted wh-NP on subject of passive

    the woman that bloody murder was cried by

12. Adjunct (gapless) relative with covert extracted wh-NP

    the day Lisa cried bloody murder

13. Adjunct (gapless) relative with PP pied-piping

    the island where Lisa cried bloody murder

14. NP gerund

    Lisa crying bloody murder

15. PRO subject

    Lisa wanted [to cry bloody murder]

16. NP gerund with PRO subject

    Lisa approves of [crying bloody murder]

A.50   n0VDAN1Pn2

1. Declarative

   Lisa makes a big deal of the accusations

2. Wh-moved subject

   who makes a big deal of the accusations

3. Imperative

   make a big deal of the accusations

4. Subject relative clause with overt extracted wh-NP

   the woman who made a big deal of the accusations

5. Subject relative clause with covert extracted wh-NP

   the woman that made a big deal of the accusations

6. Passive without by-phrase

   a big deal was made of the accusations

7. Passive without by-phrase

   a big deal was made of the accusations by Lisa

8. Wh-moved object of by-phrase

   whom was a big deal made of the accusations by

9. Wh-moved by-phrase

   by whom was a big deal made of the accusations

10. Outer passive without by-phrase

    the accusations were made a big deal of

11. Outer passive with by-phrase

    the accusations were made a big deal of by Lisa

12. Outer passive with wh-moved by-phrase

    by whom were the accusations made a big deal of

13. Outer passive without by-phrase with relative on the subject with overt extracted wh-NP

    the accusations which were made a big deal of

14. Outer passive without by-phrase with relative on the subject with covert extracted wh-NP

    the accusations that were made a big deal of

15. Outer passive with by-phrase with relative on the subject with overt extracted wh-NP

    the accusations which were made a big deal of by Lisa

16. Outer passive with by-phrase with relative on the subject with covert extracted wh-NP
    the accusations that were made a big deal of by Lisa

17. Relative with overt extracted wh-NP on subject of passive
    the woman whom a big deal was made by

18. Relative with covert extracted wh-NP on subject of passive
    the woman that a big deal was made by

19. Adjunct (gapless) relative with covert extracted wh-NP
    the day Lisa made a big deal of the accusations

20. Adjunct (gapless) relative with PP pied-piping
    the island where Lisa made a big deal of the accusations

21. NP gerund
    [Lisa making a big deal of the accusations]

22. PRO subject
    Lisa wanted [to make a big deal of the accusations]

23. NP gerund with PRO subject
    Lisa approves of [making a big deal of the accusations]

A.51   n0VAN1Pn2

1. Declarative
   Lisa makes short work of the cake

2. Wh-moved subject
   who makes short work of the cake

3. Imperative
   make short work of the cake

4. Subject relative clause with overt extracted wh-NP
   the woman who made short work of the cake

5. Subject relative clause with covert extracted wh-NP
   the woman that made short work of the cake

6. Passive without by-phrase
   short work was made of the cake

7. Passive without by-phrase
   short work was made of the cake by Lisa

8. Wh-moved object of by-phrase
   whom was short work made of the cake by

9. Wh-moved by-phrase
   by whom was short work made of the cake

10. Outer passive without by-phrase

    the cake was made short work of

11. Outer passive with by-phrase

    the cake was made short work of by Lisa

12. Outer passive with by-phrase with PRO subject

    the cakes wants [to be made short work of by Lisa]

13. Outer passive without by-phrase with PRO subject

    the cake wants [to be made short work of]

14. Outer passive with wh-moved by-phrase

    by whom was the cake made short work of

15. Outer passive without by-phrase with relative on the subject with overt extracted wh-NP

    the cake which was made short work of

16. Outer passive without by-phrase with relative on the subject with covert extracted wh-NP

    the cake that was made short work of

17. Outer passive with by-phrase with relative on the subject with overt extracted wh-NP

    the cake which was made short work of by Lisa

18. Outer passive with by-phrase with relative on the subject with covert extracted wh-NP

    the cake that was made short work of by Lisa

19. Adjunct (gapless) relative with covert extracted wh-NP

    the day Lisa made short work of the cake

20. Adjunct (gapless) relative with PP pied-piping

    the island where Lisa made short work of the cake

21. NP gerund

    [Lisa making short work of the cake]

22. PRO subject

    Lisa wanted [to make short work of the cake]

23. NP gerund with PRO subject

    Lisa approves of [making short work of the cake]

A.52   n0VN1Pn2

1. Declarative

   Bill looks daggers at the detective

2. Wh-moved subject

   who looks daggers at the detective

3. Imperative

   look daggers at the detective

4. Subject relative clause with overt extracted wh-NP
   the man who looked daggers at the detective

5. Subject relative clause with covert extracted wh-NP
   the man that looked daggers at the detective

6. Passive without by-phrase
   daggers were looked at the detective

7. Passive without by-phrase
   daggers were looked at the detective by Bill

8. Wh-moved object of by-phrase
   whom were daggers looked at the detective by

9. Wh-moved by-phrase
   by whom were daggers looked at the detective

10. Outer passive without by-phrase
    the detective was looked daggers at

11. Outer passive with by-phrase
    the detective was looked daggers at by Bill

12. Outer passive with by-phrase with PRO subject
    the detectives wants [to be looked daggers at by Bill]

13. Outer passive without by-phrase with PRO subject
    the detective wants [to be looked daggers at]

14. Outer passive with wh-moved by-phrase
    by whom was the detective looked daggers at

15. Outer passive without by-phrase with relative on the subject with overt extracted wh-NP
    the detective who was looked daggers at

16. Outer passive without by-phrase with relative on the subject with covert extracted wh-NP
    the detective that was looked daggers at

17. Outer passive with by-phrase with relative on the subject with overt extracted wh-NP
    the detective who was looked daggers at by Bill

18. Outer passive with by-phrase with relative on the subject with covert extracted wh-NP
    the detective that was looked daggers at by Bill

19. Adjunct (gapless) relative with covert extracted wh-NP
    the day Bill looked daggers at the detective

20. Adjunct (gapless) relative with PP pied-piping
    the island where Bill looked daggers at the detective

21. NP gerund
    [Bill looking daggers at the detective]

22. PRO subject

Bill wanted [to look daggers at the detective]

23. NP gerund with PRO subject
    Bill approves of [looking daggers at the detective]

A.53   n0VDN1Pn2

1. Declarative
   Bill broke the news to his wife

2. Wh-moved subject
   who broke the news to his wife

3. Imperative
   break the news to his wife

4. Subject relative clause with overt extracted wh-NP
   the man who broke the news to his wife

5. Subject relative clause with covert extracted wh-NP
   the man that broke the news to his wife

6. Passive without by-phrase
   the news were broken to his wife

7. Passive without by-phrase
   the news were broken to his wife by Bill

8. Wh-moved object of by-phrase
   whom were the news broken to his wife by

9. Wh-moved by-phrase
   by whom were the news broken to his wife

10. Outer passive without by-phrase
    his wife were broken the news to

11. Outer passive with by-phrase
    his wife was broken the news to by Bill

12. Outer passive with by-phrase with PRO subject
    his wife wants [to be broken the news to by Bill]

13. Outer passive without by-phrase with PRO subject
    his wife wants [to be broken the news to]

14. Outer passive with wh-moved by-phrase
    by whom was his wife broken the news to

15. Outer passive without by-phrase with relative on the subject with overt extracted wh-NP
    his wife who was broken the news to

16. Outer passive without by-phrase with relative on the subject with covert extracted wh-NP

       his wife that was broken the news to

17. Outer passive with by-phrase with relative on the subject with overt extracted wh-NP
       his wife who was broken the news to by Bill

18. Outer passive with by-phrase with relative on the subject with covert extracted wh-NP
       his wife that was broken the news to by Bill

19. Adjunct (gapless) relative with covert extracted wh-NP
       the day Bill broke the news to his wife

20. Adjunct (gapless) relative with PP pied-piping
       the island where Bill broke the news to his wife

21. NP gerund
       [Bill breaking the news to his wife]

22. PRO subject
       Bill wants [to break the news to his wife]

23. NP gerund with PRO subject
       Bill approves of [breaking the news to his wife]

A.54 Rn0Vn1A2

1. Declarative
       the detective is beating Bill dead

2. Passive with by-phrase
       Bill was beaten dead by the detective

3. Passive with without by-phrase
       Bill was beaten dead

4. Wh-moved subject
       who is beating Bill dead

5. Subject relative clause with overt extracted wh-NP
       the detective who is beating Bill dead

6. Subject relative clause with covert extracted wh-NP
       the detective that is beating Bill dead

7. Wh-moved object
       who did the detective beat dead

8. Object relative clause with overt extracted wh-NP
       the man who the detective is beating dead

9. Object relative clause with covert extracted wh-NP
       the man that the detective is beating dead

10. Passive with by-phrase on wh-moved object

    who was beaten dead by the detective

11. Passive with by-phrase on object relative clause with overt extracted wh- NP
    the man who was beaten dead by the detective

12. Passive without by-phrase on object relative clause with overt extracted wh- NP
    the man who was beaten dead

13. Passive without by-phrase on object relative clause with covert extracted wh- NP
    the man that was beaten dead

14. Passive with by-phrase on object relative clause with covert extracted wh- NP
    the man that was beaten dead by the detective

15. Imperative
    beat Bill dead

16. Wh-moved subject on passive without by-phrase
    who was beaten dead

17. Wh-question on object of by-phrase in subject extracted relative clauses with overt extracted wh-NP
    the detective who Bill was beaten dead by

18. Wh-question on object of by-phrase in subject extracted relative clauses with covert extracted wh-NP
    the detective that Bill was beaten dead by

19. Wh-question on object of by-phrase in subject extracted relative clauses with PP pied-piping
    the detective by whom Bill was beaten dead

20. Passive with wh-moved object of by-phrase
    who was Bill beaten dead by

21. Passive with wh-moved by-phrase
    by whom was Bill beaten dead

22. Multi anchored participial modifiers
    the beaten dead man

23. Relative clause on PP adjunct with covert extracted wh-NP
    the day the detective was beating Bill dead

24. Relative clause on PP adjunct with PP pied-piping
    the island where the detective was beating Bill dead

25. Relative clause on PP adjunct in passives with by-phrase with PP pied-piping
    the island where Bill was beaten dead by the detective

26. Relative clause on PP adjunct in passives with by-phrase with covert extracted wh-NP
    the day Bill was beaten dead by the detective

27. Relative clause on PP adjunct in passives without by-phrase with PP pied-piping

> *the island where Bill was beaten dead*

28. Relative clause on PP adjunct in passives without by-phrase with covert extracted wh-NP
    the day Bill was beaten dead

29. NP gerund
    the detective beating Bill dead

30. NP gerund with by-phrase
    Bill being beaten dead by the detective

31. NP gerund without by-phrase
    Bill being beaten dead

32. Wh-moved adjective complement
    how did the detective beat Bill

33. Passive on wh-moved adjective complement with by phrase
    how was Bill beaten by the detective

34. Passive on wh-moved adjective complement without by phrase
    how was Bill beaten

A.55   Rn0Vn1Pn2

1. Declarative
   Bill ran his shoes into pieces

2. Passive with by-phrase
   the shoes were run into pieces by Bill

3. Passive without by-phrase
   the shoes were run into pieces

4. Wh-moved subject
   who ran the shoes run into pieces

5. Wh-moved object
   what did Bill run into pieces

6. Wh-moved by-phrase
   by whom were the shoes into pieces

7. Wh-moved PP
   into what did Bill run his shoes

8. Wh-moved PP in passive with by-phrase
   into what were the shoes run by Bill

9. Wh-moved PP in passive without by-phrase
   into what were the shoes run

10. Wh-moved object of PP

    what did Bill run his shoes into

11. Wh-moved object of PP in passive with by-phrase
    what were the shoes run into by Bill

12. Wh-moved object of PP in passive without by-phrase
    what were the shoes run into

13. Relative clause on subject with overt extracted wh-NP
    the man who ran his shoes into pieces

14. Relative clause on subject with covert extracted wh-NP
    the man that ran his shoes into pieces

15. Relative clause on object with overt extracted wh-NP
    the shoes which Bill ran into pieces

16. Relative clause on object with covert extracted wh-NP
    the shoes that Bill ran into pieces

17. Wh-question on object with passive with by-phrase
    what was run into pieces by Bill

18. Wh-question on object with passive without by-phrase
    what was run into pieces

19. Wh-moved object of by-phrase
    who were the shoes run into pieces

20. Relative clause on object in passive with by-phrase with overt extracted wh-NP
    the shoes which were run into pieces by Bill

21. Relative clause on object in passive with by-phrase with covert extracted wh-NP
    the shoes that were run into pieces by Bill

22. Relative clause on object in passive with by-phrase with overt extracted wh-NP
    the shoes which were run into pieces

23. Relative clause on object in passive with by-phrase with covert extracted wh-NP
    the shoes that were run into pieces

24. Relative clause on object of by-phrase in passive
    the man whom the shoes were run into pieces by

25. Relative clause on object of by-phrase in passive
    the man that the shoes were run into pieces by

26. Relative clause on subject with PP pied-piping
    the man by whom the shoes were run into pieces

27. Relative clause on PP with PP pied-piping
    the pieces into which Bill ran his shoes

28. Relative clause on PP in passive with by-phrase with PP pied-piping
    the pieces into which the shoes were run by Bill

29. Relative clause on PP in passive without by-phrase with PP pied-piping
    the pieces into which the shoes were run

30. Relative clause on object of by-phrase with covert extracted wh-NP
    the pieces that Bill ran his shoes into

31. Relative clause on object of by-phrase with overt extracted wh-NP
    the pieces which Bill ran his shoes into

32. Relative clause on object of PP in passive with by-phrase with covert extracted wh-NP
    the pieces that the shoes were run into by Bill

33. Relative clause on object of PP in passive with by-phrase with overt extracted wh-NP
    the pieces which the shoes were run into by Bill

34. Relative clause on object of PP in passive without by-phrase with covert extracted wh-NP
    the pieces that the shoes were run into

35. Relative clause on object of PP in passive without by-phrase with overt extracted wh-NP
    the pieces which the shoes were run into

36. Imperative
    run the shoes into pieces

37. Adjunct (gap-less) relative clause with PP pied-piping
    the island where Bill ran the shoes into pieces

38. Adjunct (gap-less) relative clause with covert extracted wh-NP
    the day Bill ran the shoes into pieces

39. Adjunct (gap-less) relative clause in passive with by-phrase with PP pied-piping
    the island where the shoes were run into pieces by Bill

40. Adjunct (gap-less) relative clause in passive with by-phrase with covert extracted wh-NP
    the day the shoes were run into pieces by Bill

41. Adjunct (gap-less) relative clause in passive without by-phrase with PP pied-piping
    the island where the shoes were run into pieces

42. Adjunct (gap-less) relative clause in passive without by-phrase with covert extracted wh-NP
    the day the shoes were run into pieces

43. NP gerund
    Bill running the shoes into pieces

44. NP gerund in passive with by-phrase
    the shoes being run into pieces by Bill

45. NP gerund in passive without by-phrase
    the shoes being run into pieces

A.56 REn1VA2

1. Declarative
   the snow froze solid

2. Wh-moved subject
   what froze solid

3. Subject relative clause with overt extracted wh-NP
   the snow which froze solid

4. Subject relative clause with overt extracted wh-NP
   the snow that froze solid

5. Imperative
   freeze solid

6. Relative clause on PP adjunct with PP pied-piping
   the island where the snow froze solid

7. Relative clause with covert extracted wh-NP
   the day the snow froze solid

8. NP gerund
   the snow freezing solid

9. Wh-moved adjective complement
   how did the snow freeze

A.57 REn1VPn2

1. Declarative
   the snow is melting into a puddle

2. Wh-moved subject
   what is melting into a puddle

3. Wh-moved PP
   into what did the snow melt

4. Wh-moved object of PP
   what did the snow melt into

5. Relative clause on subject with overt extracted wh-NP
   the snow which is melting into a puddle

6. Relative clause on subject with covert extracted wh-NP
   the snow that is melting into a puddle

7. Relative clause on object of PP with overt extracted wh-NP
   the puddle which the snow is melting into

8. Relative clause on object of PP with covert extracted wh-NP

    the puddle that the snow is melting into

9. Relative clause on PP with PP pied-piping

    the puddle into which the snow is melting

10. Adjunct (gap-less) relative clause with PP pied-piping

    the island where the snow is melting into a puddle

11. Adjunct (gap-less) relative clause with covert extracted wh-NP

    the day the snow was melting into a puddle

12. NP gerund

    the snow melting into a puddle

13. Imperative

    melt into a puddle