Polysemy and Homonymy in Japanese Verbal Alternations

Stephen Nightingale

A thesis submitted in fulfilment of requirements for the degree of Doctor of Philosophy

to
Department of Theoretical and Applied Linguistics,
University of Edinburgh

December 1999

Abstract

This thesis investigates the degree to which Japanese verbal alternations can be related synchronically. In particular, it investigates the possibility of providing a polysemous analysis of the verbs which appear in simplex and conjunct forms using the conjunctive *te* form.

The data which are investigated include those verbs which participate as second conjunct in a syntactic construction which has been labelled by Hasegawa (1995) as Nuclear Conjunction. The verbs taking second conjunct position which are analysed include *miru* (see), *morau* (receive), *ageru* (give), *iru* (animate be) and *aru* (inanimate be). In the conjunct construction, the first conjunct takes the affix *te* and no arguments can intervene between the two verbs. Furthermore the Vendler-Dowty (Dowty (1979)) aspectual class of the first conjunct verb is restricted, and there is variation in the number of arguments that can be realised, depending on the properties of the second conjunct.

The analysis is developed using Head Driven Phrase Structure Grammar (HPSG) (Pollard and Sag (1994)) and Pustejovsky's (1995) Generative Lexicon theory. Polysemous analyses of the simplex and *te* form alternations of *miru*, *morau* and *ageru* are provided, based on underspecification in the syntactic *comps* and semantic *content* type hierarchies. Since current HPSG makes no provision for aspect, a type hierarchy is developed using Pustejovsky's Event Structure, under the *content* field. Variations in argument realization are shown to follow from the different modes of composing the first and second conjuncts. One particularly interesting construction in Japanese is what Matsumoto (1990) calls the *Intransitivizing Resultative* involving *V+te aru*, which as a conjunction projects the *undergoer* of the first conjunct to subject, suppressing the *actor* role. The valence alternation displayed in this construction is explained by the Agentive and Formal projections of Pustejovsky's Qualia Structure.

Some degree of polysemy is shown to hold between simplex and conjunct uses of the example verbs, but there are other syntactic phenomena to be explained. The thesis also examines the *te* conjunctions as control constructions and finds that Pollard and Sag's (1994) claim that the controlling subject is overtly structure-shared with the semantic subject of the embedded predicate, is at best not proven. A further phenomenon of the conjunct combinations of *V+te iru* and *V+te aru* is that these forms combine with adverbs compatible with the aspectual class of the second (Stative) conjunct, whereas in other *te* conjunctions, adverbs are compatible with first conjunct aspect. This phenomenon can be explained only by positing argument structure at the phrasal level, with the *V+te iru/V+te aru* phrase projecting the combined argument structure of the two conjuncts, in contradiction to the Lexicalist Hypothesis first proposed by Bresnan (1978).

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Stephen Nightingale

Acknowledgements

In the three years I have been engaged in this PhD, I have come to learn quite a fair amount of linguistics. Without the consistent attention of a supervisor dedicated to my education and to the emerging needs of this particular project, the final outcome would not nearly have been so rich. I am indebted to Ronnie Cann, my first supervisor, for a very great deal, and I thank him for his care and patience and for the faith that he has shown. I was also privileged to have the support of Caroline Heycock as second supervisor. Her knowledge of Japanese helped me to avoid many, many errors. If the data in this thesis is not "right", then the fault is mine.

Fees and living expenses are always a worry for postgraduates. I was lucky to have the burden on my shrinking pocket reduced to the extent of a University of Edinburgh Faculty of Arts scholarship, and an Overseas Research Students' (ORS) award, which together paid for one year of the university fees.

At some point during the research I was worried that I may not be able to get enough examples of Japanese, due to a shortage of available native speakers. Looking back it seems that such a worry was quite unfounded because over the past 3 years I have had the pleasure of talking with quite a significant number of people who were extraordinarily helpful in not only providing examples but also much enjoyable discussion about all things Japanese. Particular thanks must go to: Ai Hirose, Masumi Kohama, Tomoko Shimabukuro, Yoko Shomura, Madoka Aiki, Yuka Kadoi, Yoko Matsumoto Sturt, Etsuko Oishi, Yuka Kondo, Noriko ('Tenten') Nakahama, Hajime ('Jimmy') Yamauchi and Mika Itou. Kei Yoshimoto provided me with the test for distinguishing Stative verbs from Dynamic verbs.

Later I became worried that nobody would want to read my turgid works, but again, the help of a few brave souls was timely and useful, and I thank Caroline Sporleder,

Max Louwerse and Yoko Shomura for their comments on various chapters. I also wish to thank Okhee Kim, for friendship over the years.

Support of a different kind came over the network. Several times I importuned the HPSG Mailing Group with questions about Japanese and questions about HPSG. The responses were many, prompt and informative. I wish to thank all the members of the Mail Group for entertaining these questions. Particular thanks in this context must go to Bob Carpenter, with whom I had several useful e-mail exchanges.

Cassie Mayo did us all a good service by adapting Stephen Harker's thesis style package to Edinburgh requirements. For drawing the Feature Structure diagrams which are so ubiquitous in HPSG analyses, I used a LaTeXstyle file developed by Elias Papavassilopoulos.

My principal resort during these sometimes stressful times has been The Elephant House, which was opened by Iain Fraser just in time for my arrival in Edinburgh. It is a great place to chill out, and they do a good latté.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis investigates a set of Japanese verbs and hypothesizes that they are *polysemous* in subcategorizing alternate nominal or verbal complements. The verbs which participate in these alternations include *iru* (animate be), *aru* (inanimate be), *shimau* (put back; complete, regret), *miru* (see), *morau* (receive), *ageru* (give), *iku* (go), *kuru* (come) and *oku* (put, prepare). The adjective *hoshii* (want) also patterns with *morau* and carries the same meaning as the desiderative form *moraitai* (want to receive). In the case where the verb takes a nominal complement, this can be marked with a dative or an accusative case marker; in the case of a verbal complement, this is always marked with the morphological affix *te*, and these verbal combinations have been called *Nuclear Conjunctions* by Hasegawa (1995). Among these constructions the *te* marked verb is the *first conjunct* and the finite verb is the *second conjunct.*

According to an analysis reported by Martin (1975), these Nuclear Conjunctions constitute just under 50% of the uses of *te* as an affix. The prototypical form of verbs which conjoin with *te* is *te iru*, exemplified below, and 3/5 of the uses of Nuclear Conjunctions are uses of *te iru.* Examples (1) and (2) illustrate the difference between the simple present and progressive forms of a transitive verb *yomu* (read), the *te* form of which is *yonde*[1].

(1.1) Keiko ga hon o yomu.
Keiko-NOM book-ACC read
"Keiko reads a book".

[1]When the stem of a verb terminates in *m*, *n*, *g* or *b* the *te* form is realized as *de*, thus: *yomu* (read) ⟶ *yonde*, *shinu* (die) ⟶ *shinde*, *oyogu* (swim) ⟶ *oyoide*, *asobu* (play) ⟶ *asonde*, but: *akiru* (be tired of) ⟶ *akite*, *hiku* (pull) ⟶ *hiite*, *korosu* (murder) ⟶ *koroshite* and *machigau* (err) ⟶ *machigatte.*

(1.2) Keiko ga hon o yonde iru.
Keiko-NOM book-ACC read-TE be
"Keiko is reading a book".

When the verbs under investigation are examined in their simplex and conjunct forms, some interesting patterns recur. The verbs *miru*, *morau* and *ageru* seem to take the same number of arguments whether they are simplex or conjunct forms, so ditransitive *morau* takes a nominative, a dative and a third argument which may be an accusative nominal, or a *te* marked verb phrase (VP). In the simplex form, example (3), the second argument may be marked with *ni* or *kara* (from), though *kara* is the more likely to occur, whereas in the conjunct form, example (4), this argument is almost always marked with *ni*.

(1.3) Keiko ga Naoko ni/kara hon o morau.
Keiko-NOM Naoko-DAT/FROM book-ACC receive
"Keiko receives a book from Naoko".

(1.4) Keiko ga Naoko ni kami o kitte morau.
Keiko-NOM Naoko-DAT hair-ACC cut-TE receive
"Keiko had Naoko cut her hair".

Both forms involve *Keiko* receiving something from *Naoko*, which is in (3) a book, but in (4), the service of cutting her hair. The conjunctive form of *te morau* can be glossed in a number of ways, one of which is that "Naoko does something for Keiko", and another is the way I have glossed the above example, "K has N do something". These possible translations of *te morau* tend to obscure the point about the alternate uses of *morau* in that the accusative nominal strictly alternates with the *te* marked verb, which offers a possible polysemous interpretation of *morau*. Indeed, a literal translation of the conjunct form is "Keiko receives (the service of) cutting her hair from Naoko", parallel to the translation of (3) "Keiko receives a book from Naoko".

The term 'polysemy' covers a broad range of phenomena, and a comprehensive discussion is given in Section 1.1 of this introduction. My hypothesis is that these Nuclear Conjunctions illustrate the polysemy of some of the participating verbs, on the basis of polymorphic subcategorization. A second form of polysemy is exhibited in what both Martin (1975) and Matsumoto (1990) call the "Intransitivizing Resultative" *te aru* construction in which *aru* (inanimate be) combines with a transitive or ditransitive Accomplishment, suppressing the *actor* (or agent) and projecting the *undergoer* (or theme) as subject, thus *shimeru* (close) in example (5) projects

a transitive verb frame, while *shimete aru* in (6) projects intransitive, and *aru* has selected an alternative intransitive projection of *shimeru*.

(1.5) Naoko ga mado o shimeru.
Naoko-NOM window-ACC close
"Naoko closes the window".

(1.6) mado ga shimete aru.
window-NOM close-TE be
"The window is closed".

A further challenge posed by the data is how to account for the different adjunct compatibility constraints of these constructions: while *te miru*, *te morau* and others combine with adverbial adjuncts to reflect the aspectual characteristics of the first conjunct, *te iru*, *te aru* and *te shimau* combine with adverbs compatible with the second conjunct. Thus, while *hiite morau* in (7) is a conjunct composed of atelic Activity *hiku/hiite* (play/playing) and telic Accomplishment *morau* (receive), and this combines with an atelic adverb but not a telic adverb, in contrast *hiite iru* in (8) is composed of atelic Activity *hiite* and State *iru* (be), yet only the Stative adjunct *shibaraku mae kara* (since a while ago)is compatible with the conjunctive form.

(1.7) Keiko ga Taroo ni 1-jikan/*1-jikan de piano o
Keiko-NOM Taroo-DAT for an hour/*in an hour piano-ACC
hiite morau.
play-TE receives
"Taroo plays the piano for Keiko for an hour/*in an hour".

(1.8) Taroo ga shibaraku mae kara/*moo sugu piano o hiite iru.
Taroo-NOM since a while ago/soon piano-ACC play-TE be
"Taroo is playing the piano since a while ago/*soon".

While the first case illustrated by example (7) can be directly captured under standard lexicalist assumptions, the second case is more problematic. However, an approach to the analysis of (8) is offered in this thesis which posits argument structure merger. Such an explanation poses a challenge to the Lexical Integrity Hypothesis originally propounded by Bresnan (1978), and more recently repeated in Bresnan & Mchombo (1995), which stipulates that the formation of complex lexical phenomena and syntactic structures is explained by different types of processes, and that these phenomena are disjoint.

This thesis builds on a number of basic concepts mentioned above, in particular the concepts of *polysemy* and of verbal aspect, or *aktionsarten*. A general introduction to polysemy is given in Section 1.1 of this introduction, while Vendler-Dowty aspect is explained in Section 1.2.

1.1 What is Polysemy?

In one interpretation, the term *polysemy* is used to denote the multiple related meanings of a particular word, while the obverse term *homonymy* denotes the multiple different meanings of a word. As illustration of these terms the word 'bank' in examples (9) to (11) has both homonymous and polysemous forms. Cruse (1986) notes that the phonological complex 'bank' has multiple meanings, including the sense of 'side of a river' in (9) and that of 'financial institution' in (10), and these are homonyms represented in the lexicon by separate lexemes. The contrast between (10) and (11) however is between two different senses of the same lexeme: bank in its financial institution sense — as an organization in (10) and as a specific location in (11) and these are polysemous forms of the same sense.

(1.9) We sat on the bank and fished.
(1.10) The bank made a profit of fourteen trillion yen last year.
(1.11) My bank was damaged in the earthquake.

A second form of polysemy is what Cruse (1986) calls *generality* and Kempson (1977) calls *vagueness*. The word 'neighbour' can be used to describe an adjacent location, and also a person at an adjacent location, thus neighbour in (12) denotes the location, but in (13) it denotes the person who lives next door. The sense of neighbour in (14) is also one of proximity which is similar to that of (12), but the difference between them shows that the word is being used to denote a sense of relativistic proximity, more general than the strictly local sense.

(1.12) "I'll lug the guts into the neighbour room"
(Shakespeare, Hamlet, Act III, Scene IV)
(1.13) My neighbour always plays his stereo too loud.
(1.14) At 2 million miles, the Andromeda galaxy is our nearest neighbour.

Of more direct relevance to this thesis is a form of polysemy in which verbs participate, involving alternation in verb frames, and both Levin (1993) and Pustejovsky (1995) suggest these syntactic alternations are semantically driven. Levin (1993)

identifies a wide variety of alternation classes including, for example, "transitivity alternations" and "intra VP argument alternations". The first class includes causative/inchoative alternations such as 'break', which patterns as a transitive Accomplishment in (15) and an intransitive Achievement in (16). In both cases, the meaning of 'break' involves the same change of state of the window.

(1.15) Taroo broke the window.
(1.16) The window broke.

Induced action alternations fall in this class and these include the verbs 'run' and 'jump'. In (17) 'the horse' engages in an action, 'jumping', while 'the rider' causes the action in (18).

(1.17) The horse jumped over the fence.
(1.18) The rider jumped the horse over the fence.

Among "intra VP argument alternations", the verbs 'give' and 'teach' participate in the dative alternation, exemplified for 'give' in (19) and (20), and the verbs 'spray' and 'load' exhibit the locative alternation, exemplified for 'load' in (21) and (22).

(1.19) Aiko gave the dog a bone.
(1.20) Aiko gave a bone to the dog.
(1.21) Sweyn loaded the booty onto his ship.
(1.22) Sweyn loaded his ship with booty.

In all these verb frame alternations, the verbs 'break', 'jump', 'give' and 'load' subcategorize for different arguments, but the meanings of the verbs themselves denote the same consequential actions, although there may be, for example, quantificational differences in the overall meanings of the sentences. For example with 'load' in (21) and (22): 'to load the ship with booty' implies that the ship becomes full, with possible booty for which there is no more room. In contrast, 'to load the booty onto the ship' implyies that all the booty is loaded, and there may be cargo space left over. Leaving aside the quantificational effect, the meaning of 'load' involves the same action.

Pustejovsky (1995) introduces a further set of verbal alternations which involve *type shifting*. These include verbs such as 'begin' and 'try', which may take a nominal, infinitive or gerundive complement, as shown in (23) (a) to (f). It is Pustejovsky's (1995) contention that the alternations of 'begin' in (a) to (c), and the alternations

of 'try' in (d) to (f) are polysemous, and he develops a lexical semantic framework to account for them.

(1.23)
(a) Sonoko began the journey.
(b) Sonoko began to go home.
(c) Sonoko began going home.
(d) Junko tried the beer.
(e) Junko tried to go home.
(f) Junko tried going home.

I propose that a similar phenomenon is in operation with the Japanese verbs which participate in Nuclear Conjunction such as *morau* (receive) with which the accusative argument (*hon o*) in example (3) alternates as complement with the *te* marked VP (*yonde*) in example (4). The analysis of these phenomena using Pustejovsky's machinery is taken up in Chapter 4.

1.2 Dowty's Aspectual Classification

A noticeable feature of the various verbs identified as second conjunct in Chapter 2 is that the *te* form alternations are selected based on their aspectual class. It is appropriate then, prior to the presentation of the data, to introduce the Vendler-Dowty (Dowty (1979)) aspectual classification system and the diagnostic tests to determine aspect.

There are a number of primitive features underlying an aspectual classification scheme for verbs, and these features are evident in the relationships between certain types of adjuncts and verbs. These primitive features include *stativity*, *telicity* and *agentivity*, as discussed in the works of Verkuyl (1993), Molla Aliod (1997) and others. A fourth feature, *durativity*, is identifiable in Kindaichi's (1950) system, though it is not used in the primary classification of verbs. A verb is *telic* if it exhibits a change of state, and this is tested by compatibility with adverbs like 'suddenly' or 'in an hour/minute/second'. The verbs 'die' and 'write' in examples (24) and (25) are *telic*, while 'ride' in (26), which is compatible with 'for 6 hours' but not with 'in 6 hours', is *atelic*.

(1.24) Diana died suddenly.
(1.25) Karl Marx wrote Das Kapital in 8 years/*for 8 years.
(1.26) Tom Simpson rode his bike for 6 hours/*in 6 hours.

It is, however, possible to give 'write' an atelic reading by emphasizing repeated events, like 'write books' in (27), or give 'ride' a telic reading by adding a goal 'the Tour de France', in (28), and in fact Verkuyl (1993) developed the idea that

sentential aspect can differ from the aspect of the head verb if additional goals are added, or if quantification is taken into account.

(1.27) Isaac Asimov wrote books for 50 years/*in 50 years.
(1.28) Lance Armstrong rode the Tour de France in 3 weeks/*for 3 weeks.

Verbs which are compatible with adverbs like 'deliberately' or 'carefully' exhibit the feature of *agentivity*, and these include verbs like 'write' and 'ride' above, but not 'died', as shown in examples (29) to (31).

(1.29) Karl Marx deliberately wrote about the oppression of the workers.
(1.30) Lance carefully rode up the Galibier.
(1.31) *Beryl deliberately died last year.

States are incompatible with the progressive form, so its use constitutes a test for Stativity; while verbs like 'know' (32), or 'want' (33) fail this test, 'reading', 'writing' and 'exploring' are compatible with the progressive, as shown in examples (34) to (36).

(1.32) *Yasuko is knowing *The Tale of Genji.*
(1.33) *Stephen is wanting to get a real job.
(1.34) Adam is writing about international trade.
(1.35) Harpo is riding Sun Up in the 2:30.
(1.36) Antarctic explorers are dying of loneliness.

Vendler (1957) (in Vendler (1967)) noticed that verbs could be classified according to some of these properties and he devised an aspectual classification system. He found that 'verbs that admit continuous tenses' include *activity terms* such as "pushing a cart" and *accomplishment terms* such as "running a mile", and 'verbs lacking continuous tenses' include *achievement terms* such as "reaching the summit" and *state terms* such as "loving" or "believing". Vendler's taxonomy of aspectual classes is given in Table 1.1, with examples of English verbs for each class.

Dowty (1979) further developed Vendler's classification and he devised a series of tests for verbal and adverbial phenomena which test for the primitive features, and thereby developed a systematic way to classify States, Activities, Achievements and Accomplishments. In his further explanation, Dowty (1979) is also building on ideas from Generative Semantics proposed by Lakoff (1965), with the intention

State	Activity	Achievement	Accomplishment
have	walk	receive	give
know	eat	learn	teach
want	watch	die	kill
be dead	talk	come	put

Table 1.1: Vendler's Verb Classes (From Dowty (1979))

of developing a model-theoretic analysis based on Montague Grammar (Montague (1974)).

Lakoff had made the distinction between stative and non-stative verbs (i.e. between statives on the one hand, and Achievements, Activities and Accomplishments on the other); for example only non-statives take the progressive, Example (37) versus Example (38), and only non-statives occur as complements of *force* or *persuade*, illustrated by Example (39) versus (40).

(1.37) John is building a house.
(1.38) *John is knowing the answer.
(1.39) John forced Harry to run.
(1.40) *John forced Harry to know the answer.

In order to arrive at a rigorous taxonomy we need unambiguous ways to distinguish each verb class, and Dowty explores further tests to distinguish between Activities and Accomplishments, and to identify Achievements and distinguish them from Accomplishments. Time adverbials can be used to distinguish between Activities and Accomplishments: where Accomplishments can be completed "in an hour/week/year", Activities can only take place "for an hour/week/year". Thus, the Accomplishment of "painting a picture" is illustrated in Examples (41), versus (42) and the Activity of walking is shown in Example (43) versus (44)[2].

(1.41) John painted a picture in an hour.
(1.42) ??John painted a picture for an hour.
(1.43) John walked for an hour.
(1.44) *John walked in an hour.

[2]Curiously enough, though, when painting is applied to 'wall' both telic and atelic readings are fully acceptable.

It should be noted that in English, and in other languages such as Japanese, many verbs with an Activity reading can be turned into Accomplishments, with the addition of a "theme" argument, or a "goal" argument. Some intransitive activities can be turned into accomplishments by the addition of an object, so the acceptability of "John walked ..." in Examples (43) and (44) is exactly reversed in Examples (45) and (46).

(1.45) *John walked three miles for an hour.
(1.46) John walked three miles in an hour.

Vendler's fourth class, Achievement verbs, are distinguished from Activities in that they do not happily combine with "for a minute/hour/week", are not acceptable as complements of *finish*, and they are anomalous in conjunction with a class of adverbs identified by Ryle (1949), including *attentively, studiously, vigilantly.* The combination of tests in (47) to (50) demonstrate that 'notice' is an Achievement.

(1.47) John noticed the painting in a moment.
(1.48) ??John noticed the painting for a moment.
(1.49) *John finished noticing the painting.
(1.50) ??John vigilantly noticed the painting.

Dowty's compendium of tests for aspect discussed here is summarized in Table 1.2. While the aspectual classification described here gives a deeper insight into the behaviour of verbs, Dowty sought a model-theoretic basis for building on the simplest possible elements which are Stative predicates, together with only a few operators and connectives. The resultant formal model is discussed below.

1.3 Dowty's Aspectual Calculus

The basis of Dowty's system is that stative verbs can be directly represented as predicates, thus *love* and *know* in Examples (51) and (52) translate into two place predicates:

(1.51) John loves Mary = **love'(J,M)**.
(1.52) Strachey knows denotational semantics = **know'(S,D)**.

The other verb classes incorporate, and are expressed in terms of, stative predicates, so achievements, which involve a change of state, can be derived quite straightforwardly with the use of the operator BECOME with the following definition:

No.	Criterion	St.	Act.	Acc.	Ach.
1	Meets stative vs non-stative tests of Lakoff (1965).	no	yes	yes	?
2	Has habitual interpretation in simple present.	no	yes	yes	yes
3	ϕ for an hour, spend an hour ϕ'ing.	OK	OK	OK	Bad
4	ϕ in an hour, take an hour to ϕ.	Bad	Bad	OK	OK
5	ϕ for an hour entails ϕ at all times in the hour.	yes	yes	no	d.n.a
6	x is ϕ'ing entails x has ϕ'ed.	d.n.a.	yes	no	d.n.a.
7	Complement of *stop*	OK	OK	OK	Bad
8	Complement of *finish*	Bad	Bad	OK	Bad
9	Ambiguity with *almost*	no	no	yes	no
10	x ϕ'ed in an hour entails x was ϕ'ing during that hour.	d.n.a.	d.n.a.	yes	no
11	Occurs with *studiously, attentively, carefully,* etc.	Bad	OK	OK	Bad

Table 1.2: Syntactic and Semantic Tests for Verb Classification (Dowty (1979)).

- OK = The sentence is grammatical, semantically normal.
- bad = The sentence is ungrammatical, semantically anomalous.
- d.n.a. = The test does not apply to verbs of this class.
- (SN's Note: This table is copied verbatim from Dowty).

> Where ϕ is any predicate formula and t is any time, BECOME ϕ is true at t if ϕ is true at t and false at t - *1*.
>
> Dowty (1979:p.76)

In the sentence "John learned calculus" the implication is that prior to a certain time or event, John did not know calculus, and afterwards John knew/knows calculus. This is an Achievement encoded in Example (53).

(1.53) John learned calculus = BECOME [**know'(J,C)**].

Achievements which involve a loss on the other hand can be represented as the negation of a predicate. The verb *forget* is in some sense the opposite of *learn* and "John forgot the calculus" is represented with the BECOME NOT operators in Example (54).

(1.54) John forgot the calculus = BECOME NOT [**know'(J,C)**].

As a further build-up to the system, accomplishments can be interpreted as linking two sentences together with the CAUSE operator in the relation [ϕ CAUSE ψ], where ϕ may be an achievement or activity sentence, and ψ is an achievement. Some notion of "do something" is presumed in the initial CAUSE operation of the formal translation of Accomplishments and as illustration, Dowty notes the class of homicidal verbs including *electrocute, strangle, poison, drown, hang* as giving a specific method of bringing about death, so in the [ϕ CAUSE ψ] model, ϕ can be construed as something like "John strangled/poisoned/drowned Bill", and ψ is invariably [BECOME **dead'(Bill)**].

Activity verbs are more problematic for Dowty and he tentatively introduces the operator DO, inherited from Ross (1972) and he supposes that both stative and active verbs are constructed from the same homogeneous class of primitive stative predicates, and the presence of the operator DO distinguishes the meaning of active verbs from stative verbs. Examples are drawn from Rogers (1971), with a list of cognitive (stative) verbs paired with their active counterparts, thus:

Cognitive	**Active**
see	look at, watch
hear	listen to
feel	feel
smell	smell
taste	taste

Dowty's proposed use of *DO* actually differs from Ross' conception of the underlying semantics of Action verbs in which "every verb of action is embedded in the object complement of a two-place predicate whose phonological realization in English is **do**'". The class of Activity verbs, however, includes a set of intransitives which are non-agentive and therefore cannot be modified by the operator DO. These verbs include particular verbs of motion, such as *vibrate, rotate, hum, run* and *rumble* and *roll.* His failure to satisfactorily account for the semantics of these in the existing model led Dowty to propose an interval based semantics, with a function *move(x)* in which "L is a function which assigns a place to each individual at each moment in time" and is specified as follows:

> move(x) is true at interval i if there is a place (p) such that L(x) = p at the lower bound of i and L(x) $\neq$ p at the upper bound of i.
>
> Dowty (1979:p.169)

This is encoded using the operators already introduced as:

Verb Class	Logical Structure
State	predicate'(x,y).
Activity	do'(x,[predicate'(x,y)]).
Achievement	become[State LS].
Accomplishment	do'(x,[predicate'(x,y)]) cause become[State LS].

Table 1.3: Formal Model of Aspect Classes (Dowty 1979)

be-at'(x,y) & BECOME NOT be-at'(x,y).

Instances of *fall* and *roll* with this logical structure include Examples (55) and (56).

(1.55) The vase fell off the mantelpiece.

(1.56) The ball rolled down the ramp.

In summarizing the aspect calculus, Dowty incorporates the operator DO in spite of his reservations, and the system incorporates the following:

1. a set of individual variables,
2. a set of individual constants,
3. a set of n-place predicate constants (statives),
4. the logical symbols of first order logic plus:
5. the symbols AT, BECOME, CAUSE and DO.
6. It may be supposed to be a two sorted logic with variables $t_1 ... t_n$ and constants ranging over times, and variables $x_1 ... x_n$ and constants ranging over ordinary individuals.
7. The formation rules for first order predicate logic plus:
   - i if ϕ is a formula, then BECOME ϕ is a formula,
   - ii if ϕ and ψ are formulae then [ϕ CAUSE ψ] is a formula,
   - iii if ϕ is a formula and α is a term denoting an individual, then DO(α,ϕ) is a formula (subject to the like subject constraint, or alternatively treating DO as a predicate modifier),
   - iv if ϕ is a formula and τ is a term (variable or constant) denoting a time, then AT(τ,ϕ) is a formula.

Dowty (1979:p.122)

The verb classes and their logical structures are summarized in Table 1.3. Dowty's aspect calculus is incorporated into the early version of the Role and Reference Grammar framework (Foley & Van Valin (1984)) and used by Hasegawa (1995) in her analysis of the Japanese *te* verb constructions. An alternative system which

was developed for Japanese prior to Vendler's work is that of Kindaichi (1950), discussed next.

1.4 Kindaichi (1950)

The Vendler-Dowty aspectual system can be adapted to the classification of Japanese verbs (or indeed any other language), though a set of tests relevant to Japanese is also required. A prior classification system had already been developed by Kindaichi (1950), however, predating even that of Vendler (1957), and this is based on compatibility with *te iru*, the conjunctive form which denotes the progressive. Kindaichi argued that there were four classes: Stative I, Instantaneous, Continuative and Stative IV, distinguished as follows: Statives such as *iru* (animate be) and *aru* (inanimate be) do not occur in conjunction with *te iru*, thus **atte iru* is ungrammatical; Stative IV verbs always occur with *te iru* to denote a resultative state, as with *sobiete iru* (tower over) or *nite iru* (resemble)[3]. Instantaneous verbs yield an imperfective reading: *shinu* (die) $\longrightarrow$ *shinde iru* (is dead), *aku* (open) $\longrightarrow$ *aite iru* (is open), and Continuative verbs yield a progressive reading: *oyogu* (swim) $\longrightarrow$ *oyoide iru* (is swimming), *aruku* (walk) $\longrightarrow$ *aruite iru* (is walking). Kindaichi's system is defined purely in terms of *te iru*, so it does not independently classify first conjunct verbs with any accuracy, for example this system is unable to classify the subset of continuative verbs which combine with *te aru*, though the Vendler-Dowty system identifies these as Accomplishments. Nor can Kindaichi's system identify the fact that all but Stative I verbs take *te shimau*. Kindaichi's system, then, is not independently useful. Since this study seeks out characteristics of other Nuclear Conjunctions, which the Vendler-Dowty system provides, their system is preferred over that of Kindaichi.

1.5 Diagnostic Tests

Although there is no direct correspondence between Kindaichi's system and the Vendler-Dowty system, Jacobsen (1991) pointed out that there is a commonality. He showed that the Continuative verbs denote the progressive and therefore bear a resemblance to Activities and Accomplishments. On reclassification by Jacobsen, the Accomplishments actually straddle the Continuative and Instantaneous categories of Kindaichi. McClure (1994) develops a syntacticized view of aspect in

[3]Martin (1975) demonstrates examples where the imperfect is used: ... *kanarazu-shimo anzen sei ga sugureru to wa ienai* 'we cannot say that they are always superior in safety'.

Verb Form	**Stative I**	**Instantaneous**	**Continuative**	**Stative IV**
te iru	not possible	perfective	progressive	mandatory
Main Verb	iru (be)	sinu (die)	oyogu (swim)	sobieru (tower over)
	aru (be)	iku (go)	hashiru (run)	niru (resemble)
	iru (need)	aku (open)	aruku (walk)	-

Table 1.4: Kindaichi's Aspect Table for *te iru*

Japanese, in which he adopts Vendler-Dowty. To develop the primary aspectual classification of verbs, he adopts a set of diagnostic tests developed by Moriyama (1988). These tests include compatibility with raising forms *V-i hajimeru* (begin to V), *V-i tuzukeru* (continue to V) and *V-i owaru* (finish V-ing). Activities, Accomplishments and Achievements are all compatible with *V-i hajimeru*, while States are not, so *shinihajimeru* (begin to die), *oyogihajimeru* (begin to swim) and *tukurihajimeru* (begin to bake) are good and **arihajimeru* (begin to be) is not. Both Activities and States are compatible with *V-i tuzukeru*, so *oyogituzukeru* (continue to swim) and *arituzukeru* (continue to be), are good, **shinituzukeru* (continue to die) is bad, because *shinu* (die) entails a change of state, and only the result state can be continued, but not the change. It is problematic to apply this test to Accomplishments because a single 'baking' event is telic, but the repeated activity is atelic: *keeki o tukurituzukeru* (?continue baking a cake) is questionable, because at some stage the cake becomes baked. The diagnosis is further confused by the fact that in Japanese, nominals are underspecified for number, thus 'to continue baking one cake' (*hitotsu no keeki*) is questionable, but a plural reading of *keeki o* as cakes in general is acceptable. The compatibility of Accomplishments with *V-i owaru* (finish V-ing) is similarly ambiguous: *tukuriowaru* (finish baking) is acceptable if the Activity is the focus, but *keeki o tukuriowaru* (finish baking a cake) is odd, because the change of state defines termination[4] though underspecified *tukuriowaru* is okay. Only Activities are fully compatible with *V-i owaru*, as in *oyogiowaru* (finish swimming). Achievements are telic, and **shiniowaru* (finish dying) is just bad. For the opposite reason, States are incompatible with *V-i owaru*, and **ariowaru* (finish being) is bad because States are unbounded.

[4]The situation is further complicated if an Incremental Theme (Dowty (1979)) is assumed, because the change of state is gradual, so while you can leave a cake unfinished, there is still a change of state involved. A satisfactory solution for this aspectual problem is still awaited.

No.	Criterion	St.	Act.	Acc.	Ach.
1	*V-i hajimeru* (begin to V)	no	yes	yes	yes
2	*V-i tuzukeru* (continue to V)	yes	yes	yes	no
3	*V-i owaru* (Finish V'ing)	no	yes	ambig	no
4	*N-jikan kakatte V* (take N hours to V)	no	no	yes	yes
5	*N-jikan V* (V for N hours)	yes	yes	ambig	no
6	*V-te iru* compatibility	not poss	prog	prog	perf
7	*N-jikan de V* (V in N hours)	no	no	yes	yes
8	*guuzen ni V* (accidentally V)	no	no	no	yes
9	*shibaraku mae kara* (since a while ago)	yes	no	no	no
10	*moo sugu* (soon)	no	yes	yes	yes

Table 1.5: Syntactic and Semantic Tests for Japanese Verb Classification

The test for telicity in Moriyama's scheme is compatibility with *sanjikan kakatte V* (take 3 hours to V) contrasting with *sanjikan V* (V for 3 hours), and while Achievements and Accomplishments are compatible with the former, Activities and States are compatible with the latter. This test is replaced in Hasegawa (1995) with a direct translation of Vendler's 'in an hour'/'for an hour' tests, thus *3-jikan de V* (V in 3 hours) goes with telic verbs. Kindaichi's observation that *te iru* is perfect in combination with Achievements is also added as a test by Moriyama, so *shinde iru* (is dead) is perfect. In conjunction with Activities and Accomplishments *oyoide iru* (is swimming), *keeki o tukutte iru* (is baking a cake), *te iru* is progressive, and it is incompatible with States **atte iru* (is being??).

What is missing from this compendium of tests is a test for agentivity, to distinguish Activities and Accomplishments which are agentive, from Achievements and States, which are not. Hasegawa (1995) notes that only Achievements are compatible with adverbs *guuzen ni* (accidentally) and *ukkari to* (absent-mindedly), so their incompatibility with Activity and Accomplishment verbs is a test for agentivity. More positively, the adverbs *shinchyou ni* (deliberately) and *otonashiku* (quietly) require an agentive subject. A test for the difference between Stative and non-stative (or *dynamic*) verbs is also required: States are compatible with the adverbial *shibaraku mae kara* (since a while ago) and dynamic verbs are compatible with *moo sugu* (soon)[5] The entire range of tests are summarized in Table 1.5 and this scheme is used in identifying the constraints on the conjunct forms of Japanese verbs.

[5]This test furnished by Kei Yoshimoto (personal communication).

1.6 Organization of the Thesis

The structure of this thesis is as follows: the data of Nuclear Conjunctions and their component verbs are introduced in Chapter 2, following the introductory concepts discussed below. First conjunct forms are discussed generally in terms of Vendler-Dowty (Dowty (1979)) aspect classes, and the second conjunct forms of *miru* (see), *morau* (receive), *ageru* (give), *iru* (animate be), *aru* (inanimate be), *shimau* (put back; complete or regret) and *oku* (put, prepare) are examined for their simplex and conjunctive behaviours. An investigation of polysemy implies a lexical analysis of the data, and is essentially an enterprise in lexical semantics. A formal framework which is used to explain both syntactic and semantic phenomena is the lexicalist framework of current HPSG, and the essential concepts relevant to the analysis of these Japanese verbal phenomena are explained in Chapter 3. One limitation of HPSG, however, is that it makes no provision for the representation of aspect. While Dowty's (1979) framework covers the concepts, it is not particularly adapted to a computational representation. In a branch of the aspectual tradition going through Moens & Steedman (1988), Pustejovsky (1995) addresses questions of polysemy in nominals and verbals using an Event Structure framework which builds on the aspectual classification system and which is compatible with a computational framework using typed feature structures. As part of the analysis of polysemy in verbal alternations, Pustejovsky's framework is introduced in Chapter 4, applied to the Japanese data, and an analysis is developed with the second conjunct verbs *miru* (see), *morau* (receive) and *ageru* (give) as polysemous. These verbs are found to be *control* constructions in their conjunctive forms. Various control analyses of English verbs are reviewed, particularly those of Pustejovsky (1995) and Chierchia (1984), and discussed with respect to the Japanese verbs. The verbs of existence *iru* (animate be) and *aru* (inanimate be) have simplex and conjunct behaviours which lead to more complex lexical analyses, for which the architecture of HPSG does not permit full generalization of the polysemous possibilities. These verbs and their conjunctions are discussed in the context of *sense enumerativity* (defined by Pustejovsky), and the analyses developed in Chapter 5. The verbs *iru* and *aru* in their conjunctive forms exhibit a pattern of behaviour consistent with their having phrasal level argument structure, and I argue for this analysis in Chapter 6. The thesis is concluded in Chapter 7.

Chapter 2

Simplex and Conjunct Verb Forms

The range of use of the connective *te* covers broad ground, and it participates in verbal conjunctive phenomena characterized by Hasegawa (1995) as *Clausal, Core and Nuclear Conjunction.* These phenomena always involve the first conjunct with its morphologically bound *te* marker being subordinated to the second conjunct verb, which bears a finite inflection and is the head of the clause. The differences between the three conjunctive forms involve the presence or otherwise between the two conjuncts of syntactically expressed arguments and adjuncts of the head verb. In the introduction I showed that the Nuclear Conjunctive form involves the conjunction between two verbs, in which no arguments or adjuncts may appear between the two conjuncts[1]. Core Conjunction involves the conjunction between two verbs which share a subject, as shown in example (1) where the subject 'Joan' is the victim of an accident, 'breaking her leg' of which the consequence is that she was taken to hospital. In this instance the meaning of the *te* linkage is interpreted as a *cause.* Other possible interpretations for verbs in Core Conjunction include *manner*, *means*, *concession* and *relation* interpretations.

(2.1) Jyoan ga ashi o otte byooin ni hakobareta koto.
Joan-NOM leg-ACC break-TE hospital-DAT be carried to fact
"The fact that Joan broke her leg and was carried to hospital".

[1]The only thing which can appear between the two conjuncts of the Nuclear Conjunctive form is one of a range of focus particles, such as *wa* (topic marker) *mo* (focus marker) or *sae* ('even'). An example with intervening topic marker *wa* is given below.

Keiko ga hon o yonde wa iru.
Keiko-NOM book-ACC read-TE TOP be
"What Keiko is reading is a book".

This thesis will not be concerned with the analysis of these particles.

Verb	Class	Simplex Use	Conjunct Use	Valence
ageru	Accomp	give	give (Speaker as Agent)	V-I
aru	State	be (inanimate)	Resultative; "Past Perfect"	V-R; V-M
hoshii	State	want	want (someone to do)	V-I
iru	State	be (animate)	Progressive, Perfective	V-M
iku	Accomp	go	go (physical/ metaphorical)	V-I; V-M
kuru	Accomp	come	come (physical/ metaphorical)	V-I; V-M
kureru	Accomp	give	give (3rd party as Agent)	V-I
miru	Act	see	sample; see if	V-M
miseru	Accomp	show	show	V-M
morau	Accomp	receive	receive	V-I
oku	Accomp	put	Prepare for	V-M
shimau	Accomp	put,store	Regret, Completion	V-M

Table 2.1: Verbs which License *te* Arguments.
Key:

- V-M: Valence Maintaining.
- V-I: Valence Increasing.
- V-R: Valence Reducing.

Verbs which combine in Clausal Conjunction similarly involve a range of interpretations, including *cause* and *contrast* relations. Under this form, though, the subjects of both predicates are present, and the particle *te* acts as a coordinator, similar to English 'and'. Example (2) shows a situation which is interpreted by Hasegawa (1995) as a *causal* relation: it was Joan's buying the car, which caused Bill to pay.

(2.2) Jyoan ga kuruma o katte Biru ga okane o haratta.
Joan-NOM car-ACC buy-TE Bill-NOM money-ACC pay-PAST
"Joan bought a car and Bill paid for it".

Hasegawa (1995) gives a considerable discussion of all three conjunctive forms, however my focus is on the relationship between the simplex and conjunct forms of the verbs listed in Table 2.1, whose conjunct forms are characterized as Nuclear Conjunction[2]. There will therefore be no further discussion of Core and Clausal conjunctive forms in this thesis. Along with each verb, the table identifies its Vendler-Dowty

[2]There are also associated with some of these verbs honorific forms, so for example *ageru* has an honorific form *sashiageru*, and *iru* has humble form *oru* and exalted form *irassharu*.

(Dowty (1979)) aspectual classification, simplex and conjunct uses, and the characterization of the valence potential of the conjunct form using Hasegawa's (1995) typology, which may be *Valence Maintaining* or *Valence Changing*[3] By Valence-Maintaining she means that the addition of the second conjunct verb has no surface effect on the argument expressing properties of the first conjunct. For example *shinu* (die) is an unaccusative verb with a single *ga* marked argument as subject. The *te iru* conjunctive form, *shinde iru* (be dead) takes the same *ga* marked argument. A Valence-Changing conjunction results in the addition of an argument (Valence-Increase) or the deletion of an argument (Valence-Reduction), differing from the patterns shown when the first conjunct verb is expressed as simplex. These various valence effects are illustrated in examples (3) to (6) below. The verb *shimeru* is a transitive verb meaning 'close', taking nominative (*ga*) and accusative (*o*) marked arguments as shown in (3). As a Valence Maintaining form, *shimete iru* in (4) exhibits the same argument pattern, but has a progressive reading. Uniquely among the Nuclear Conjunctive forms, *te aru* is Valence Reducing, and the subject of *shimeru* is suppressed and its object is projected as the *ga* marked argument of the conjunction, shown in (5). Valence Increasers include *te morau*, the benefactive form, in which an extra subject is added, and the argument (*Natsuko*) which is the *ga* marked subject of *shimeru* in simplex form, is the *ni* marked dative of conjunct *shimete morau*, in example (6).

(2.3) Natsuko ga doa o shimeru.
Natsuko-NOM door-ACC close
"Natsuko closes the door".

(2.4) Natsuko ga doa o shimete iru.
Natsuko-NOM door-ACC close-TE be
"Natsuko is closing the door".

(2.5) doa ga shimete aru.
door-NOM close-TE be
"The door is/has closed".

(2.6) Fusako ga Natsuko ni doa o shimete morau.
Fusako-NOM Natsuko-DAT door-ACC close-TE receive
"Natsuko closes the door for Fusako".

The table summarizes simplex and conjunct forms of a verb, and the entry for *oku*, for example, is to be interpreted as follows: *oku* is an Accomplishment which as

[3]Although Hasegawa (1995) identifies *Valence Changing*, I have further articulated this into *Valence Reducing* (V-R) and *Valence Increasing* (V-I) forms, as this study aims to develop a lexical analysis of these verbs.

a simplex verb means 'put something somewhere', and is therefore ditransitive, as shown in (7). As a Nuclear Conjunction it means 'do something in preparation', and in this form it is a Valence-Maintainer, so the conjunct form *shimete oku* reflects the valence pattern of simplex *shimeru*, with *ga* and *o* marked arguments, in (8).

(2.7) haha ga hondana ni kabin o oita.
mother-NOM bookshelf-DAT vase-ACC put-PAST
"My mother put the vase on the bookshelf".

(2.8) Natsuko ga doa o shimete oita.
Natsuko-NOM vase-ACC close-TE put-PAST
"Natsuko closed the door (in preparation)".

In this chapter, both simplex and conjunct forms of the Nuclear Conjunctive verbs are examined, with respect to their aspectual class, their valence potential and their meaning; if it can be concluded that there is a similarity between the simplex and conjunct forms with respect to these characteristics, then a polysemous analysis can be entertained. If, on the other hand, there are significant differences, then the separate forms are homonymous and must be separately listed in the lexicon.

2.0.1 Aspect

The simplex forms of the verbs under investigation have *aspectual class* or *aktionsarten* of States (*iru*, *aru*), Activities (*miru*), and Accomplishments (*morau*, *ageru*, *oku*, *shimau*, *iku* and *kuru*). In classifying the conjunct forms of these verbs there are two patterns which emerge, reflecting different adjunct scopes. The verbs *miru*, *morau*, *ageru* and *oku*, in conjunct form, take adverbial adjuncts which are compatible with the aspectual class of the first conjunct verb, not the second. This is illustrated for *miru* and *morau* in examples (9) and (10). The adverbial *1-kagetsu de* is a telic adjunct which means 'in a month' and patterns with telic verbs such as Achievements and Accomplishments, but not with Activities. Since *tateru* is an Accomplishment and *miru* is an Activity, *1-kagetsu de* is compatible with the first, not the second conjunct. Similarly, *30-ppunkan* is an atelic adverbial meaning 'for 30 minutes'. It is compatible with Activity *odoru* (dance), the first conjunct, and not with an Accomplishment.

(2.9) Saburo-san ga 1-kagetsu de ie o tatete mita.
Saburo-NOM in 1 month-ADV house-ACC build-TE see
"Saburo saw to building his house in a month".

(2.10) John ga Mary ni 30-ppunkan odotte moratta.
John-NOM Mary-DAT for 30 minutes-ADV dance-TE receive-PAST
"Mary danced for John for 30 minutes".

The verbs *iru*, *aru* and *shimau*, in conjunct form, take adverbial adjuncts which are compatible with the second conjunct, as illustrated for *iru* in example (11): *shibaraku mae kara* means 'since a while ago' and is compatible only with States. Since *shinu* (die) is an Achievement, the *te iru* form indicates second conjunct adjunct compatibility.

(2.11) yumeina hito ga shibaraku mae kara takusan shinde iru.
famous people-NOM since a while ago many die-TE be
"Many famous people have died since a while ago".

The verbs *te iku* and *te kuru* in conjunct form exhibit both of these adjunct compatibility phenomena. Example (12) shows first conjunct adjunct compatibility, since *guuzen ni* (accidentally) is *non-agentive* and *ochiru* (fall) is an Achievement.

(2.12) neko ga guuzen ni ido no naka ni ochite itta.
cat-NOM accidentally well-GEN inside-DAT fall-TE went
"The cat accidentally went falling down the well".

Example (13) shows second conjunct adjunct compatibility, since *5-funkan de* (in 5 minutes) is telic, yet *motsu* (hold) is an Activity.

(2.13) Naoko-san ga 5-funkan de gakkoo ni hon o motte itta.
Naoko-NOM in 5 minutes school-DAT book-ACC hold-TE go-PAST
"Naoko took the book to school in 5 minutes".

This split in adjunct compatibility patterns suggests a typological partition, so in the balance of this chapter, the data is organized into verb forms which reflect the aspectual classification of the first conjunct, in Section 2.1, and verb forms which reflect the aspectual classification of the second conjunct, in Section 2.2. The remaining verbs, which can reflect the aspect of either conjunction, are described separately, with Valence Maintaining/Valence Reducing *te aru* in Section 2.3, and Valence Maintaining/Valence Increasing *te iku* and *te kuru* in Section 2.4.

2.0.2 Valence

The valence patterns of simplex verbs fall into predictable patterns across aspectual classes. States have two arguments, with (*ga, ga*) or (*ga, ni*) case marking; Achievements which take one *ga* marked argument are sometimes called unaccusatives (see Perlmutter (1978)), and there are other Achievements which take two arguments, (*ga, ni*) marked. The definition of transitivity given by Jacobsen (1991) requires that an Accusative *o* marked argument be present, and only Activities and Accomplishments can be transitive so even though a State or an Achievement may have two arguments, it is still intransitive, as neither argument is accusative. Transitive Activities take (*ga, o*) marked arguments and intransitive Activities, sometimes called unergatives, take a single *ga* marked argument. There are, exceptionally, some ditransitive Activities such as *oshieru* (teach) with (*ga, ni, o*) argument marking. Accomplishments show three distinct case marking patterns: *ga, ni* for verbs such as *noboru* (climb) and *kaeru* (go home), (*ga, o*) in transitives such as *tukuru* (bake), and ditransitives, with (*ga, ni, o*) marking such as *oku* (put) and *nokosu* (leave). In the description of the conjunct data in this section, the primary classification is by aspect, though examples with any of these different verb frames may appear, and the tables summarizing conjunct effects for each second conjunct verb take account of the various verb frames. Table 2.2, here, summarizes the verb frames for each aspect class, with example verbs and their valences. I have also identified the cases in this table: *ga* is nominative (nom), *o* is accusative (acc) and *ni* is dative (dat). All verbs with one argument assign nominative case (unaccusatives and unergatives), transitive Activities and Accomplishments assign accusative case, and dative case may appear in any aspect class.

2.1 First Conjunct Aspect

The common characteristic of the verbs presented in this section is that in their conjunct forms, adjuncts are compatible with the aspectual class of the first conjunct. The verbs *miru*, *morau*, *ageru* and *oku* are presented in successive subsections, starting from 2.1.1 with *miru*; the section concludes with a comparative discussion in 2.1.5.

2.1.1 miru

In its simplex form, *miru* is a transitive activity commonly glossed as 'to see' or 'to look at'. In fact it covers much the same ground as English 'see' in that it

Verb	Gloss	Example	Example Gloss.
States			
aru(Nom,Nom)	be	*koko ga okane ga aru*	It is here that there is money.
aru(Nom,Dat)	be	*okane ga koko ni aru*	There is money here.
Achievements			
kieru(Nom)	extinguish	*hi ga kieta*	The fire is out.
kakureru(Nom,Dat)	hide	*tsuki ga kumo ni kakurete iru*	The moon is hidden by clouds.
Activities			
odoru(Nom)	dance	*Mai-san ga odotte iru*	Mai is dancing.
motsu(Nom,Acc)	hold	*boku ga hon o motte iru*	I am holding a book.
oshieru(Nom,Acc,Dat)	teach	*sensei ga kodomo ni gengogaku o oshieru*	The teacher teaches the kids linguistics.
Accomplishments			
noboru(Nom,Dat)	climb	*boku ga Fujisan ni nobotta*	I climbed Mt Fuji.
tukuru(Nom,Acc)	make	*boku ga keeki o tukutta*	I baked a cake.
oku(Nom,Acc,Dat)	put	*jyouchyuu ga hondana ni kabin o oku*	The maid puts the vase on the bookshelf.

Table 2.2: Verb Frames

can be used to express weakly agentive, or non-agentive activities such as 'seeing a mountain', in the sense of the mountain coming into view, in (14), and strongly agentive activities such as 'watching television' in (15), given in its progressive, *te iru* form. Indeed, the range of possible glosses recorded in Hiroo (1993) for *miru* is 'to see', 'look at', 'watch', 'try' or 'look after'.

(2.14) kinoo, boku ga Fujisan o mita.
yesterday I-NOM Fujisan-ACC saw
"Yesterday, I saw Mount Fuji".

(2.15) Naoko-san ga terebi o mite iru.
Naoko-NOM television-ACC see-TE be
"Naoko is watching television".

As a demonstration that it is an Activity rather than an Achievement or Accomplishment, compatibility with *10-ppunkan* (for 10 minutes), but not with *10-ppunkan de* (in 10 minutes) shows that *miru* is atelic, and the inchoative test showing

compatibility with *Vi-hajimeru* (begin to V), demonstrates that it is an Activity rather than a State, as States are incompatible with this test.

(2.16) Mai-san ga 10-ppun/*10-ppun de shinbun o mita.
Mai-NOM for/*in 10 minutes newspaper-ACC looked at-PAST
"Mai-san looked at the newspaper for(/*in) 10 minutes".

(2.17) Mai-san ga terebi o mihajimeta.
Mai-NOM television-ACC see-begin-PAST
"Mai began watching television".

In its conjunct form, *te miru* also patterns with the English verb 'see', in that it can be used in the senses of 'see to doing something' or 'see about doing something'. Although it is often glossed as 'try', Martin (1975) claims that it is more correctly translated as 'see to doing' or 'sample something and see'. Martin calls *te miru* an exploratory form, since it is used to check one's reaction to something. The use of 'try' is not accurate in the sense that it does not necessarily denote completion, i.e. you can 'try to ride a bike' or 'try to build a house', and fail to complete the act. However, *te miru* implicates that the event is completed. The contrast is most marked in considering the similarities and differences between present and past forms of 'try' and *te miru* respectively. In both cases, "I will try riding the bike" and "boku ga jitensya ni notte miru" indicate a willingness to experiment. The outcomes differ for the past tense forms though. If I say "I tried riding the bike", it is usually accompanied by an indication of failure. Conversely, if I say "boku ga jitensya ni notte mita", the indication is that I have had a little bike-riding experiment which was successful. In this case, therefore, translating *te mita* with "try doing" misses the proper semantic interpretation. An example is given of *te miru* in conjunction with *ie o tateru* (build a house) in (18).

(2.18) Saburo-san ga ie o tatete mita.
Saburo-NOM house-ACC build-TE see to-PAST
"Saburo-san saw to building a house".

In conjunction, *te miru* combines with an agentive verb (Activity or Accomplishment), while a State or Achievement as first conjunct is unacceptable. Conjunction with an Accomplishment, *tatete miru* in (18) above, and with an Activity, *odotte miru* in (19), below, demonstrate compatibility with agentives, while example (20), showing conjunction with an Achievement *odoroku* (be surprised) is unacceptable.

The use of *miru* presupposes that the initiator has some control over the event denoted by the first conjunct verb, but *odoroite* (being surprised) is not inherently a controllable event, so one cannot 'try' it or 'sample' it.

(2.19) Boku ga odotte mita.
I-NOM dance-TE saw to-PAST
"I tried dancing (to see what it was like)".

(2.20) *hannin ga odoroite mita.
criminal-NOM surprise-TE see to-PAST
"*The criminal tried to be surprised".

Seq	Verb	Aspect	Args	Output	Args	Effect
a.	*irimasu*	State	ga,ga	none	none	N/A
c.	*kikoeru*	State	ga,ga	"	none	N/A
d.	*umareru*	Achiev	ga	none	none	N/A
d.	*odoroku*	Achiev	ga	"	none	N/A
e.	*tsuku*	Achiev	ga,ni	"	none	N/A
f.	*odoru*	Act	ga	Act	ga	V-M
g.	*miru*	Act	ga,o	"	ga,o	V-M
h.	*iku*	Accomp	ga,ni	Accomp	ga,ni	V-M
i.	*akeru*	Accomp	ga,o	" "	ga,o	V-M
i.	*tateru*	Accomp	ga,o	" "	ga,o	V-M
j.	*oku*	Accomp	ga,o,ni	" "	ga,o,ni	V-M

Table 2.3: Verb Frames and Aspectual Readings for *te miru*

The conjunctive pattern of *te miru* is summarized in Table 2.3. While States such as *iru* (be) and *kikoeru* (be audible) and Achievements *umareru* (be born), *odoroku* (be surprised) and *tsuku* (arrive) lead to unacceptable combinations, Activities *odoru* (dance) and *miru* (see), and Accomplishments *iku* (go),*tateru* (build),*akeru* (get up) and *oku* (put) lead to combinations which reflect the valence pattern and the aspect class of the first conjunct, as indicated in the 'Output' column. This is exemplified for unergative *odoru* (dance) in (21) and (22), and for transitive *tateru* (build) in (23) and (24). Both *odoru* in (21) and *odotte miru* in (22) project a single *ga* marked subject, and pattern with the atelic *1-jikan* (for an hour), while *tateru* in (23) and *tatete miru* in (24) project a *ga* marked subject and an *o* marked object, and pattern with the telic *1-kagetsu de* (in a month).

(2.21) Boku ga 1-jikan/*1-jikan de odotta.
I-NOM for an hour/*in an hour dance-PAST
"I danced for an hour (*in an hour)".

(2.22) Boku ga 1-jikan/*1-jikan de odotte mita.
I-NOM for an hour/*in an hour dance-TE see-PAST
"I tried dancing for an hour (*in an hour)".

(2.23) Saburo-san ga 1-kagetsu de/*1-kagetsu ie o tateta.
Saburo-NOM in a month/*for a month house-ACC build-PAST
"Saburo-san built his house in a month/*for a month".

(2.24) Saburo-san ga 1-kagetsu de/*1-kagetsu ie o tatete mita.
Saburo-NOM in a month/*for a month house-ACC build-TE see-PAST
"Saburo tried building his house in a month/*for a month".

The problem with a single test for the aspectual class of a conjunct is that the scope of the test is ambiguous. In example (24) *tateru* is an Accomplishment, while *miru* is an Activity, and it is possible that the two verbs retain separate adjunct scoping domains. A possible way to determine this is to see if the adjunct can appear immediately between the two conjuncts, as in example (25). However, except for certain focus particles such as *wa* and *mo*, nothing can intervene between the conjuncts of a *te miru* construction, so there is no decisive test of scoping.

(2.25) *Saburo-san ga ie o tatete 3-kagetsu mita.
Saburo-NOM house-ACC build-TE for 3 months see-PAST

It is, therefore, an inference that the adjunct is associated with the first conjunct, rather than with the conjunctive pair.

2.1.2 morau

As a benefactive verb, the broad use of *morau* in its simplex form is to signify that person A receives something from person B, but like *miru* (see), it covers a range of agentivity from a weakly agentive 'receive' to a more strongly agentive 'take': indeed, it is listed in Hasegawa (1995) separately as an Achievement glossed as 'receive', and as an Accomplishment glossed as 'obtain'. Hiroo (1993) lists the glosses for *morau* as 'to get', 'receive' or 'to marry'[4]. Example (26) illustrates the basic 'receive' usage, with the *ga* marked *Kukiko* as recipient and the *ni* marked *Keiko* as donor.

(2.26) Kukiko ga Keiko ni keeki o moratta.
Kukiko-NOM Keiko-DAT cake-ACC receive-PAST
"Kukiko got a cake from Keiko".

[4]This sense of *morau* as 'to marry' is parallel with the archaic English usage of 'to take a wife', which seems to require that only one of the parties be agentive.

Both Hiroo (1993) and Martin (1975) note that in simplex form, the donor of *morau* may be marked by *ni* dative or by postpositional *kara* (from), as shown in (27). Martin claims that the oblique argument in this simplex use is more likely to be marked by *kara*, and in the conjunct use it is more likely to be *ni*.

(2.27) boku ga tomodachi ni/kara tegami o moraimashita.
I-NOM friend-DAT/from letter-ACC receive-POL-PAST
"I received a letter from a friend".

The conjunct form of *morau* is Valence-Increasing in the terms of Hasegawa (1995). While a transitive verb such as *tateru* (build) has a *ga* marked subject and an *o* marked object in (28), conjunction with *morau* adds a subject and converts the embedded subject to a dative marked argument, in (29).

(2.28) Saburo ga ie o tateta.
Saburo-NOM house-ACC build-PAST
"Saburo built a house".

(2.29) Yasuko ga Saburo ni ie o tatete moratta.
Yasuko-NOM Saburo-DAT house-ACC build-TE receive-PAST
"Saburo built a house for Yasuko".

(2.30) Yasuko ga Saburo ni ie o moratta.
Yasuko-NOM Saburo-DAT house-ACC receive-PAST
"Yasuko received a house from Saburo".

There is an alternative interpretation available, opposed to Hasegawa's Valence-Increasing categorization, however. The difference between the simplex and conjunct forms of *morau* is that in the simplex case there is an accusative object *ie o* (a house), but in the conjunct case, in the same position there is a VP *ie o tatete* (building a house). Hasegawa's categorization is based on the valence alternation between the simplex form of the first conjunct verb, and the valence of the conjunction. If the valence alternation is expressed with respect to the second conjunct, then both *te miru* and *te morau* are Valence-Maintaining, and the direct object alternates between a *te* marked VP in (29) and an accusative NP in (30). For the sake of consistency in the description of data in this chapter, however, Hasegawa's typology with respect to valence is respected. The aspectual range of verbs that can appear as first conjunct to *te morau* is less restricted than the range appearing with *te miru*, and examples of States, Achievements, Activities and Accomplishments have all been found and these are recorded in examples (31) - (34) respectively. The example of a Stative first conjunct in (31) reported by Ishikawa (1985) is,

however, pragmatically rather odd, given that Japanese citizenship is not so easily acquired in the first place, nor is it so easy to repudiate or expunge "Japanese-ness".

(2.31) ?John ga Mary ni Nihonjin de ite moratta
John-NOM Mary-DAT Japanese COP be-TE receive-PAST
"John received from Mary the favour of (her) remaining a Japanese." (Ishikawa)

(2.32) John ga tenshi ni arawarete moratta
John-NOM angel-DAT appear-TE receive-PAST
"The angel appeared for John".

(2.33) John ga Mary ni odotte moratta
John-NOM Mary-DAT dance-TE receive-PAST
"John received from Mary the favour of her dancing (with him)."

(2.34) John ga Mary ni jibun no iken o nobete moratta
John-NOM Mary-DAT self-GEN opinion-ACC express-TE receive-PAST
"John received from Mary the favour of (her) expressing self's (John or Mary) opinion." (Ishikawa)

Tests for the overall aspect of *te morau* forms pattern with *te miru* in that the aspectual class of the first conjunct verb is reflected, so Achievement and Accomplishment first conjuncts combine with telic adjuncts (*1-jikan de* (in an hour)) as in examples (35) and (37), and Activities combine with atelic adjuncts (*1-jikan* (for an hour)), as in (36).

(2.35) boku ga 3-pun de / 3-punkan go ni kare ni kokage ni
I-NOM in 3 minutes / after 3 minutes he-LOC shade-DAT
kakurete moratta.
hide-TE receive-PAST
"He was hidden in the shade for me
in 3 minutes/after 3 minutes".

(2.36) John ga 30-pun Mary ni utatte moratta.
John-NOM Mary-DAT for 30 minutes dance-TE receive-PAST
"John received from Mary the favour of her dancing with him for 30 minutes."

(2.37) John ga 30-pun de Mary ni yonde moratta.
John-NOM in 30 minutes Mary-DAT read-TE receive-PAST
"Mary read (it) for John in 30 minutes."

The aspectual range and aspectual classification patterns for the *te morau* conjunction are summarized in Table 2.4.

Seq	Verb	Aspect	Args	Output	Args	Effect
a.	*imasu*	State	ga,ga	none	none	N/A
b.	*imasu*	State	ni,ga	"	ga,ni,de	V-I
c.	*wakaru*	State	ni,ga	"	ni,ga	V-I
d.	*arawareru*	Achiev	ga	Achiev	ga,ni	V-I
e.	*kakureru*	Achiev	ga,ni	" "	ga,ni,ni	V-I
f.	*odoru*	Act	ga	Act	ga,ni	V-I
f.	*utau*	Act	ga	"	ga,ni	V-I
g.	*homeru*	Act	ga,o	"	ga,ni,o	V-I
h.	*noboru*	Accomp	ga,ni	Accomp	ga,ni,ni	V-I
i.	*noberu*	Accomp	ga,o	" "	ga,ni,o	V-I
g.	*oku*	Accomp	ga,o,ni	" "	ga,ni,o,ni	V-I

Table 2.4: Verb Frames and Aspectual Readings for *te morau*

An interesting construction related to *te morau* is the *te hoshii* construction. Although *hoshii* is an adjective meaning 'want', shown in example (38), in conjunct form it patterns with *te morau* as a ditransitive conjunction (39). Martin (1975) notes that it has the same meaning as *te moraitai* (want to receive), in (40), and that while *te hoshii* is more commonly used by Kansai speakers, the *te moraitai* form is more acceptable to Kanto speakers.

(2.38) Kukiko ga keeki ga hoshikatta.
Kukiko-NOM cake-NOM want-PAST
"Kukiko wanted a cake".

(2.39) Oguchi-san ga Keiko ni ie o tatete hoshikatta.
Oguchi-san-NOM Keiko-DAT house-ACC build-TE want-PAST
"Oguchi-san wanted Keiko to build a house (for him)".

(2.40) Oguchi-san ga Keiko ni ie o tatete moraitakatta.
Oguchi-san-NOM Keiko-DAT house-ACC build-TE receive-DESID-PAST
"Oguchi-san wanted Keiko to build a house (for him)".

As a simplex predicate *hoshii* is a State, but in conjunct form both first conjunct compatible and second conjunct compatible adjuncts are acceptable. In examples (41) and (42), *shuuri suru* (to repair) is an Accomplishment; in (41) the conjunct *shuuri shite hoshikatta* is compatible with the Stative adjunct *shibaraku mae kara* (since a while ago), and in (42) it is compatible with telic *1-jikan de* (in an hour).

Seq	Verb	Args	Output	Args	Effect
a.	imasu	ga	N/A	none	N/A
b.	imasu	ni,ga	N/A	none	N/A
c.	kikoeru	ga,ga	State	ga,ni,ga	V-I
d.	arawareru	ga	both	ga,ni	V-I
e.	kakureru	ga,ni	"	ga,ni,ni	V-I
f.	odoru	ga	both	ga,ni	V-I
g.	homeru	ga,o	"	ga,ni,o	V-I
h.	noboru	ga,ni	both	ga,ni,ni	V-I
i.	noberu	ga,o	"	ga,ni,o	V-I
j.	oku	ga,o,ni	"	ga,ni,o,ni	V-I

Table 2.5: Verb Frames and Aspectual Readings for *te hoshii*

(2.41) boku ga shibaraku mae kara Tanaka-san ni jitensha o shuuri site
I-NOM since a while ago Tanaka-san-DAT bicycle-ACC repair-DO-TE
hoshikatta.
wanted
"I wanted Mr Tanaka to repair the bicycle since a while ago."

(2.42) boku ga 2-jikan de Tanaka-san ni jitensha o shuuri site hoshikatta
I-NOM in 2hours Tanaka-san-DAT bicycle-ACC repair-DO-TE wanted

"I wanted Mr Tanaka to repair the bicycle in 2 hours."

Example (43) shows compatibility with both adjuncts, and *shibaraku mae kara* goes with *hoshikatta* (wanted), while *2-jikan de* goes with *syuuri suru* (repair). The conjunct forms of *te hoshii* are summarized in Table 2.5, showing Valence-Increase over the first conjunct as simplex and for dynamic verbs, both aspect possibilities are reflected.

(2.43) boku ga shibaraku mae kara 2-jikan de Tanaka-san ni jitensha o
I-NOM since a while ago in 2 hours Tanaka-san-DAT bicycle-ACC
shuuri site hoshikatta
repair-DO-TE wanted

"Since a while ago I wanted Mr Tanaka to repair the bicycle in 2 hours."

Although *te hoshii* patterns closely with *te morau*, the difference between simplex and conjunct forms of *hoshii* is more marked than with the comparable forms of *morau*, noted above. The simplex form of *hoshii* has *ga*, *ga* argument marking, but the conjunct form has the *ga*, *ni*, *V-te* pattern of *te morau*, and this is not so much a Valence-Increase, as a complete change of arguments.

Seq	Verb	Aspect	Args	Output	Args	Effect
b.	*imasu*	State	ni,ga	State	ni,ga	N/A
c.	*kikoeru*	State	ga,ga	N/A	none	N/A
d.	*umareru*	Achiev	ga	N/A	none	N/A
d.	*odoroku*	Achiev	ga	"	none	N/A
e.	*tsuku*	Achiev	ga,ni	"	none	N/A
f.	*odoru*	Act	ga	Act	ga,ni	V-I
g.	*miru*	Act	ga,o	"	ga,ni,o	V-I
h.	*iku*	Accomp	ga,ni	Accomp	ga,ni,ni	V-I
i.	*tateru*	Accomp	ga,o	" "	ga,ni,o	V-I
j.	*oku*	Accomp	ga,ni,o	" "	ga,ni,ni,o	V-I

Table 2.6: Verb Frames and Aspectual Readings for *te ageru*

2.1.3 ageru

Both *ageru* and *morau* pattern alike, as they are ditransitive transfer verbs with a donor, a recipient and a 'gift', and both have simplex and conjunct forms. In the case of *ageru*, it is a verb with the basic meaning of 'give', exemplified in (44) in which the *ga* marked subject is the donor and the *ni* marked oblique is the recipient. It is an Accomplishment, as demonstrated by compatibility with telic *kyuu ni* (suddenly), in (45) and with agentive *shinchyou ni* (deliberately), in (46).

(2.44) Oguchi-san ga Keiko-san ni kisu o ageta
Mr Oguchi-NOM Keiko-DAT kiss-ACC give-PAST
"Mr Oguchi gave Keiko a kiss."

(2.45) Oguchi-san ga kyuu ni Keiko-san ni kisu o ageta
Mr Oguchi-NOM suddenly Keiko-DAT kiss-ACC give-PAST
"Mr Oguchi suddenly gave Keiko a kiss."

(2.46) Oguchi-san ga shinchyou ni Keiko-san ni kisu o ageta
Mr Oguchi-NOM deliberately Keiko-DAT kiss-ACC give-PAST
"Mr Oguchi deliberately gave Keiko a kiss."

One way in which the conjunct form of *ageru* differs from *morau*, however, is in the more restricted range of first conjunct verbs with which it combines. The Achievement conjunctions in examples (47) and (48) are incoherent, while Activities and Accomplishments in (49) and (50) respectively are acceptable. As these examples also show, adverbial adjuncts are compatible with the aspectual class of the first conjunct, with atelic *1-jikan* (for an hour) combining with Activity *odotte* (dancing) in (49) and telic *1-jikan de* (in an hour) combining with Accomplishment *tatete* in (50).

(2.47) *Taroo ga Keiko-san ni ido no naka ni ochite ageta.
Taroo-NOM Keiko-DAT well-GEN inside-DAT fall-TE give-PAST
"Taroo fell down the well for Keiko".

(2.48) *Taroo ga Keiko-san ni shinde ageta.
Taroo-NOM Keiko-DAT die-TE give-PAST
"Taroo died for Keiko".

(2.49) Oguchi-san ga Keiko ni 1-jikan/*1-jikan de odotte ageta.
Oguchi-NOM Keiko-DAT for an hour/*in an hour dance-TE give-PAST
"Mr Oguchi danced for Keiko for an hour/*in an hour".

(2.50) Oguchi-san ga Keiko ni 3-kagetsu de ie o tatete ageta.
Oguchi-NOM Keiko-DAT in 3 months house-ACC build-TE give-PAST
"Mr Oguchi built a house for Keiko in 3 months".

The pattern of admissible and inadmissible conjunctions for *te ageru* is summarized in Table 2.6.

2.1.4 oku

A ditransitive Accomplishment in its simplex form, *oku* is different from *morau* and *ageru* insofar as it has one less argument in its conjunct form, and is Valence-Maintaining, in Hasegawa's scheme. The simplex form is illustrated in example (49) with the meaning 'put something (the accusative object) somewhere (the dative argument)'. It is telic, as evidenced by compatibility with *kyuu ni* (suddenly) in (50), and agentive, as shown by compatibility with *shinchyou ni* (deliberately) in (51).

(2.51) haha ga hondana ni kabin o oita.
Mother-NOM bookshelf-DAT vase-ACC put-PAST
"My mother put the vase on the bookshelf".

(2.52) haha ga kyuu ni hondana ni kabin o oita.
Mother-NOM suddenly-ADV bookshelf-DAT vase-ACC put-PAST
"My mother suddenly put the vase on the bookshelf".

(2.53) haha ga shinchyou ni hondana ni kabin o oita.
Mother-NOM deliberately-ADV bookshelf-DAT vase-ACC put-PAST
"My mother deliberately put the vase on the bookshelf".

Also in contrast with *morau*, *ageru* and *miru*, there is no direct relationship between the meaning of the simplex form 'put' and the meaning of the conjunctive form, 'do something in preparation'. Similar to *te ageru*, conjunctions with *te oku* are acceptable only with agentive verbs, as in (54) with *neru* (sleep) an Activity, and

in (55) with *oku* (put) an Accomplishment; *te oku* is incompatible with States and Achievements, as shown in (56) and (57) respectively.

(2.54) yuube wa yoku nete oita kara
last night-TOP well sleep-TE put-PAST because
kyoo no tesuto wa sinpai nai
today-GEN test-TOP worry not
"I have slept well last night, so today's test
will be no problem". (Jacobsen (1991))

(2.55) haha ga hondana ni kabin o oite oita.
Mother-NOM suddenly-ADV bookshelf-DAT vase-ACC put-TE put-PAST
"My mother left the vase on the bookshelf".

(2.56) *kyoujyu wa nihon ni ite okimashita.
Professor-TOP Japan-DAT be-TE prepare-POL-PAST
"*The professor was in Japan in preparation".

(2.57) *shitai ga hyoumen ni ukande okimashita.
corpse-NOM surface-DAT float-TE prepare-POL-PAST
"??The corpse floated on the surface (in preparation for something)".

The adjunct is compatible with the aspect of the first conjunct, in (58).

(2.58) boku ga 10-jikan/*10-jikan de nete oita.
last night-TOP for 10 hours/*in 10 hours sleep-TE leave-PAST
"I have slept for/*in 10 hours (so I'm ready)".

Seq	Verb	Args	Output	Args	Effect
a.	imasu	ga	N/A	none	N/A
b.	imasu	ni,ga	"	none	N/A
c.	kikoeru	ga,ga	"	none	N/A
d.	shinu	ga	N/A	none	N/A
e.	kakureru	ga,ni	"	none	N/A
f.	odoru	ga	Act	ga	V-M
g.	oshieru	ga,o	"	ga,o	V-M
h.	noboru	ga,ni	Accomp	ga,ni	V-M
i.	tukuru	ga,o	" "	ga,o	V-M
j.	oku	ga,o,ni	" "	ga,o,ni	V-M

Table 2.7: Verb Frames and Aspectual Readings for *te oku*

The valence, conjunction and aspect information for *te oku* is summarized in Table 2.7. A summary discussion of the various verbs *miru*, *morau*, *hoshii*, *ageru* and *oku* is given in Subsection 2.1.5, next.

2.1.5 Summary of the Data

The common factor of the verbs that appear in this section is that when they appear as second conjunct, adverbs compatible with the aspectual class of the first conjunct are admissible. The adjective *hoshii* (want), in second conjunct form, also has compatibility with adverbial adjuncts, so the *te hoshii* construction is *underspecified* for adjunct compatibility. There is a regularity observable in the data which is common to some, but not all, of the verbals: in the cases of *miru*, *morau* and *ageru*, if the *te* marked first conjunct verb is interpreted as an embedded verb phrase, then this alternates with the accusative object of the simplex forms. The simplex/conjunct alternation in *oku* does not show the same regularity: in simplex form it is a ditransitive, and as a conjunction it is transitive, and the dative argument is not retrievable. Simplex and conjunct forms of *hoshii* are similarly anomalous, in that as a simplex it is an adjective with *ga, ga* marking, and as a conjunction it patterns with *te morau*.

A lexicalist analysis of these verbs is taken up in Chapter 4, and the regular valence pattern seen in the verbs *miru*, *morau* and *ageru* suggests a possible polysemous account in that verbs with the same meaning take verb frames in which one argument alternates between a nominal and a verbal. In contrast, *hoshii* and *oku* show more homonymous behaviour in taking different numbers of arguments between their simplex and conjunct forms.

2.2 Second Conjunct Aspect

The verbs presented in this section include *iru* the verb 'to be' and *shimau*, a verb which in its simplex form means 'to put something away'. Both these verbs have simplex and *te* marked conjunct forms, and in contrast to the previous section, and conjunctions with these verbs reflect the aspectual class of the second conjunct.

2.2.1 iru

Japanese has two verbs expressing existence[5] and these are *iru* and *aru*; both have a locative argument but they are distinguished by the fact that *iru* has an ani-

[5]The copula is a third form of locative 'be', but with highly defective verbal morphology. The example below shows that a copula marked nominal does not take a normal case marker.

- kore wa hon desu.
  this-TOP book is-POL
  "This is a book".

mate argument in addition to the locative, while *aru* has an additional inanimate argument. The simplex and conjunct forms of *iru* are described here, and *aru* is described in the next section. In common with other Stative verbs, the case marking patterns of *iru* include *ga, ga* and *ga, ni* marking, shown in examples (59) and (60) respectively. The *ni* marked argument indicates a location, but when converted to *ga* it represents what Kuno (1973) calls the *exhaustive listing* reading, so in (59) "it is the park, and only the park, where children are".

(2.59) kooen ga kodomo ga iru.
park-NOM children-NOM be
"The park is where the children are"

(2.60) kooen ni kodomo ga iru.
park-DAT children-NOM be
"The children are in the park".

In describing the constructions involving *iru*, one problem is to determine which of the two arguments is the subject. Tateishi (1994) claims that the argument which participates in *ga/ni* alternation, i.e. the locative argument, is the subject, but if the animate argument of *iru* is a definite, or proper noun, such as *Taroo* rather than an indefinite form such as *kodomo* (kids), then the locative argument is marked *ni* and does not alternate with *ga*, as shown in the contrast between examples (61) and (62).

(2.61) Taroo ga koko ni iru.
Taroo-NOM here-DAT be

(2.62) *koko ga Taroo ga iru.
here-NOM Taroo-NOM be

When the animate argument is general, rather than a designated individual, the locative argument alternates between *ni* and *ga*, thus *kooen* (park) in (63) and (64) and *Hokkaido* in (65) and (66), and is therefore the subject in Tateishi's theory.

(2.63) hito ga kooen ni takusan iru.
people-NOM park-DAT many be

(2.64) kooen ga hito ga takusan iru.
park-NOM people-NOM many be

The various forms of the copula are investigated in Nightingale (1996) and will not be further investigated here.

(2.65) tsuru ga Hokkaido ni takusan iru.
cranes-NOM Hokkaido-DAT many be

(2.66) Hokkaido ga tsuru ga takusan iru.
Hokkaido-NOM cranes-NOM many be

In a simplex predicate *jibun* (self) binding is taken as a diagnostic for subjecthood[6] and example (67), with a designated subject, analogous to (61), shows that *Taroo* binds *jibun*, so in this case the animate nominative argument *Taroo* is the subject of *iru*. This casts some doubt on Tateishi's theory, since *jibun* binding is a more general relation based on antecedence among the arguments of a predicate, while *ni/ga* alternation is a rather arbitrary criterion for subjecthood, with no extra justification.

(2.67) Taroo ga jibun no ie ni iru.
Taroo-NOM his-GEN house-DAT be
"Taroo is in his own house".

As a conjunction, *te iru* combines with Achievements, Activities, Accomplishments and certain derived States, but not with pure States. In (68) *shinu* (die) is an Achievement and in conjunction with *iru* it yields a perfective reading, while *asobu* (play) in (69) is an Activity and yields a purely progressive reading. Accomplishments also yield a progressive reading, in conjunction with *te iru* (70).

(2.68) kare ga shinde iru
he-NOM die-TE be
He is dead. (perfective achievement)

(2.69) kodomo ga asonde iru
children-NOM play-TE be
The children are playing. (progressive activity)

(2.70) Keiko ga keeki o tukutte iru
Keiko-NOM cake-ACC bake-TE be
Keiko is baking a cake. (progressive accomplishment)

There are some verbs which are States, which appear to be derived potentials: *kikoeru* (can hear) is related to Activity *kiku* (hear, listen) and *mieru* (can see) is related to Activity *miru* (see, look at). These verbs allow *te iru* conjunction, as in example (71). McClure (1994) argues that there are only 3 verbs which are pure States: *iru* (animate be), *aru* (inanimate be) and *iru* (need). In contrast, these

[6]The general case of *jibun* binding is more complex. See Iida (1995).

verbs do not allow conjunction with *te iru* (be) and so *ite iru* (*iru* + *te iru*) (is needing), in (72), is unacceptable.

(2.71) kanojo no na ga seken ni kikoete iru
her-GEN name-NOM people-DAT can hear-TE be
"Her name is known by everyone".

(2.72) *atarashii kutsu ga (boku ni) ite imasu
new shoes-NOM (me-DAT) need-TE be-POL
"*I am needing new shoes".

In its overall aspect the *te iru* conjunction is in contrast with the *te morau* types of construction: *te iru* always displays Stative aspect. This is demonstrated with Achievement, Activity and Accomplishment conjunctions in (73) through (75) respectively.

(2.73) yumeina hito ga shibaraku mae kara takusan shinde iru.
famous people-NOM since a while ago many die-TE be
"Since a while ago a lot of famous people have been dying".

(2.74) denwa ga shibaraku mae kara natte iru.
telephone-NOM since a while ago ring be
"The telephone has been ringing since a while ago".

(2.75) haha ga imoto o isha ni shibaraku mae kara misete iru.
mother-TOP little sister-ACC doctor-DAT since a while ago show-TE be
"My mother took my little sister to see the doctor a while since".

Seq	Verb	Aspect	Args	Output	Args	Effect
a.	*irimasu*	State	ga,ga	N/A	N/A	N/A
b.	*irimasu*	State	ni,ga	"	N/A	N/A
c.	*mieru*	State	ga,ga	State	ga,ga	V-M
d.	*shinu*	Achiev	ga	State	ga	V-M
e.	*kakureru*	Achiev	ni,ga	" "	ni,ga	V-M
f.	*naru*	Act	ga	State	ga	V-M
g.	*yomu*	Act	ga,o	" "	ga,o	V-M
h.	*iku*	Accomp	ga,ni	State	ga,ni	V-M
i.	*kaku*	Accomp	ga,o	" "	ga,o	V-M
j.	*miseru*	Accomp	ga,o,ni	" "	ga,o,ni	V-M

Table 2.8: Verb Frames and Aspectual Readings for *te iru*

The patterns of conjunction showing invariant aspect and Valence-Maintenance are summarized in Table 2.8. The central problem with this data is to determine the relationship between the simplex and the conjunct forms of *iru*. Simplex *iru* has two arguments, a dative *ni* alternating with *ga* marking, and a nominative *ga* marked argument. The binding tests show that either one of them can act as the subject, in some but not all circumstances. The conjunct form of *te iru* is associated with a *te* marked dynamic[7] verb phrase and a *ga* marked argument as subject. Indeed, the *te* marked verb is strongly attached to its head *iru* and this is suggestive of a process of grammaticalization which is coming to distinguish simplex and conjunct constructions. In the general process, complements are grammaticalized before subjects (see Hopper & Traugott (1993)).

2.2.2 shimau

As a simplex verb *shimau* is a ditransitive (*ga,ni,o*) Accomplishment meaning to 'put something away' as indicated in example (76).

(2.76) sore o tsukue no hikidashi ni shimatta.
it-ACC desk-GEN drawer-DAT put away-PAST
"(I) put it away in the desk drawer". (Hiroo)

Martin (1975) indicates that in Old Japanese, *shimau* is an intransitive verb meaning 'to finish', and Hasegawa (1995) notes that a form which she indicates to be transitive, and meaning 'to finish completely' survives in idiomatic expressions such as *mise o shimau* (to quit business). In contemporary Japanese then, simplex *shimau* has two forms, represented by transitive (*ga, o*) and ditransitive (*ga, ni, o*) case frames. These may or may not be historically related, but the separate case frames, and the distinct meanings that go with them, suggest that *shimau* is homonymous.

The pattern of combination of *te shimau* is similar to that of *te morau*, since it combines with Stative-Potentials, Achievements, Activities or Accomplishments and does not alter the valence of the first conjunct, examples (77) - (80).

(2.77) onaka ga miete shimatta
belly-NOM be visible-TE put-PAST
"My belly was showing, to my embarrassment". (Hasegawa)

[7]A dynamic verb is non-Stative, following Lakoff (1965). See Chapter 1 of this thesis.

(2.78) jyoon ga siken ni ukatte shimatta
Joan-NOM exam-DAT pass-TE put-PAST
"Joan passed the exam, to my regret/surprise". (Hasegawa)

(2.79) nikunde wa ikenai to omoi nagara nikunde shimatta
hate-TE TOP must-NEG QUOT think while hate-TE put-PAST
"While I thought I shouldn't hate him/her, I did hate him/her". (Hasegawa)

(2.80) Jyoon ga guuzen ni tegami o sutete shimatta
Joan-NOM accidentally letter-ACC throw away-TE put-PAST
"Accidentally, Joan threw the letter away". (Hasegawa)

Hasegawa gives a detailed account of the semantic complexities of *te shimau* as a conjunct, centering on the observation that *shimau* adds some combination of the modalities of completion and of regret (from the point of view of the speaker). In its aspect, *te shimau* forms a non-agentive reading, as demonstrated by its compatibility with *guuzen ni* (accidentally) in examples (81) and (84). It does not combine with *shibaraku mae kara* (since a while ago), in (82), and so conjunct *shimau* is aspectually an Achievement. The aspectual and valence data for conjunction with *shimau* is summarized for the various aspectual classes and verb frames in Table 2.9.

(2.81) onaka ga guuzen ni miete shimatta.
belly-NOM accidentally be visible-TE regret-PAST
"My belly was accidentally showing, to my embarrassment".

(2.82) *onaka ga shibaraku mae kara miete shimatta.
belly-NOM since a while ago be visible-TE regret-PAST
"My belly was showing since a while ago, to my embarrassment ".

(2.83) haha ga fuun ni mo taorete shimatta.
mother-NOM unfortunately fall over-TE regret-PAST
"My mother unfortunately fell over, to my regret".

(2.84) Jyoon ga guuzen ni tegami o sutete shimatta.
Joan-NOM accidentally letter-ACC throw away-TE finish-PAST
"Accidentally, Joan threw the letter away".

The fact that *guuzen ni* (accidentally) as an adverbial is *-agentive* indicates there is a problem in relating the simplex and conjunct forms of *shimau* in synchronic data, because as an Accomplishment, the simplex form is agentive, and conjunction with Activities and Accomplishments should be *+agentive*. That the overall aspect of the conjunction is *-agentive* is indicative that one or other of the conjuncts is *-agentive*, so this must be the second conjunct form of *shimau*, and as Martin (1975)

Seq	Aspect	Args	Output	Args	Effect
a.	State	ga	N/A	none	N/A
b.	State	ni,ga	"	none	N/A
c.	St-Pot	ga,ga	Achieve	ga,ga	V-M
d.	Achiev	ga	Achiev		
e.	Achiev	ga,ni	" "	ga,ni	V-M
f.	Act	ga	Achiev	ga	V-M
g.	Act	ga,o	" "	ga,o	V-M
h.	Accomp	ga,ni	Achiev	ga,ni	V-M
i.	Accomp	ga,o	" "	ga,o	V-M
j.	Accomp	ga,o,ni	" "	ga,o,ni	V-M

Table 2.9: Verb Frames and Aspectual Readings for *te shimau*

surmises, this form is probably related to the classical intransitive form, rather than the synchronic transitive Accomplishment form.

2.2.3 Summary

The only regularity in the data of *iru* and *shimau* is that of second conjunct aspect compatibility, and even that relies on an interpretation of the conjunct form of *shimau* as having an aspectual classification different from that of the simplex form(s). The simplex form of *shimau* has a *ni* marked locative argument which is not present in the conjunct form. This difference resembles the pattern between simplex and conjunct forms of *oku*, except that *te shimau* and *te oku* pattern differently with respect to their aspectual classification. The semantics of *te shimau* also show a broad range of meanings involving modalities of completion and regret, which are, however, different from the simplex meaning.

There is perhaps, a closer relationship in meaning between simplex and conjunct forms of *iru*, since in both cases some notion of 'being' is involved: in simplex form an animate argument is 'being' in relation to a location, and the conjunct form involves 'being' in relation to an event. Simplex *iru* has two arguments, a *ga* marked animate argument and a *ni* or *ga* marked locative. Referring back to the previous section, *miru* (see) has two arguments, a *ga* marked subject and an *o* marked accusative object, and it is possible to identify an alternation between the accusative argument in the simplex case, and the *te* marked VP in the conjunct case. If a similar alternation can be identified between simplex and conjunct forms of *iru* then it seems possible that the *ni* marked locative argument alternates with the *te* form: the *ga* marked argument of simplex *iru* is invariably animate, and *te*

iru combines with Activities, Accomplishments and Achievements, and the first two of these have a *ga* marked subject, which is defined to be agentive. The subject of an Achievement is non-agentive, and this brings an anomaly for conjunction with *te iru*, and indeed Achievement + *te iru* yields a perfective meaning, rather than the progressive readings which Activity and Accomplishment conjunctions bring.

2.3 *aru*

In concert with *iru* (animate be), *aru* is a verb of 'being' with two arguments, one of which is a locative marked by dative *ni*, or alternatively by nominative *ga*. The other, *ga* marked argument is generally inanimate, and this is what distinguishes simplex *aru* from simplex *iru*. The two alternatives of *aru* are shown in examples (85) and (86). The inanimate argument is *handoru* (steering wheel) in both cases, and the locative *kuruma* (car) marked by dative *ni* in (85) and nominative *ga* in (86).

(2.85) kuruma ni handoru ga aru.
car-DAT steering wheel-NOM be
"The car has a steering wheel".

(2.86) kuruma ga handoru ga aru.
car-NOM steering wheel-NOM be
"The car has a steering wheel".

Although it is generally regarded as inanimate, it is not invariably the case that the nonlocative argument must be inanimate: there are exceptional situations in which an animate argument can be "objectified", usually using a collective term such as *kodomo* (children) as reported by Martin (1975) and shown in (87).

(2.87) dare ni/ga kodomo ga arimasu ka.
who-DAT/NOM kids-NOM be-POL QN
"Who has kids"?

in the presentation of the data of *iru* (animate be) I have already pointed out Tateishi's (1994) contention that the argument which alternates *ni/ga* is subject, but that this possibly contradicts the evidence from *jibun* binding. Example (88) below supports Tateishi's contention, because the possessive *dare ni* (whose) binds *jibun*. Unfortunately the *jibun* tests cannot be used to determine whether the pure *ga* marked argument is subject, as the animacy requirement forces the use of *iru*.

It may be that *aru* patterns with *iru* as underspecified for subject projection, but the simplex data alone is inconclusive for this hypothesis.

(2.88) dare ni/ga jibun no kodomo ga arimasu ka.
who-DAT/NOM own-GEN kids-NOM be-POL QN
"Who has their own kids"?

The conjunct forms of *aru* show a greater range of patterns than the conjunct forms of *iru*, and this is another difference between the 'be' verbs. There is a Valence-Maintaining form of *te aru* which takes an Activity or Accomplishment first conjunct, shown in examples (89) and (90). Achievements are inadmissible in this form, as shown in (91).

(2.89) watashi wa takusan nete aru wa yo
I-TOP a lot sleep-TE be TOP EMPH
"I've slept a lot". (Hasegawa)

(2.90) Hurukawa wa nimotsu o issai reikusando
Hurukawa-TOP luggage-ACC all Lakesand
hoteru ni nokosite atta to iu
hotel-DAT leave-TE be-PAST QUOT say
"They say that Hurukawa had left all his
luggage at the Lakesand Hotel". (Jacobsen)

(2.91) *neko ga ido no naka ni ochite aru.
cat-NOM well-GEN inside-DAT fall-TE be

The Valence-Reducing form applies only to transitive and ditransitive Accomplishments. A transitive verb such as *tukuru* (bake) in (92) takes a *ga* marked subject and an *o* marked object in its simplex form, but in conjunction with *te aru* (93), the subject of *tukuru* is elided and its object is marked with *ga* and appears as the subject of the conjunction. A similar process applies to ditransitives as in (94) and (95). The verb *nokosu* means 'to leave something somewhere' and has *ga, ni, o* case marking. In conjunction with *te aru* the subject is elided as above, the accusative object promoted to subject and the *ni* marked locative remains. This *te aru* construction has been dubbed as an *Intransitivizing Resultative* by Matsumoto (1990).

(2.92) Eiko ga keeki o tukutta
Eiko-NOM cake-ACC bake-PAST
"Eiko baked a cake".

(2.93) keeki ga tukutte aru
cake-NOM bake-TE be
"The cake is/has baked".

(2.94) Eiko ga nimotso o hoteru ni nokoshita
Eiko-NOM luggage-ACC hotel-DAT leave-PAST
"Eiko left her luggage in the hotel".

(2.95) nimotso ga hoteru ni nokoshite aru
luggage-NOM hotel-DAT leave-TE be
"The luggage is left in the hotel".

Activities do not participate in this Valence-Reducing construction, and there are some Accomplishments which do not. The verb *oshieru* means 'teach' and may be used in transitive or ditransitive form. Example (96) shows that the accusative argument is the subject taught, *kezaigaku* (economics). It is not admissible to put this in a resultative construction, so (97) is bad. Matsumoto (1990) argues that in order to participate in Resultative *te aru*, the outcome of the verb must leave a "visible trace". The outcome of "baking a cake" is "a cake, baked", but there is no visible trace following an Activity such as 'teaching'. Similarly, the verbs *noboru* (climb) and *kaeru* (return (home)) take a *ni* marked locative argument, as for *noboru* in (98). These verbs participate in the Valence-Maintaining form of *te aru* conjunction (99), but not in the Valence-Reducing form (100).

(2.96) Kenjo ga keizaigaku o oshieru.
Kenjo-NOM economics-ACC teach
"Kenjo teaches economics".

(2.97) *keizaigaku ga oshiete aru.
economics-NOM teach-TE be

(2.98) boku ga Hakusan ni nobotta.
I-NOM Hakusan-DAT climb-PAST
"I climbed Hakusan".

(2.99) boku ga Hakusan ni nobotte aru.
I-NOM Hakusan-DAT climb-TE be
"I have climbed Hakusan".

(2.100) *Hakusan ga nobotte aru.
Hakusan-NOM climb-TE be

Valence-Reducing *te aru* patterns with *te iru* in its overall aspect: compatibility with adverbs patterns with the second conjunct, as Stative, Examples (101) and (102) showing the contrast.

(2.101) keeki ga shibaraku mae kara tukutte aru
cake-NOM since a while ago bake-TE be
"The cake is baked since a while ago".

(2.102) *keeki ga moo sugu tukutte aru
cake-NOM soon bake-TE be
"*The cake is baked soon".

In its Valence-Maintaining form, however, te aru takes the aspect of the first conjunct, which may be Activity (103) or Accomplishment (104).

(2.103) Jyoon ga soto ni kuruma o tomete aru
Joan-NOM outside-DAT car-ACC stop-TE be
"Joan has parked the car outside". (Hasegawa)

(2.104) Jyoon ga shinchyou ni/*shibaraku mae kara soto ni
Joan-NOM deliberately/since a while ago outside-DAT
kuruma o tomete aru
car-ACC stop-TE be
"Joan has deliberately/*since a while ago parked the car outside".
(Hasegawa (1995))

Seq	Verb	Args	Output	Args	Effect
a.	imasu	ga,ga	N/A	none	N/A
b.	imasu	ni,ga	"	none	N/A
c.	mieru	ga,ga	"	none	N/A
d.	shinu	ga	N/A	none	N/A
e.	kakureru	ga,ni	"	none	N/A
f.	odoru	ga	Act	ga	V-M
g.	oshieru	ga,o	"	ga,o	V-M
h.	kaeru	ga,ni	Accomp	ga,ni	V-M
i.	shimeru	ga,o	" "	ga,o	V-M
j.	oku	ga,o,ni	" "	ga,o,ni	V-M
k.	shimeru	ga,o	State	ga	V-R
l.	oku	ga,o,ni	" "	ga,ni	V-R

Table 2.10: Verb Frames and Aspectual Readings for *te aru*

The data for *te aru* is summarized in Table 2.10. Simplex *aru* is a State with a pure *ga* marked argument and an argument with *ni/ga* alternate case marking which shows independent evidence of subjecthood. Conjunct *te aru* has a *ga* marked

subject and participates in Valence-Maintaining and Valence-Reducing constructions with the restriction that the first conjunct must be agentive in the Valence-Maintaining case, and a (di)transitive Accomplishment in the Valence-Reducing case. The overall conjunction takes Stative aspect in this case.

2.4 Mixed Valence Maintaining and Valence Increasing Forms

A complex and subtle analysis of the various meanings and usages of the *te iku* and *te kuru* constructions is offered in Hasegawa (1995), including actual motion and various metaphorical motion uses. The present study is interested in the valence and aspectual characteristics of simplex and conjunct forms of *iku* (go) and *kuru* (come), which display both Valence-Maintaining and Valence-Increasing characteristics. In addition, the verbs of motion, in first conjunct position, provide a special case in the Valence-Maintaining analysis. The data description is divided into subsections addressing each of these phenomena: Valence-Maintaining in 2.4.1, Motion verbs in 2.4.2 and the cases of Valence-Increase in 2.4.3.

2.4.1 Valence-Maintaining Forms

The verbs *iku* and *kuru* are the motion verbs for 'going' and 'coming' respectively. Their behaviour is sufficiently similar that they are best described together. Both are intransitive Accomplishments taking two arguments, a *ga* marked agentive subject and a *ni* marked locative, or directional argument. Canonical use is shown by examples (105) and (106) where the agent 'Joan' participates in an act of coming or going to the location, the store.

(2.105) Jyoon ga ano mise ni kuru.
Joan-NOM that store-DAT comes
"Joan comes to the store". Hasegawa (1995)

(2.106) Jyoon ga ano mise ni iku.
Joan-NOM that store-DAT goes
"Joan goes to the store". Hasegawa (1995)

Conjoined uses include Achievements, Activities and Accomplishments, all of which exhibit Valence-Maintaining variants, but there are special problems relating to verbs of motion such as *aruku* (walk) and *hashiru* (run), and certain idiosyncratic verbs which result in a Valence-Increase, with an extra *ni* marked locative.

The verb *ochiru* (fall) in example (107) is an Achievement with (*ga, ni*) undergoer and locative arguments, so 'the cat' is the unwitting undergoer of falling into the well. Conjunction with *te iku* or *te kuru* creates what Hasegawa calls a **point-of-view** construction, adding "atmosphere and vivid imagery to the statement". While example (107) is just a statement about the cat, the alternatives in (108) also offer the point-of-view, in terms of physical location, of the speaker. In the case of *ochite itta* (went falling), the speaker is somewhere at the top of the well. In the case of *ochite kuru* (comes falling), the speaker is specifically at the bottom of the well, and the cat has landed.

(2.107) neko ga ido no naka ni ochita.
cat-NOM well-GEN middle-LOC fall-PAST
"The cat fell down the well". Hasegawa (1995)

(2.108) neko ga ido no naka ni ochite itta/kita.
cat-NOM well-GEN middle-LOC fall-TE go-PAST/come
"The cat went/came falling down the well". Hasegawa (1995)

This construction takes the aspect of the first conjunct and so as in (109) the event combines with adverbial *guuzen ni* (accidentally), which cannot combine with an agentive verb.

(2.109) neko ga guuzen ni ido no naka ni ochite itta.
cat-NOM accidentally well-GEN middle-LOC fall-TE go-PAST
"The cat accidentally went falling down the well". Hasegawa (1995)

In this combination the *te iku/kuru* construction seems to pattern with *te morau*, which also takes an Achievement as first conjunct and reflects its aspect, in conjunction. Valence-Maintaining Activities pattern in the same way, with a difference in interpretation. The conjunction *odotte iku* (dance and go) patterns aspectually with the first conjunct, as an Activity in example (110). Notice that this is different from 'go to dance' which is expressed using *odori ni iku* (111), a construction related to the simplex usage of *iku*, not its conjunct use.

(2.110) Mai-san ga 3-jikan/*3-jikan de odotte itta.
Mai-NOM for 3 hours/*in 3 hours dance-TE go-PAST
"Mai-san danced for 3 hours/*in 3 hours and went".

(2.111) Mai-san ga 3-jikan de odori ni itta.
Mai-NOM in 3 hours dance-INF-DAT go-PAST
"Mai-san went to dance in 3 hours".

Conjunction of *te iku/kuru* with an Accomplishment has the same interpretation, (112), although since both conjuncts are Accomplishment, the overall aspect is the same. By inference with the way that Achievements and Activities work in these conjunctions, the Accomplishment conjunction is assumed to pattern with the first conjunct, (113).

(2.112) Keiko-san ga gohan o tabete itta.
Keiko-NOM meal-ACC eat-TE go-PAST
"Keiko ate a meal and went.".

(2.113) Keiko-san ga 1-jikan de gohan o tabete itta.
Keiko-NOM in 1 hour meal-ACC eat-TE go-PAST
"Keiko ate a meal in an hour and went.".

If these examples, (105) to (113) were all that there is to the *te iku/kuru* construction then it would be easy to classify together with *te miru/morau/ageru* as a *te* conjunction which takes the aspect of the first conjunct, but the range of phenomena is broadened by the interaction of *te iku* and *te kuru* with motion verbs, and also certain other idiosyncratic verbs which lead to a Valence-Increasing analysis.

2.4.2 Motion Verbs

There is a range of verbs which Martin (1975) classifies as *quasi-intransitive* in that their range of valences includes a traversal object. The valence patterns of these verbs include unergative *ga*, quasi-transitive *ga, o* as Activities, and intransitive *ga, ni* as Accomplishment. Verbs included in this class are *wataru* (cross), *tooru* (pass through), *tobu* (fly), *hashiru* (run) and *aruku* (walk). As with English, it is possible to make any of the following expressions in (114) to (118).

(2.114) Oguchi-san ga hashiru.
Oguchi-NOM runs
"Mr Oguchi runs".

(2.115) Oguchi-san ga 1-jikan/*1-jikan de hashiru.
Oguchi-NOM for an hour/*in an hour runs
"Mr Oguchi runs for an hour/*in an hour".

(2.116) Oguchi-san ga rooka o aruku.
Oguchi-NOM hall-ACC walks
"Mr Oguchi walks the halls".

(2.117) Oguchi-san ga 1-jikan/1-jikan de rooka o aruku.
Oguchi-NOM for an hour/in an hour hall-ACC walks
"Mr Oguchi walks the halls for an hour/in an hour".

(2.118) Oguchi-san ga eki ni/made hashitta.
Oguchi-NOM station-DAT/-towards run-PAST
"Mr Oguchi ran to the station".

It is not possible however to have both a traversal object and a dative goal, in (119), so these arguments are strictly alternatives.

(2.119) *Oguchi-san ga michi o eki ni hashitta.
Oguchi-NOM streets-ACC station-DAT run-PAST
"*Mr Oguchi ran the streets to the station".

These phenomena create a problem for the simple aspectual classification of verbs, since in each case the motion verb, *hashiru* (run) for instance, is interpreted as an Accomplishment when completed with a *ni* marked goal (*eki ni* to the station), and is underspecified for Activityhood and Accomplishmenthood in the case of a traversal object (i.e. *rooka o* the halls). This can be partly explained by the observation of Verkuyl (1993), that the aspect of a sentence is not solely determined by the aspect of the matrix verb, but is composed from other elements as well, including goal adjuncts and quantificational information. Example (120) shows a telic Accomplishment, but when the object is pluralized, the sentence may be interpreted as atelic (121). Similarly, while the goal argument creates a telic sentence when composed with an atelic verb, it acquires an atelic interpretation when quantified, as shown in the contrast between Examples (122) and (123).

(2.120) I baked a cake (in 30 minutes/*for 30 minutes).
(2.121) I baked cakes for 3 hours.
(2.122) I cycled to the station in 30 minutes/*for 30 minutes.
(2.123) I cycled to the station every day for 5 years/*in 5 years.

The goal form and traversal object form of a motion verb pattern differently with respect to aspect. In (124) the goal form *eki ni hashitte* (running to the station) is strictly telic, while in (125) the traversal object *rooka o aruite* (walking the halls) is underspecified for telicity, being compatible with both *1-jikan* (for an hour) and *1-jikan de* (in an hour).

(2.124) Keiko-san ga *5-funkan/5-funkan de eki ni hashitte itta.
Keiko-NOM *for 5 minutes/in 5 minutes station-DAT run-TE go-PAST
"Keiko-san went to the station running in 5 minutes/*for 5 minutes".

(2.125) Oguchi-san ga 1-jikan/1-jikan de rooka o aruite iku.
Oguchi-NOM for an hour/in an hour hall-ACC walk-TE go
"Oguchi walked the halls for an hour/in an hour".

The broad pattern for *te iku* and *te kuru* is that of a Valence-Maintaining conjunction that takes the aspect of the first conjunct. What remains to be described are the verbs which in conjunction result in a pure Valence-Increase, and this follows.

2.4.3 Valence-Increasing Forms

The verbs *motsu* (hold), *kiru* (wear) and *tureru* (accompany) are transitive Activities, while *nageru* (throw) is a transitive Accomplishment. None of these verbs in its simplex form is 'natural' with a *ni* locative argument, so while the transitive form shown in (126) is canonical, addition of the *ni* locative in (127) is bad. A locative adjunct, *gakkoo de* (at school), in (128), is acceptable.

(2.126) Naoko-san ga hon o motta.
Naoko-NOM book-ACC hold-PAST
"Naoko-san held the book".

(2.127) Naoko-san ga *gakkoo ni hon o motta.
Naoko-NOM school-DAT book-ACC hold-PAST
"*Naoko-san held the book to the school".

(2.128) Naoko-san ga gakkoo de hon o motta.
Naoko-NOM book-ACC hold-PAST
"Naoko-san held the book at the school".

The verb *motsu* (hold) is atelic, as demonstrated by association with time-adjuncts *1-jikan* (for an hour) but not *1-jikan de* (in an hour), in (129).

(2.129) Naoko-san ga 1-jikan/*1-jikan de hon o motta.
Naoko-NOM for an hour/*in an hour book-ACC hold-PAST
"Naoko-san held the book for an hour/*in an hour".

What is unique among the varied phenomena of *te* conjunctions is that, in conjunction with *te iku*, these verbs license a *ni* locative and the entire conjunction reads as an Accomplishment, as is illustrated in Examples (130) and (131).

(2.130) Naoko-san ga gakkoo ni hon o motte itta.
Naoko-NOM school-DAT book-ACC hold-te go-PAST
"Naoko-san took the book to school".

(2.131) Naoko-san ga *5-funkan/5-funkan de gakkoo ni hon o
Naoko-NOM *for 5 minutes/in 5 minutes school-DAT book-ACC
motte itta.
hold-TE go-PAST
"Naoko-san took the book to school for 5 minutes/in 5 minutes".

Although it is possible to move the locative argument between the two conjuncts, this then forms a different construction, the Core Conjunction, discussed by Hasegawa (1995).

(2.132) Naoko-san ga hon o motte gakkoo ni itta.
Naoko-NOM book-ACC hold-TE school-DAT go-PAST
"Naoko-san went to school holding the book".

Hasegawa shows that the verb *kiru* (put on, wear), behaves in the same way as *motsu*. In (133) *kiru* does not admit a *ni* marked argument, while the conjunction in (134) does so. As with the case of *motsu* above, the conjuncts may be split by the locative, in (135), again indicating a Core Conjunction.

(2.133) *Marii ga sogizyoo ni akai huku o kita.
Mary-NOM funeral-DAT red dress-ACC wear-PAST
"*Mary wore a red dress to the funeral". (intended)

(2.134) Marii ga sogizyoo ni akai huku o kite kita.
Mary-NOM funeral-DAT red dress-ACC wear-TE come-PAST
"Mary came to the funeral wearing a red dress".

(2.135) Marii ga akai huku o kite sogizyoo ni kita.
Mary-NOM red dress-ACC wear-TE funeral-DAT come-PAST
"Wearing a red dress Mary came to the funeral ". (Hasegawa (1995))

2.4.4 Summary of iku/kuru

As with all other *V-te V* constructions, co-occurrence with pure states is ruled out, though conjunction with derived states such as *mieru* or *kikoeru* is admissible. All other aspect classes are generally admissible and most of these cases result in a Valence-Maintaining conjunction with the overall aspect of the first conjunct. Although *te iku* and *te kuru* pattern together in relation to the various verbal phenomena, there are occasions when either *iku* or *kuru* is not admissible. Hasegawa notes

Seq	Verb	Aspect	Args	Output	Args	Effect	kuru	iku
	States							
a.	*iru*	State	ga,ga	N/A	none	N/A	no	no
b.	*iru*	State	ni,ga	"	none	N/A	no	no
c.	*mieru*	State	ga,ga	State	ga,ga	V-M	yes	yes
	Achievements							
d.	*kieru*	Achiev	ga	Achiev	ga	V-M	yes	yes
e.	*ochiru*	Achiev	ga,ni	" "	ga,ni	V-M	yes	yes
	Activities							
f.	*odoru*	Act	ga	Act	ga	V-M	yes	yes
f.	*furu*	Act	ga	"	ga	V-M	no	yes
g.	*oyogu*	Act	ga,o	Act	ga,o	V-M	yes	yes
g.	*hashiru*	Act	ga,o	"	ga,o	V-M	yes	yes
f.	*oshieru*	Act	ga	Act	ga	V-M	yes	yes
g.	*oshieru*	Act	ga,o	"	ga,o	V-M	yes	yes
g+.	*oshieru*	Act	ga,o,ni	"	ga,o,ni	V-M	yes	yes
	Accomplishments							
h.	*kaeru*	Accomp	ga,ni	Accomp	ga,ni	V-M	yes	yes
h.	*noboru*	Accomp	ga,ni	" "	ga,ni	V-M	yes	yes
i.	*korosu*	Accomp	ga,o	" "	ga,o	V-M	yes	yes
i.	*taberu*	Accomp	ga,o	" "	ga,o	V-M	yes	yes
j.		Accomp	ga,o,ni					
	Motion Verbs							
m1.	*aruku*	Accomp	ga,ni	Accomp	ga,ni	V-M	yes	yes
m2.	*aruku*	Act	ga,o	Act/Acc	ga,o	V-M	yes	yes
m3.	*aruku*	Accomp	ga,o	Act/Acc	ga,o	V-M	yes	yes
	Valence Increasers							
g1.	*motsu*	Act	ga,o	Accomp	ga,o,ni	V-I	yes	yes
g1.	*tureru*	Act	ga,o	" "	ga,o,ni	V-I	yes	yes
g1.	*kiru*	Act	ga,o	" "	ga,o,ni	V-I	yes	yes
i.	*nageru*	Accomp	ga,o	" "	ga,o,ni	V-I	yes	yes

Table 2.11: Verb Frames and Aspectual Readings for *te iku/kuru*

that transfer verbs such as *nageru* (throw) co-occur only with *kuru* and not with *iku*. Achievements which connote appearance occur with *kuru*, such as *otozurete kuru* (come visiting) and *arawarete kuru* (gradually appear). Those which connote disappearance occur with *iku*, such as *kiete iku* (gradually die out) or *tokete iku* (gradually melt). The range of motion verbs, *aruku* (walk) and *hashiru* (run) for examples, are ambiguous between Activity and Accomplishment readings when associated with a traversal object, and are Accomplishments when associated with a goal argument (*eki ni hashiru* - run to the station). These interpretations are borne out in conjunction, too. Finally, the most puzzling phenomena to occur in any *V-te*

Pattern	Verb	Class	Conjs	Output	Valence
2	*aru*	State	Accomp	State	2nd onj
1	*aru*	State	agentive	agentive	1st conj
2	*iru*	State	dynamic	State	2nd conj
3	*shimau*	Accomp	dynamic	Achieve	1st conj
1	*miru*	Act	agentive	agentive	2nd conj
1	*morau*	Accomp	all	all	2nd conj
1	*ageru*	Accomp	agentive	agentive	2nd conj
1/2	*hoshii*	State	all	both	'2nd conj'
1	*oku*	Accomp	agentive	agentive	1st conj
1	*iku*	Accomp	dynamic	dynamic	2nd conj
2	*iku*	Accomp	agentive	Accomp	V-I
1	*kuru*	Accomp	dynamic	dynamic	2nd conj
2	*kuru*	Accomp	agentive	Accomp	V-I

Table 2.12: Overall Summary of *V-te V* Data

V conjunction are those idiosyncratic verbs such as *motsu - motte iku* (hold - take ... to), which result in a Valence-Increase with an apparently unmotivated locative argument. The range of phenomena of *te iku* and *te kuru* are summarized in Table 2.11.

2.5 Summary and Preliminary Phrase Structural Analysis

The overall correlation of *te* conjunctions shows three patterns of behaviour. The predominant pattern is for the aspect of the first conjunct to be compatible with adverbial adjuncts, though admissibility of the class of first conjuncts depends on the head (second) verb. Conjunctions which pattern this way are identified in Table 2.12 as Pattern 1. Both *hoshii* and *oku* are variants of this pattern, because their simplex arguments cannot be mapped into the conjunct form. Pattern 2 is characterized by conjunctions taking the overall aspect of the second verb. Notable among these forms are Stative verbs *iru*, (Valence-Reducing) *aru*, and (Valence-Increasing) Accomplishments *iku* and *kuru*. The adjective *hoshii* is notable as being underspecified for both patterns. The third pattern, applicable only to *te shimau*, is a variation of Pattern 2, since the output aspect is invariant, but as an Achievement it differs from the simplex form.

The development of a typology which is sensitive to aspect is also useful in showing patterns of restriction on conjunction. For example *miru* and *ageru* combine with an agentive verb as first conjunct, *iku*, *kuru* and *iru* combine with a dynamic verb; *morau* combines with a verb of any aspect class; and *aru* combines with an

Accomplishment in the Valence-Reducing case, or with an agentive verb in Valence-Maintaining. In some sections of her thesis, Hasegawa (1995) also makes some allusion to aspect, though it is never central to her account. I find, however, that the use of aspect as a typological classifier for the data offers a natural bridge for the development of selectional restrictions among the entire range of *te* conjunctions. This is demonstrated in the analysis later in this thesis.

The valence patterns tell a different story. If the *te* marked verb is treated as an alternate argument of the second conjunct, alternating with an accusative nominal in the cases of *miru* (see), *morau* (receive) and *ageru* (give), and with a dative nominal in the cases of *iru* (animate be), *iku* (go) and *kuru* (come), then the hypothesis that these verb alternations are polysemous becomes viable, and an investigation along these lines is conducted in Chapter 4. The verbs *oku* (put) and *shimau* (put), and the adjective *hoshii* (want) show simplex/conjunct alternations in which the valence patterns are different, and these are listed in the lexicon as homonymous forms. The Valence-Increasing *te iku* and *te kuru* forms are anomalous, since if the *te* conjunct is counted as an argument which alternates with the locative of *iku/kuru*, then the extra locative licensed by the conjunction cannot be clearly identified as an argument of either of the separate conjuncts.

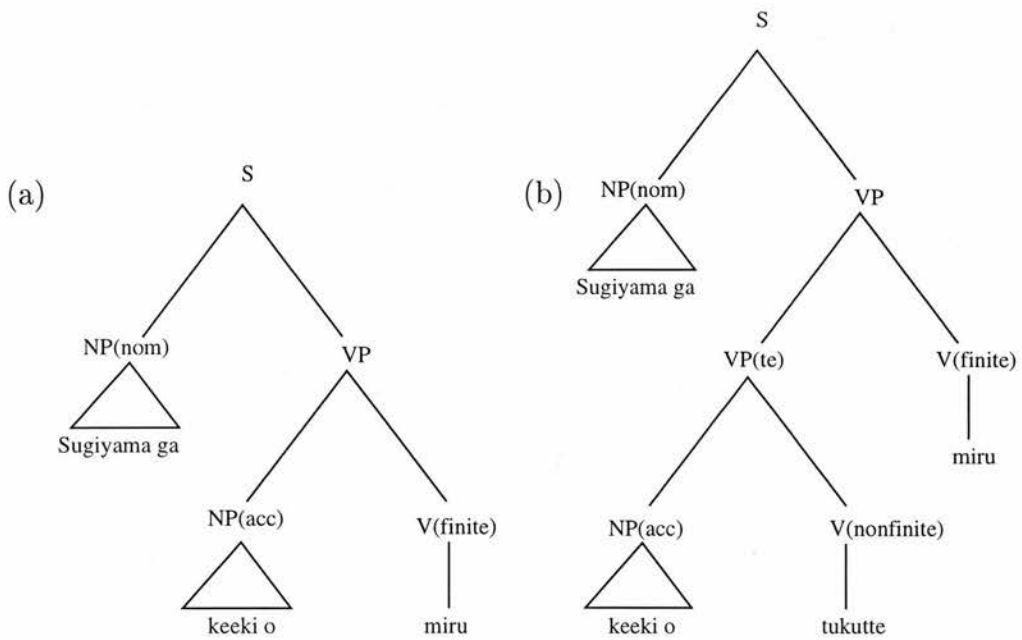


Figure 2.1: Phrase Structure Trees for Simplex and Conjunct *miru*

Since this thesis develops a lexicalist analysis of the data presented in this chapter, a preliminary phrasal analysis is useful at this point. The key to the analysis is in

the basic phrase structures generated by the head verbs, *miru*, *morau*, *iru*, etc. The range of patterns for *miru* is sketched in figures 2.1 and 2.2.

In Figure 2.1(a) *miru* is a finite verb which projects a *ga* marked NP as subject and an *o* marked NP as direct object. The postpositional case markers in this configuration are conventionally assigned nominative case for *ga* and accusative case for *o*, respectively. Figure 2.1(b) shows the projection of the *V-te miru* construction. Again, *miru* projects a *ga* marked NP as subject, but the object position is taken by a *te* marked VP. A *te* marked verb is regarded as non-finite (analogous to English V-ing), and its subject is not expressed in the syntax.

A further extension to these phrasal structures comes with the consideration of adjunct attachment. Adverbial adjuncts in Japanese are free word order, and can scramble freely with the arguments. Tests for aspect are performed with adverbials such as *1-jikan* (for an hour), testing for atelicity, and *1-jikan de* (in an hour), testing for telicity. In the simplex case, the adverbial can attach anywhere before the verb, as shown in Figure 2.2(a). In the conjunct case of *V-te miru*, the adverbial is compatible with the aspect class of the first conjunct verb, not with the second, and this is independent of adjunct ordering in the surface structure. Figure 2.2(b) illustrates.

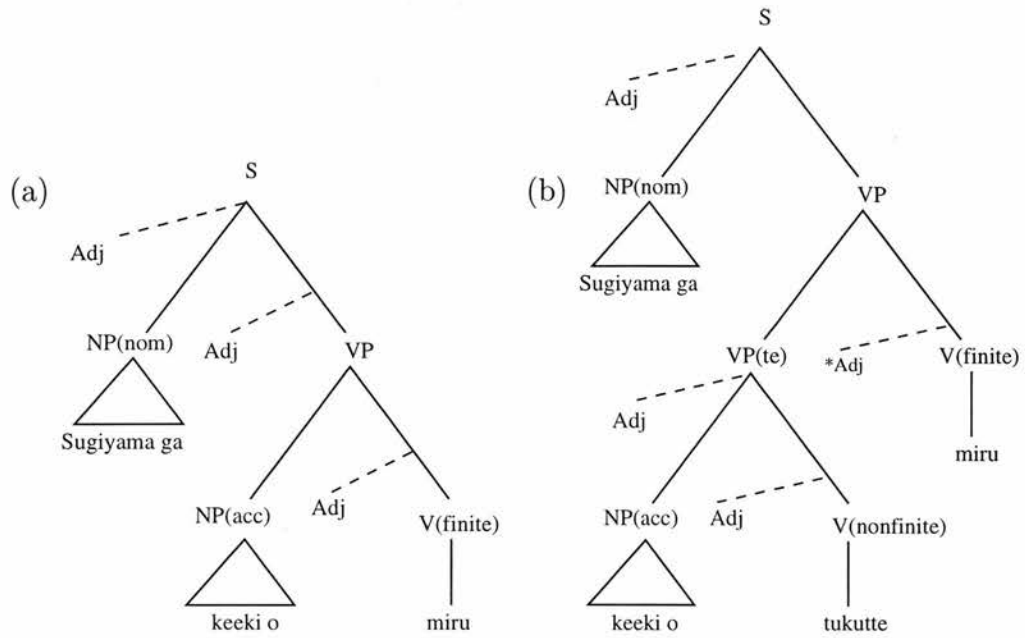


Figure 2.2: Adjunct Attachments for Simplex and Conjunct *miru*

Chapter 3

The HPSG Framework

Head Drive Phrase Structure Grammar (HPSG) (Pollard & Sag (1987), Pollard & Sag (1994), Sag (1997)) is a lexicalist theory of grammar which uses a theory of types (see Carpenter (1992)) to define lexical items of considerable complexity, and combine them under unification (see Shieber (1986)) with a small set of phrasal schemata. Since the reemergence of phrase structure as a principle of syntactic combination with Generalized Phrase Structure Grammar (GPSG) (Gazdar *et al.* (1985)), there has been significant work done to reduce the proliferation of phrase structural rules by abstracting generalizations from the lexical components of language. This reduction in the number of phrase structure rules goes together with an increase in the complexity of lexical items. The most general unit of linguistic representation in HPSG is the *sign*, used to represent lexical or phrasal items. The sign is a complex type which comprises a set of properties, known in HPSG as *features*, which take values that may themselves be simple or complex types. A simple or a complex type may be articulated into two or more subtypes, and every subtype inherits all the features of its immediate parent. Flickinger (1987) develops a theory of the lexicon in which multiple inheritance is employed to cross classify word classes and feature classes so as to allow redundant information to be minimized in the specification of lexical items and phrasal conjunctions, and the mechanism of inheritance works as follows: Single inheritance typing looks like a tree where each local (parent) node has one or more daughters. All features introduced by the parent are automatically inherited by the daughters, and each daughter may introduce its own features, not shared with the parent or with siblings. This is illustrated in Figure 3.1(a), where both Daughter1 and Daughter2 inherit FeatureA from their parent. Daughter1 introduces its own feature, FeatureB and Daughter2 introduces FeatureC. The children of Daughter1 inherit all its features and generate

zero or more new features. The new features introduced in a subtree are shared only by direct descendants, thus FEATUREC introduced by Daughter2 is not shared with Daughter1 or its descendants.

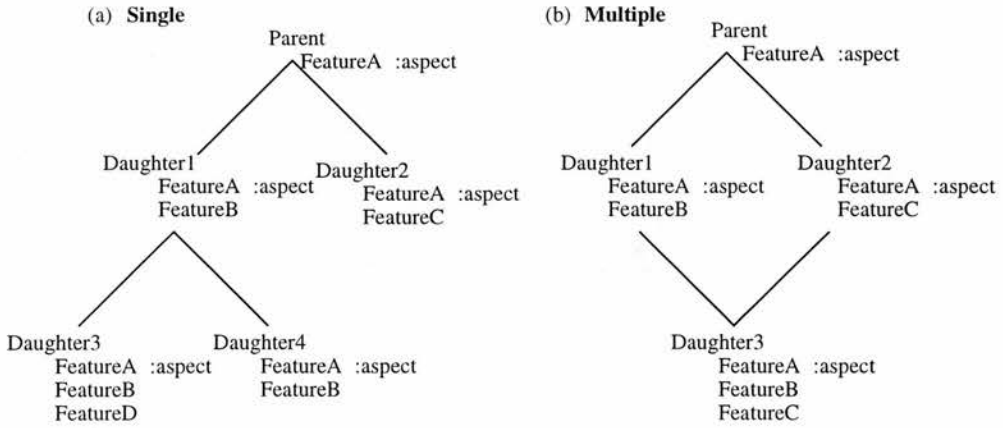


Figure 3.1: Single and Multiple Inheritance

In the multiple inheritance model of 3.1(b), Daughter1 and Daughter2 inherit FEATUREA from their parent, as before, and introduce their own features: FEATUREB for Daughter1 and FEATUREC for Daughter2. Daughter3 is the common descendant of both Daughter1 and Daughter2, and it inherits all the features of its ancestors.

Each feature has a value which may also be a hierarchical type. In Figure 3.1 FEATUREA has a value *aspect*[1], which has subtypes *dynamic* and *state*; the *dynamic* subtype has further subtypes of *telic* and *agent*. If FeatureA is introduced in the parent with value *aspect*, then daughters 1 and 2 inherit the same value or a more restricted subtype. In the multiple inheritance model in 3.1(b), both daughter 1 and 2 can have the value *dynamic*, or indeed daughter 1 can have the more restricted value *telic* while daughter 2 has the value *agent*. Daughter3 in Figure 3.1(b), with two parents, inherits a value of the feature compatible with the values of both its parents. The only value in the *aspect* hierarchy (Figure 3.2) which is a subtype of both *agent* and *telic* is *accomplishment*: so FEATUREA on Daughter3 must carry this value. In a computational system such as ALE (Carpenter & Penn (1998)), the attempt to create a lexical entry with a complex type of Daughter3, having a value for FEATUREA with type anything other than *accomplishment*, results in a type error. This is what enforces the constraints in a strongly typed system.

In the aspect hierarchy of Figure 3.2, subtypes of *telic* are Accomplishment and Achievement, subtypes of *agent* are Activity and Accomplishment, and the value of

[1] Although they are obligatorily present, the values on other features are not shown in Figure 3.1.

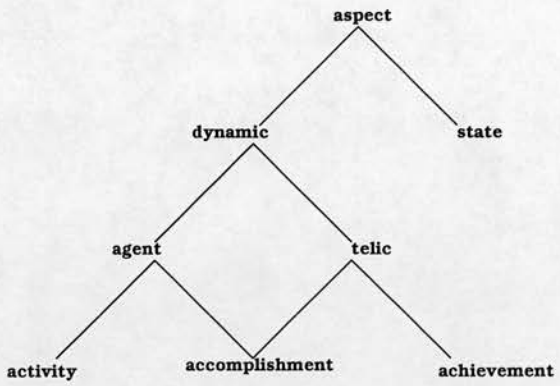


Figure 3.2: Aspect Type Hierarchy

FEATUREA on daughter3 is constrained to bear the the value Accomplishment, as subtype of both *telic* and *agentive*[2].

Inheritance in the type hierarchy is the predominant organizing principle in the lexicon. In Section 1.1 the representation of lexical items for Japanese is discussed, and related to the principles of lexical organization used by HPSG for the analysis of other languages. In addition to this theory of lexical organization, HPSG is a syntactic theory, and lexical signs are combined together under phrase structure schemata and constrained by universal and language specific principles. The phrase structures necessary for an account of the data of Japanese verb alternations are described in Section 3.2. In the development of linking patterns between syntactic and semantic roles, argument structure (ARG-S) developed through two separate lines of evolution: through the ROLES list of Wechsler (1995), and from the SUBCAT list of Pollard & Sag (1994) through to the ARG-S used in Manning (1994), as the locus of binding theory. These two contrasting approaches to the same phenomenon are compared and resolved in Section 3.3, and the resultant ARG-S feature is used as prerequisite to the discussion of Davis (1995) and his linking theory over a reduced set of semantic roles, in Section 3.4. The phrasal type hierarchy of Sag (1997) is introduced in Section 3.5, and the chapter is concluded in 3.6 with a review of the support which HPSG provides for polysemous analyses.

[2]For more detail on the mathematics of inheritance systems see Touretzky (1986). HPSG also incorporates techniques of unification and structure-sharing (Kay 1979, Shieber 1986). The theory is formalized in Carpenter (1992) so that lexical information and phrase structure are defined in a totally well typed and sort resolved system. A computational realization of the system is provided in Carpenter & Penn (1999) Attribute Logic Engine (ALE).

3.1 The Lexicon

The information in a lexical item is organized under the *sign* and contains syntactic, semantic and phonological information. This thesis is not concerned with the phonology of Japanese so wherever the PHON feature appears, its use is confined to identifying the romanized orthography of a word, and the lexical entry is usually identified by this orthography. The verb *nomu* (drink) is used as an example. The syntactic information is gathered under a complex SYNSEM feature, with feature CAT denoting the syntax and feature CONT denoting the semantics. All lexical items are defined to be of some categorial head such as verb, noun, marker or adjective, and *nomu* is a *verb*, which is a value of the feature HEAD under *cat*. Categorial heads are arranged in a type hierarchy under the supertype *head*, so the feature HEAD has a value *head*, or one of its subtypes. *Heads* are partitioned into substantive *subst* and case marker *mark*, and *nouns*, *verbs* and *adjectives* are arranged as subtypes of *subst*, as illustrated in Figure 3.3.

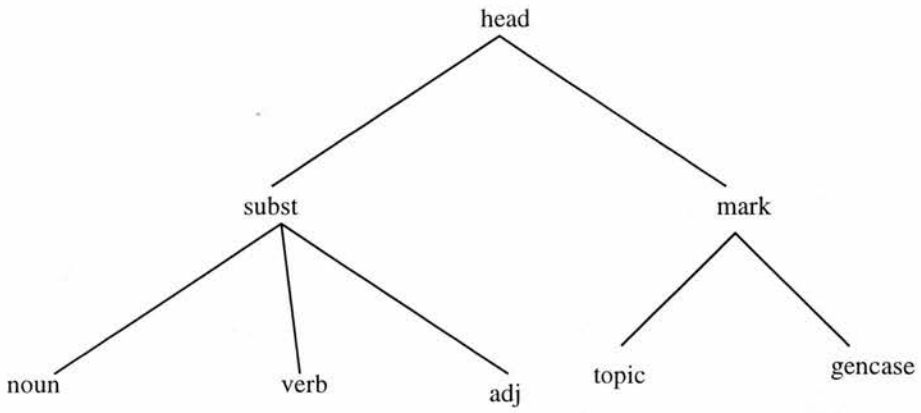


Figure 3.3: The *head* Type Hierarchy

A verb is a predicator and can project a subject and one or more complements. The SUBJ and COMPS features of the verb *nomu* are lists of length one, and its subject and complement have HEADS of type noun. Except for Stative verbs, subjects in Japanese are marked with *ga*[3] and accusative objects are marked with *o*, and the marking value of the phrasal signs for the noun carry this value, finite verbs are unmarked and this is reflected in the marking feature of CAT. The CAT value of a Japanese transitive verb such as *nomu* is shown in Figure 3.4.

[3]Following the presentation of data in the last chapter, the subject of a State may be either dative *ni* marked, or nominative *ga* marked.

$$\begin{bmatrix} \text{PHON } \langle nomu \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } \textit{verb}[\text{VFORM } \textit{fin}] \\ \text{SUBJ } \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS } \langle \boxed{2}[\text{MKNG ACC}] \rangle \\ \text{MKNG UNMARKED} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.4: Syntax of *nomu*.

The semantics of a word is encoded in its content field, including quantificational and non-quantificational ('nuclear') material. Quantificational matters are not addressed in this thesis, so I will focus here on reviewing the nuclear content. Each verb expresses a unique relation with one, two or three arguments and an argument may be nominal, verbal or a complementizer phrase. In Pollard & Sag (1994), Chapter 9, the semantic roles defined by a relation are defined under the NUCLEUS feature. Rather than defining a small general set of thematic roles, Pollard and Sag treat every set of semantic roles as being unique to the particular relation defined by the verb. As a transitive verb, *nomu* is a two-place relation with the roles DRINKER and DRUNK, shown in Figure 3.5.

$$\begin{bmatrix} \text{CONT} \begin{bmatrix} \text{NUC} \begin{bmatrix} \text{RELN } \textit{drink} \\ \text{DRINKER } \textit{index} \boxed{1} \\ \text{DRUNK } \textit{index} \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.5: The *drink* Relation.

These roles are the INDICES of *nominal-objects* linked to the syntactic arguments through the SUBCAT feature, a canonical list of the subject (SUBJ), complements (COMPS) and specifier (SPR) valences.[4] This linking of syntactic arguments and semantic roles is illustrated in Figure 3.6.

For comparison, an intransitive verb such as *odoru* (dance) denotes a one-place relation with a single DANCER role linked to the subject; and a ditransitive verb such as *ageru* (give) denotes a three-place relation with a GIVER role, a BENEFICIARY and a GIFT. This form of role assignment is used in the discussion of Wechsler (1995) in Section 3.3, but replaced with a minimal set of general thematic roles advocated by Davis (1995) and introduced in Section 3.4. Many languages including Japanese have a particularly rich inflectional morphology for the different finite and non-finite verb forms, and various honorific forms. For most verbs, the forms are regularly

[4] I have omitted the SUBCAT list from the discussion of syntactic feature CAT, as it is superseded by ARG-S, following the discussion in Section 3.3.

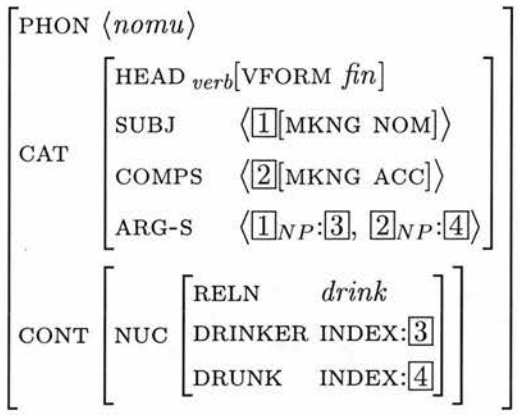


Figure 3.6: Linking *drink*.

derived, and rather than list all forms of all verbs systematically in the lexicon, a set of lexical rules is used to generate alternate forms from the base form. As an example, the past tense lexical rule is shown for deriving *nonda* (drank) from *nomu* (drink) in Figure 3.7. Differences are in the phonology: *nomu* becomes *nonda*; and in the value of the HEAD feature TENSE which is changed from *present* to *past*.

In the general situation a noun does not project any arguments, so its SUBJ and COMPS fields are empty, and as with verbs, a bare noun has no marking value. In Pollard & Sag's (1994) scheme, the CONTENT value of a noun includes an INDEX and it is this value which is structure shared with the verb when the two items combine together. In English, the INDEX is used to specify agreement features for person, number and gender. Japanese nouns do not have these features, but there is possible agreement between honorific prefixes and exalted verbs. However these take no part in this analysis. The verb selects for a marked noun, and the marking value is supplied under combination with a case marker such as nominative *ga*, for which the lexical entry is shown here in Figure 3.9. The analysis of case markers is an area of contention in the literature. Tomabechi (1989) hypothesizes that they are MARKERS, arguing that they are often dropped in informal speech, cf example (2), as compared with the argument marking of the accusative object *biiru o* (the beer) in (1).

(3.1) Oguchi-kun ga biiru o ippai nonda.
Oguchi-NOM beer-ACC complete-ADV drink-PAST
"Oguchi completely drank the beer".

(3.2) biiru ippai nonda.
beer(-ACC) complete-ADV drink-PAST
"(I) drank all the beer".

In subcategorization, the verb selects arguments based on a noun's marking value, but the semantic information of the noun is salient, and the markers contain little or no semantic content. In assesing the grammatical role of case markers, Yoshimoto (1996) follows Tomabechi's account, but on the other hand Gunji (1987) claims case markers as heads assuming that they subcategorize for the noun, essentially without any argument. Siegel (1996) also claims that case markers are heads, arguing by analogy with non-case marking postpositions. In Nightingale (1996), I argue from symmetry that in keeping with a head-final language, case markers are heads. There is a third way: that specifiers are *secondary heads* is the proposal of Cann (1999), where the mother of a head and its specifier inherits the union of (or some function of) the features of both daughters. However a particular constraint on the categorization of case marked noun phrases is that the semantic content of nominal arguments need to be accessible to the verb, for the HPSG formulation of semantic relations to work, and this suggests that case markers are better analysed as of type *mark*. The lexical item for marker *ga* projects no arguments, but has a marking value of *nominative*. Markers select their heads in HPSG, and this is

$$\left[\begin{array}{ll}\text{PHON} & \langle nomu \rangle \\ \text{CAT} & \left[\begin{array}{ll}\text{HEAD} & \left[\text{VERB} \left[\begin{array}{ll}\text{TENSE} & pres \\ \text{VFORM} & fin\end{array}\right]\right] \\ \text{SUBJ} & \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG ACC}] \rangle \\ \text{ARG-S} & \langle \boxed{1}_{NP}:\boxed{3}, \boxed{2}_{NP}:\boxed{4} \rangle\end{array}\right] \\ \text{CONT} & \left[\text{NUC} \left[\begin{array}{ll}\text{RELN} & drink \\ \text{DRINKER} & index:\boxed{3} \\ \text{DRUNK} & index:\boxed{4}\end{array}\right]\right]\end{array}\right]$$

$$\Longrightarrow$$

$$\left[\begin{array}{ll}\text{PHON} & \langle nonda \rangle \\ \text{CAT} & \left[\begin{array}{ll}\text{HEAD} & \left[\text{VERB} \left[\begin{array}{ll}\text{TENSE} & past \\ \text{VFORM} & fin\end{array}\right]\right] \\ \text{SUBJ} & \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG ACC}] \rangle \\ \text{ARG-S} & \langle \boxed{1}_{NP}:\boxed{3}, \boxed{2}_{NP}:\boxed{4} \rangle\end{array}\right] \\ \text{CONT} & \left[\text{NUC} \left[\begin{array}{ll}\text{RELN} & drink \\ \text{DRINKER} & index:\boxed{3} \\ \text{DRUNK} & index:\boxed{4}\end{array}\right]\right]\end{array}\right]$$

Figure 3.7: The Past Tense Lexical Rule

$$\begin{bmatrix} \text{PHON } \textit{biiru} \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } \textit{noun} \\ \text{SUBJ } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \\ \text{MKNG } \textit{unmarked} \end{bmatrix} \\ \text{CONTENT } \textit{index} \end{bmatrix} \end{bmatrix}$$

Figure 3.8: Noun Lexical Entry.

achieved by a head feature SPEC which specifies a noun. Markers have no semantic content.

$$\begin{bmatrix} \text{PHON } \textit{ga} \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } \textit{mark}[\text{SPEC } \textit{noun}] \\ \text{SUBJ } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \\ \text{MKNG } \textit{nom} \end{bmatrix} \\ \text{CONTENT } \textit{content} \end{bmatrix} \end{bmatrix}$$

Figure 3.9: Marker Lexical Entry.

The range of case markers includes *ga*, *o* and *ni*, though sometimes a phrase headed by a postposition such as *kara* (from) is subcategorized as an argument. This is treated as a homonymous form, subtype of mark and not related to the postpositional form. This follows the treatment of Siegel (1996). The *head* type hierarchy for case markers is given in Figure 3.10(a).

All categorial heads have a MARKING value, and markers are associated with *case*, being either nominative (*nom*), accusative (*acc*), dative (*dat*) or genitive (*gen*) Subjects may take either nominative or dative case. A hierarchy of marking values reflecting these case relationships is given in Figure 3.10(b).

Looking back to the semantic relation of *nomu* (drink) in Figure 3.6, the semantic roles are DRINKER and DRUNK. Comparing this with the arguments in example (1), *Oguchi-kun ga* (Mr Oguchi) comprises a noun and a nominative marker and the noun *Oguchi-kun* is a personal name denoting a possible drinker; *biiru o* comprises a noun and an accusative marker, and *biiru* (beer) denotes something drinkable; and nothing in the case markers *ga, o* indicates a relation with the roles DRINKER and DRUNK in the lexical entry of *nomu*. So the noun must be the semantic head of its phrase, and therefore (ruling out the possibility that it is an adjunct of the

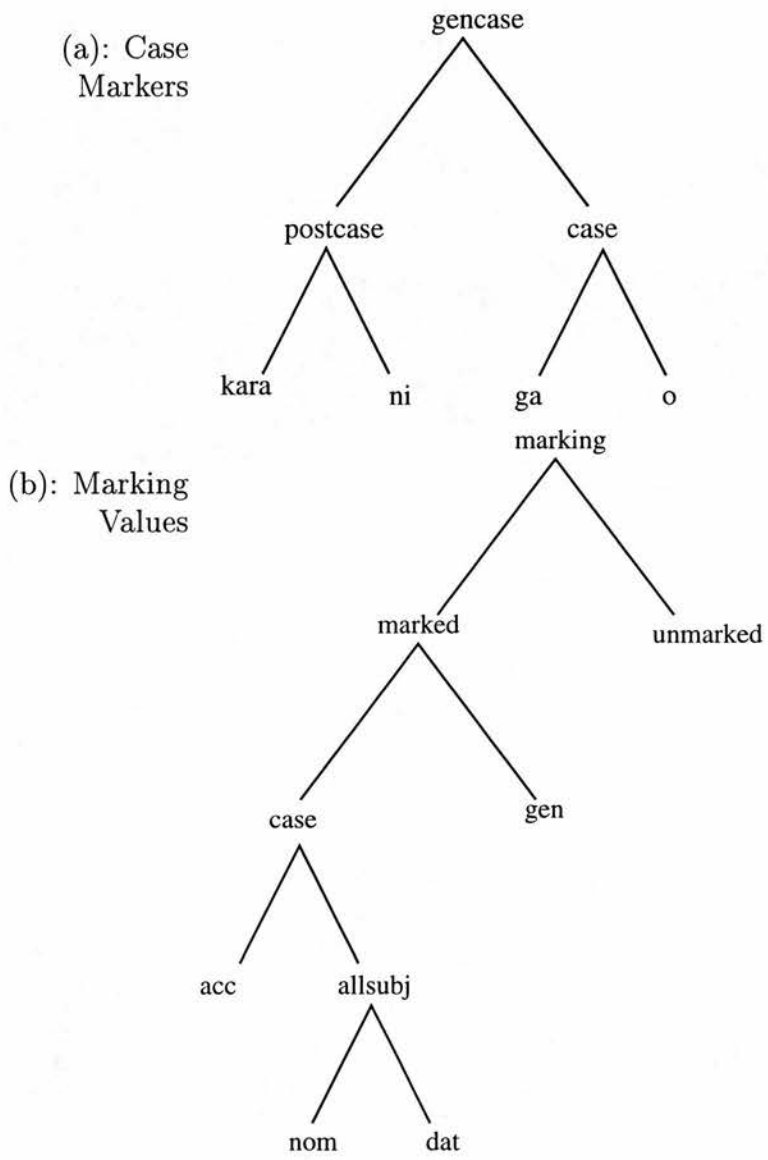


Figure 3.10: Type Hierarchies for (a) *Case Markers* and (b) *Marking* Values

marker), the syntactic head also. The entire Head-Marker phrase is subcategorized by the verb and the operation of phrasal structures is discussed in the next section.

3.2 Phrasal Structures

The sentence given in example (1) is used to illustrate the application of phrasal structures in HPSG, and this is depicted as a tree in Figure 3.11, with Subject-Head, Complement-Head and Head-Marker structures, with the adjunct omitted for simplicity.

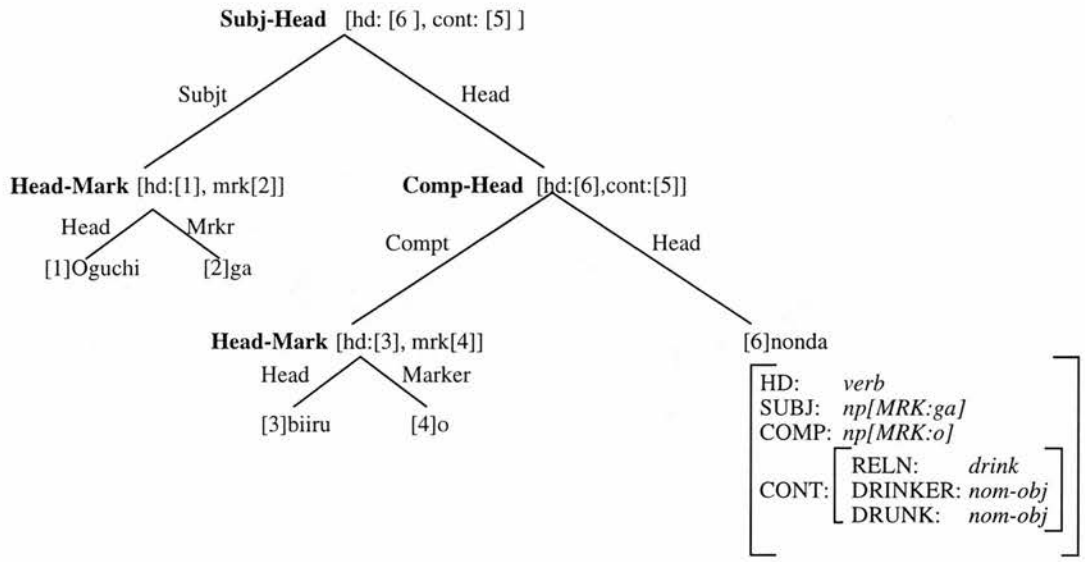


Figure 3.11: HPSG Representation of the "biiru" Sentence

In the figure, *Oguchi-kun ga* and *biiru o* are Head-Marker structures, with the nouns *Oguchi-kun, biiru* as head and the case particles *ga, o* as marker. The lexical entry for the marker has a feature SPEC which selects for a noun. The Head-Marker schema under which they combine illustrates some of the universal and local principles of HPSG. Under the Head Feature Principle (HFP), the parent structure takes its syntactic value from the head daughter, the noun. The HFP simply states:

> **The Head Feature Principle**
> The head value of a headed phrase is identical to the head value of the head daughter.
>
> Borsley (1996:p.50)

Since the head daughter in this case is a noun, the phrasal structure in which it participates is a Noun Phrase. The Semantics Principle also operates to ensure that the CONTENT of the parent structure is identical to the CONTENT of the semantic head, which is the same as the syntactic head in a Head-Marker structure.

> **The Semantics Principle**
> In a headed phrase, the CONTENT value is token-identical to that of the adjunct daughter if the DTRS value is of sort *adj-head-ph*, and with that of the head daughter otherwise.
>
> Pollard and Sag (1994:p.319)

Finally, the Marking Principle operates to ensure that the MARKING value of the Marker daughter and that of the parent are identical (and the marking value of the head daughter is ignored). The Marking Principle is formulated as follows:

> **The Marking Principle**
> In a headed phrase, the MARKING value is token-identical with that of the marker daughter, if any, and with that of the head daughter otherwise.
>
> Pollard and Sag (1994:p.400)

The SYNSEM value of the resulting parent of a Head-Marker structure for the phrase *biiru o* contains the features shown in Figure 3.12, and the only way this differs from the SYNSEM value of noun is that the marking value of the parent is *accusative*, following the operation of the Marking Principle.

$$\begin{bmatrix} \text{PHON } \textit{biiru o} & \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SUBJ} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle \\ \text{MKNG} & \textit{acc} \end{bmatrix} \\ \text{CONTENT} & \textit{index} \end{bmatrix} \end{bmatrix}$$

Figure 3.12: Head-Marker *synsem* Value

There is a local, or language specific principle in operation in this structure, which is that in Japanese, Heads precede their Markers, unlike English where Markers precede their Heads, and this contrast is shown in examples (3) and (4), where in (3) the marker *o* follows its head *biiru*, but in (4) the marker *for* precedes its head *Salman to go to Iran.*

(3.3) biiru o
beer ACC
"the beer".

(3.4) For Salman to go to Iran is insane.

The consequence of this linear ordering is that in a computational implementation, the phrasal structure for English assumes that the first word is the Marker and the second word is the Head and is usually designated as a Marker-Head structure. The phrasal structure for Japanese assumes the opposite ordering, and is designated a Head-Marker structure. Similarly, complements precede their heads

in Complement-Head structures, specifiers precede their heads in Specifier-Head structures, and adjuncts precede their heads in Adjunct-Head structures.

The verb which heads the entire sentence, illustrated in Figure 3.11, *nonda* (drank) selects two arguments, a subject which is a noun with nominative marking, and a complement which is a noun with accusative marking. Semantically, the subject has the role DRINKER and the complement has the role DRUNK. In the lexicon, these valence specifications are potential, so they are said to be *unsaturated.* When the arguments are expressed as in example (1) the valences are said to be *saturated.* The complement combines with the verb under a Complement-Head structure and the subject combines with the Complement-Head phrase to form a Subject-Head phrase, to complete the parse of the sentence. The configuration of the tree depicted in Figure 3.11[5] is determined by the interaction of other principles in HPSG, such as the Valence Principle and the Empty Complements Constraint. The Valence Principle operates over the features SUBJ, COMPS and SPR to ensure that the head daughter's value for each such feature is cancelled off when combined under Subject-Head, Complement-Head and Specifier-Head phrases respectively. Formally, the principle is:

> **Valence Principle**
> In a headed phrase, for each valence feature F, the F value of the head daughter is the concatenation of the phrase's F value with the list of SYNSEM values of the Fdtrs value.
>
> Pollard and Sag (1994:p.348).

In the figure, the verb *nonda* has two valence potentials, for a subject and a complement, and these are 'cancelled off' as each one becomes saturated. The verb selects for a nominal complement with accusative case marking; this unifies with the Head-Marker structure for *biiru o* (the beer) and under the semantics principle the CONTENT of the complement is unified with the head daughter in the phrasal mother. The resultant Complement-Head phrase has a SYNSEM value with a verbal head having a saturated COMPS list. If the subject is omitted from the phrase, then this phrasal structure retains a SUBJ value which is unsaturated. When combining with the subject, the verb selects a nominal complement with nominative case

[5]Sag (1997) points out that phrasal signs are more faithfully represented as feature structures, not least because they are better able to illustrate structure sharing and multiple inheritance relationships inherent in this framework. This use of phrasal signs is amplified in Section 3.4. however the use of a tree is still a convenient approximation for the current purpose.

marking to form a Subject-Head phrase in which the CONTENT value of the subject is unified with the head daughter under the semantics principle. All values in the Subject-Head phrase are saturated. There is nothing in the Valence Principle to prevent subject and complements saturation in any order, so extra constraints are needed to ensure that argument saturation occurs in the correct order. Sag (1997) assumes the Empty Complements Constraint, paraphrased as:

> **Empty Complements Constraint**
> The head daughter of a headed phrase contains a saturated complements list unless it is an instance of a headed phrase which says otherwise.
>
> Sag (1997:p.10)

The effect of the ECC is to ensure that the subject combines with a saturated Complement-Head phrase, so a sentence such as example (5), although correct Japanese, cannot be accepted by an ALE implementation of the grammar because its subject cannot combine with the head daughter of a transitive verb unless the complement is also present.

(3.5) Oguchi ga ippai nonda.
Oguchi-NOM all drank
"Oguchi drank (it) all".

This system, then, cannot account for languages in which free argument drop occurs. However, argument drop and scrambling are generally difficult problems in language, and for the purposes of this thesis, I require all arguments which can be expressed in the syntax, to be expressed.

Schema	Used Here
Head-Marker	Yes
Subject-Head	Yes
Complement-Head	Yes
Head-Subject-Complement	No
Specifier-Head	No
Adjunct-Head	Yes
Filler-Head	No

Table 3.1: HPSG Schemata

HPSG defines seven phrasal schemata, all of which can be used with English, and these are listed in Table 3.1. The data analysed in this thesis makes use of only

four of these: in addition to the Head-Marker, Complement-Head and Subject-Head schemata, the Adjunct-Head schema is admitted for combining adverbial adjuncts with their verbal heads, such as *ippai nonda* (completely drank) in example (1). Since the publication of *Head Driven Phrase Structure Grammar* in 1994, the SUBCAT list, and its heir, ARG-S has been the nexus of various lines of research. Some of the principal work is reviewed in the next section.

3.3 Notions of Argument Structure

In earlier work on HPSG (Pollard & Sag (1987), Pollard & Sag (1994) chapters 1-8) all the arguments of a predicate are combined together on a single SUBCAT list, ordered according to

> a version of the traditional obliqueness hierarchy, thus, subjects appear first (leftmost) followed by other complements in the order primary object, secondary object, then oblique PP and verbal and/or predicative complements.
>
> Pollard and Sag (1994:p.24)

Arguments are cancelled off under the operation of a phrase structure schema similar to the Subject-Head and Complement-Head schemata discussed in the last section, but with all the valences combined in a single list. The advantage of organizing the arguments of a predicate in this way is that a version of the binding theory can be developed which relies on the ordering principle of obliqueness, leading to a rather simple and straightforward formulation of binding, developed in Pollard & Sag (1994). This is in contrast with the more complex theory developed by Chomsky (1986), which relies on the configuration of the syntactic tree. The SUBCAT list is a syntactic phenomenon, defined as a feature under CAT.

Although HPSG was developed with significant influence from GPSG, the separation of subjects in GPSG was not originally imported. Following Borsley's (1987) arguments that subjects should be kept separate from the SUBCAT list, the Chapter 9 reformulation of Pollard & Sag (1994) reorganized subjects, complements and specifier into separate valences, as introduced above. Rather than redefine the HPSG binding theory over the separate valences, Pollard and Sag chose to retain the SUBCAT list as the canonical append of SUBJ, COMPS and SPR valences so that the obliqueness based binding theory is retained. In later work, it is claimed that this formulation of SUBCAT "has become similar to certain notions of argument

structure" (Manning and Sag (1994:p.3)), and the feature has been renamed as ARG-S. The traditional, published, view of argument structure is as an ordered list of semantic roles, exemplified by the work of Grimshaw (1990) and in this work argument structure is not a syntactic feature at all, but it is the locus of linking between syntax and semantics.

In the subsequent development of the HPSG framework, there seems to have been a split in the view of what is the role of ARG-S in a lexical item. Wechsler (1995) develops a theory of linking between semantic roles and syntactic arguments, and he introduces a ROLES list as the locus of linking. This ROLES list he later renames ARG-S, and Wechsler's view of argument structure is not inconsistent with the traditional view represented by Grimshaw (1990). In a syntactic account of *ergativity*, Manning (1994) develops an account of argument structure as the locus of binding theory, without concern for linking. The works of Wechsler and Manning represent the divided view of argument structure in HPSG. In this section, the HPSG binding theory is introduced in 3.3.1, and the accounts of Wechsler and of Manning are reviewed in subsections 3.3.2 and 3.3.3 respectively.

3.3.1 HPSG Binding Theory

The variety of nominal objects includes pronominal or nonpronominal objects and the pronominals include anaphors and personal pronouns. Finally, the anaphors are divided into reflexives and reciprocals. These are organized into a type hierarchy, depicted in Figure 3.13

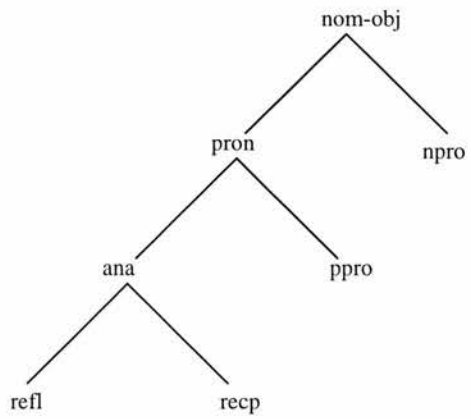


Figure 3.13: The Type Hierarchy of Nominal Objects

If there are two SYNSEM objects Y and Z on the SUBCAT list, then Y locally o-commands Z if Y is the less oblique element on the list; more generally, if Y o-commands X which o-commands Z, then Y o-commands Z. If in addition Y and Z are co-indexed then Y o-binds Z[6].

HPSG Binding Theory

Principle A: A locally o-commanded anaphor must be locally o-bound,
Principle B: a personal pronoun must be locally o-free,
Principle C: a non-pronoun must be o-free.

(Pollard & Sag 1994).

Binding theory explains how anaphors are coindexed with their antecedents, and thus how Examples (6) and (7) are correct, while (9) is not.

(3.6) (a) John showed Mary a picture of himself.
(b) $\langle John_i, Mary_j, picture, himself_i \rangle$

(3.7) (a) John showed Mary a picture of herself.
(b) $\langle John_i, Mary_j, picture, herself_j \rangle$

(3.8) (a) *John thinks Mary admires himself.
(b) $\langle John_i, \langle Mary_j, himself_i \rangle \rangle$

(3.9) (a) John thinks Mary admires herself.
(b) $\langle John_i, \langle Mary_j, herself_i \rangle \rangle$

In Examples (6) and (7) there is a single SUBCAT list associated with the ditransitive verb 'show' with ordering $\langle John_i, Mary_j, picture, himself_i \rangle$ according to Pollard & Sag's (1994) obliqueness ordering reported above. In this case both John and Mary locally o-command the preposition phrase 'of himself', so in either case co-indexing

[6]There is an extra principle which has been developed to account for long-distance binding phenomena that occur in languages such as Chinese, Japanese and Korean, and this is named Principle Z.

Principle Z: a long-distance anaphor must be o-bound.

(Xue *et al.* 1994).

Jibun is a Japanese reflexive anaphor underspecified for person, roughly equivalent to "one's", rather than "his" or "hers". In the ditransitive example here, *jibun* coindexes with the subject *Taroo* and not with the oblique *Keiko*. The SUBCAT list shows the obliqueness relations.

- Taroo_i ga Keiko_j ni $\text{jibun}_{i,/*j}$ no hon o miseta.
  Taroo-NOM Keiko-DAT self-GEN book-ACC show-PAST
  "Taroo_i showed his$_i$ own book to Keiko_j".
- subcat$\langle Taroo_i, Keiko_j, hon, jibun_{i,*j} \rangle$

is permissible and these examples are correct. In (8) and (9) "Mary admires self" is a subordinate predicate to 'thinks', with its own SUBCAT list. While John o-commands 'self' in (6), Mary locally o-commands it and therefore o-binds 'herself' in (9), and in virtue of this local o-command relation, John does not o-bind 'himself' in (8).

As it stands, retaining the SUBCAT list as part of the theory seems to be simply an expedient to avoid having to reformulate the binding theory over the separate valence lists, as is pointed out in Manning & Sag (1994) although other work, including that of Manning (1994) argues for independent motivation of a separate SUBCAT list, renamed as ARG-S, following the influence of Lexical-Functional Grammar (Bresnan (1982)). Manning argues for a separated level of argument structure, which he calls a *syntacticized* arg-s, to account for the difference in linkage patterns between ergative and accusative languages. He has little to say about the mapping between ARG-S and semantic roles, however. In contrast, Wechsler develops ideas of linkage between SUBCAT (now ARG-S) and semantic roles, in a theory where the ordering of arguments on the ROLES list follows a set of general semantic principles which he proposes.

3.3.2 Wechsler (1991/1995)

In their introduction to the verbal sign, Pollard & Sag (1994) show that the semantic roles are coindexed with items on the SUBCAT list. The SUBCAT list is ordered according to obliqueness principles quoted in the previous section but the ordering of semantic roles is not considered. Wechsler notes the SUBCAT ordering and suggests, following Crimmins & Perry (1989), that in the ditransitive 'donate' relation there is a surface syntactic ordering, and an underlying semantic ordering, shown in (10) and (11) respectively, and consequently there is a need to find this ordering and to find the principles of linkage between the ordered roles and the elements of SUBCAT.

(3.10) John donated his books to the library.

(3.11) DONATE(giver:John, recipient:the library, gift:John's books).

The surface arguments and the semantic roles are organized into the SUBCAT list and the ROLES list, respectively in Figure 3.14.

The ROLES list is equivalent to argument structure of e.g. Bresnan & Kanerva (1989), or Grimshaw (1990), and ordering is organized according to a Thematic Hierarchy in this tradition. Although many writers have tried to explain semantic

$$\begin{bmatrix} \text{PHON} & \textit{donate} \\ \text{SUBCAT} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{3}, \text{PP(TO):}\boxed{2} \rangle \\ \text{ROLES} & \langle \boxed{1}\text{GIVEN}, \boxed{2}\text{RECIPIENT}, \boxed{3}\text{GIFT} \rangle \end{bmatrix}$$

Figure 3.14: *Subcat* and *Roles* Lists

ordering, developing various ideas about the contents and ordering of such hierarchies, no overall consensus has emerged. Wechsler hypothesizes that "the appropriate semantic basis involves primitives of an even more abstract and general sort, and that they are very few in number" (Wechsler 1995, p.3). Figure 3.15 illustrates the problem. In this figure, the "principles of SUBCAT structure" are the obliqueness principles introduced earlier, the "principles of linking" and the "principles of argument structure" are the concern of Wechsler's thesis.

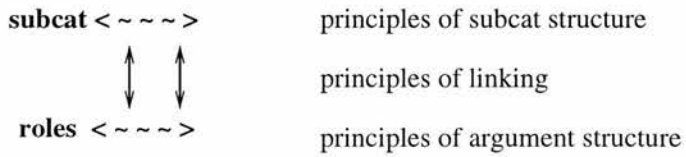


Figure 3.15: Ordering Principles and Mappings

Bresnan & Kanerva (1989) propose that linking in transitive and ditransitive verbs is constrained by an Isomorphy Condition in which "the mapping between unrestricted roles and complements is an order isomorphism, i.e. lines of association between unrestricted roles and the complements filling them cannot cross[7]". Wechsler adopts this as his principle of linking. For the principles of argument structure he develops a set of rules called **The Notion Rule**, **The Nuclear Role Rule** and **The Part Rule** arguing that these are sufficient to determine the semantic ordering rules for transitive verbs. In my view these rules are paraphrases of certain conditions holding on aspectual classes, and they actually presuppose a more limited set of semantic roles than the traditional Thematic Hierarchy. I will discuss each rule in turn starting with the Notion Rule.

The Notion Rule

A lexical sign meeting this description is ill-formed:

- Rel **R**
- ROLES $\langle ..role_1..role_2.. \rangle$

if the following holds:

[7]Roughly, restricted roles are oblique and unrestricted roles are direct, nominatives and accusatives.

- $\forall x, y \Box[R(Role_1 : y, Role_2 : x) \longrightarrow conceive(x, y)] \wedge$
  $\neg\forall x, y \Box[R(Role_1 : x, Role_2 : y) \longrightarrow conceive(y, x)]$.

Paraphrasing in words this is "In a lexical sign with two arguments x and y, the roles are ordered (x,y) if x has a notion of y but not necessarily y has a notion of x".

The verb 'like' takes two syntactic arguments, X and Y, and has two semantic roles, LIKER and LIKEE. The contention is that if X has a notion of Y, then X is designated as the LIKER and ordered first on the ROLES list. To be sure it is not unambiguously possible to designate John as the subject in likes(John,Mary) if Mary also likes John. It is more clear that the LIKER and LIKEE roles are so ordered if a non-animate object is substituted as the LIKEE. Thus, in likes(John,ice-cream) a situation where the ice-cream likes John is never possible, so with the verb *like*, the liker will always be mapped to subject and the likee to object. Wechsler's examples of verbs subject to the Notion Rule include a variety of verbs which as it turns out are Activities and Accomplishments: see, hear, touch, smell, murder, chase, flee; and also some psychological verbs: want, like, fear, expect, which are, in English, Stative verbs. Tests for Stativity are incompatibility with the progressive, and incompatibility with purposive adverbs.

(3.12) I want an ice cream / *I am wanting an ice cream.
(3.13) I like Mary / *I am liking Mary.
(3.14) I expect Thabo / I am expecting Thabo.

Of these examples, (12) and (13) are clearly Stative, while expect in (14) is completely compatible with the progressive. Stative verbs are not compatible with purposive adverbs, and the purposefulness tests (15) and (16) are decisive however.

(3.15) *I deliberately want an ice cream.
(3.16) *I expect Thabo on purpose.

The SUBCAT and ROLES lists are organized as in Figure 3.16, the isomorphy condition holds, so the mapping between subject and LIKER, object and LIKEE applies.

$$\begin{bmatrix} \text{SUBCAT} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2} \rangle \\ \text{ROLES} & \langle \boxed{1}\text{LIKER}, \boxed{2}\text{LIKEE}, \rangle \end{bmatrix}$$

Figure 3.16: Isomorphic Mapping

Wechsler's claim is that the Notion Rule is a sort of general test for *agentivity* and this fits the ideas of e.g. Grimshaw (1990) and Dowty (1991) that agents are always realized as the subject. There are, however, large classes of verbs for which the Notion Rule is inconclusive, including pure Achievement, and *causative/unaccusative* verbs. Consider the predicate crush(x,y). The Notion Rule has no bearing on the projection of CRUSHER role to subject since the CRUSHER need not have a notion of the CRUSHEE, as in Example (17) where the rock has no notion of Michael, although the rule "applies" to the arrangement of arguments given in (18). For a valid application of the rule, it has to apply generally to any possible argument of the verb. The verb 'absorb' is an example of a verb which is not obviously agentive, so the Notion Rule is of no use in correlating ABSORBER and subject in Example (19).

(3.17) The rock crushed Michael.
(3.18) Michael crushed the rock.
(3.19) The sponge absorbed the water.

For cases such as these Wechsler introduces the Nuclear Role Rule, where the focus is on the affectedness of the object.

The Nuclear Role Rule
A lexical sign meeting this description is ill-formed:

- Rel **R**
- ROLES $\langle ..role_1..role_2..\rangle$

where Role1 is $+$ *nuclear* and Role2 is $-$*nuclear*

Wechsler's intuition is that in an example such as (19), the water is the more affected by the situation, is therefore $+$*nuclear* and projected to object. In the *crush* verbs the subject role is underspecified for agentivity, but the object is the more affected and so it is designated $+$*nuclear* and linked with the CRUSHEE role. In case (18) the rock is powdered, therefore $+$*nuclear* while Michael is unaffected. In his illustration of these examples Wechsler reverts to traditional thematic roles of Theme and Agent, with the Agent designated as $-$*nuclear* and the Theme designated as $+$*nuclear*. Themes are always $+$*nuclear*, and always linked to object in transitive verbs. The diagnostic use of the verbal prefix *re-* is offered as an aid to determining $+/-$ *nuclear'hood'* as in (20), which presumes the door was open before, but not necessarily that John opened it. This diagnostic use of the *re-* prefix is by no means generally applicable however. Consider the bizarreness of 'redissolving', in (21).

(3.20) John reopened the door.

(3.21) ??The acid redissolved the metal.

The *+nuclear* distinction is the more relevant to arguments of *telic* verbs - Achievements and Accomplishments. With an Activity like 'ride', if Magnus rides a bike, there is no obvious effect on the bike afterwards, so it is not *+nuclear*. However, this case, and Activities generally, are covered by the Notion Rule such that Magnus has a notion of the bike, and the 'rider' is projected to subject. The problem with the Nuclear Role Rule is that there are no clear principles, or tests (excepting the *re-* prefix test, which is not generally applicable), to help in determining which role is the more affected by the verb, and so which role should attract the feature designation *+nuclear*.

Wechsler uses these two rules, the Notion Rule and the Nuclear Role Rule, for diagnosing role-ordering in what turn out to be the *dynamic* verbs: Activities, Achievements and Accomplishments. The Notion Rule is useful in English for certain Stative verbs, but as it turns out none of these are Stative verbs in Japanese: 'want' is either a bound morpheme, *-tai*, or a predicate adjective *hoshii*, 'like' is a predicate adjective *suki*, and 'expect' is most often used as a phrasal postposition *yoki suru* which may take an accusative argument, suggesting that it is an Activity. Some pure Stative verbs are handled in Wechsler's scheme by the Part Rule, in which a verb expresses a Whole-Part relationship where the Whole must be expressed as subject while the Part is expressed as object.

The Part Rule

A lexical sign meeting this description is ill-formed:

- Rel **R**
- ROLES $\langle..role_1..role_2..\rangle$

where the following lexical entailment holds:

- $\forall x, y \Box R(Role_1 : y, Role_2 : x) \longrightarrow Part(Whole : x, Part : y)$

Examples (22) and (23) are used to illustrate. In (22) the appendix, as Part, cannot be presented as Subject, while in (23) the term VP, as Whole, must be projected as subject.

(3.22) The book includes an appendix.

(3.23) The VP dominates the NP.

The verb 'dominate' (Wechsler's examples) is, perhaps, not a particularly good choice of example as it is underspecified for agentivity. In (24) 'The Earl' certainly has a notion of his vassals and the verb is an Activity, compatible with the progressive in (25), while as a non-agentive it is incompatible with the progressive, and therefore a State, as in (26).

(3.24) The earl dominates his vassals.
(3.25) The earl is dominating his vassals.
(3.26) ??The VP is dominating the NP.

Not all Stative verbs are of the Whole-Part variety. Indeed, the verb *to be* may have a locative interpretation (27) or may be used as an equative (28).

(3.27) The tree is in the garden.
(3.28) His lordship is the scoundrel.

It may be possible to treat locative relationships analogously with Part-Whole relationships, with the FIGURE as PART, but the equative case is not accounted for by any of Wechsler's three rules.

In Chapter 4 of his (1995) book, Wechsler updates his arguments for mapping between SUBCAT and ROLES lists to account for the HPSG separation into valences. He suggests that:

> The list we call ROLES corresponds roughly to a list called SUBCAT or ARG-S (mnemonic for "argument structure") in other work within this revised model. We retain the term ROLES for consistency with previous chapters of this book. The old mapping between the ROLES and SUBCAT lists is now replaced with a mapping between the ROLES list and the VALENCE lists, SUBJ and COMPS".
>
> Wechsler (1995) p.132-133.

If we go back to Figure 3.15, the SUBCAT list is the list of syntactic elements and is also the locus of binding, ordered according to traditional obliqueness principles. Adding the valences gives an extra level of syntactic linking, in Figure 3.17(a)

What Wechsler is suggesting in the passage quoted above is that the ROLES list replaces the SUBCAT list, now ARG-S, so there is a mapping from ROLES to valences,

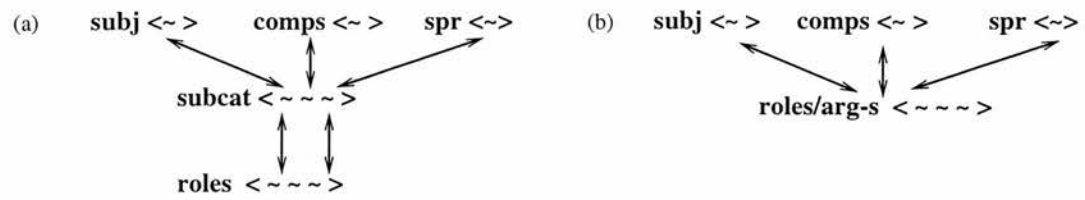


Figure 3.17: Ordering Principles and Mappings

shown in 3.17(b). This means there is a conflict in ordering of the ROLES/ARG-S list, which is either ordered by obliqueness following Pollard & Sag (1994), or ordered according to the Notion, Nuclear Role and Part rules. Wechsler's thesis is primarily concerned with developing a linking theory, to resolve the ordering of semantic roles of a predicate with the surface syntactic order of arguments, and his theory does not address the constraints imposed by binding.

3.3.3 Manning (1994)

In his thesis on *ergativity*, Manning develops an analysis of Tagalog, Inuit, Balinese, Toba Batak and other syntactically ergative languages. The essential point about ergativity is that it displays a case marking pattern systematically different from that of an *accusative* language. In the ergative case intransitive verb subjects and transitive verb objects are case marked the same, while the argument which corresponds with the subject is marked in an *ergative* case. His examples from West Greenlandic (Inuit) follow.

(3.29) Oli-p neqi neri-vaa.
Oli-ERG meat-ABS eat-IND.TR.3sg.3sg
"Oli eats meat".

(3.30) Oli sinippoq.
Oli-ABS sleep-IND.INTR.3sg
"Oli sleeps".

These are contrasted with examples from Japanese, which in spite of its free word ordering is a syntactically accusative language, and subjects are case-marked with nominative *ga* (except when topicalized using *wa* or *mo*).

(3.31) kisya ga tuita.
steam train-NOM arrive-PAST
"The steam train arrived".

(3.32) Orinoko-san ga kisya o orita.
Orinoko-NOM steam train-ACC alight-PAST
"Orinoko got off the steam train".

So where an *accusative* language has nominative and accusative cases, an *ergative* language has ergative and absolutive cases. The relationship between them is explained by Dixon's (1979) mapping through "A, S and O" arguments where A is the agent-like argument of a transitive verb, S is the single actant of intransitive verbs, and O is the patient-like argument of a transitive verb as illustrated by Figure 3.18.

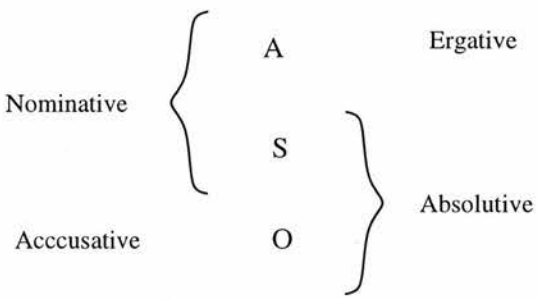


Figure 3.18: Ergative/Accusative Case Marking

In the reformulation of HPSG in Pollard & Sag (1994), Chapter 9, there is direct linkage between grammatical arguments, expressed as valences SUBJ, COMPS and SPR, and semantic roles. The SUBCAT list is retained strictly for ease of formulation of the binding theory. The relationships within ergative languages are inverted relative to accusative languages. Manning suggests that the SUBCAT/ARG-S list ordering is invariant with respect to ergative and accusative languages, and it is just the syntactic linkage between ARG-S and the valences that is inverted, as shown in his figure for accusative (3.19(a)) and ergative (3.19(b)) mapping.

Manning's thesis does not address the issues of mapping to semantic roles, although his tacit assumption is that role to ARG-S mapping is the same for ergative and accusative languages. A related paper by Manning & Sag (1999) does address this mapping. In that paper they show the progression in HPSG from the SUBCAT/semantic roles mapping model to the ARG-S/semantic roles model, and yet they claim that ARG-S is a syntactic representation, and should not be used as a substitute for a semantic representation. The latter set of mappings for (a) accusative languages and (b) ergative languages is shown as an extension to Manning's ergative/accusative linking figure, in Figure 3.20.

It is hard to see what Manning & Sag (1999) are trying to protect by protesting that ARG-S is a syntactic representation. The fact is, there is a need to represent both

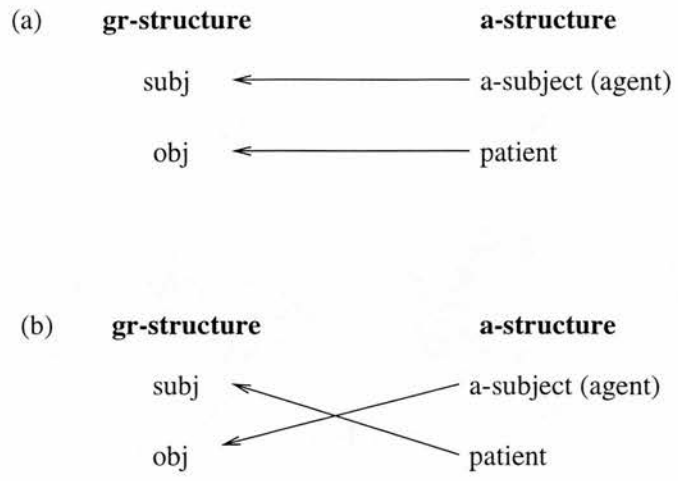


Figure 3.19: Accusative (a) and Ergative (b) Mappings

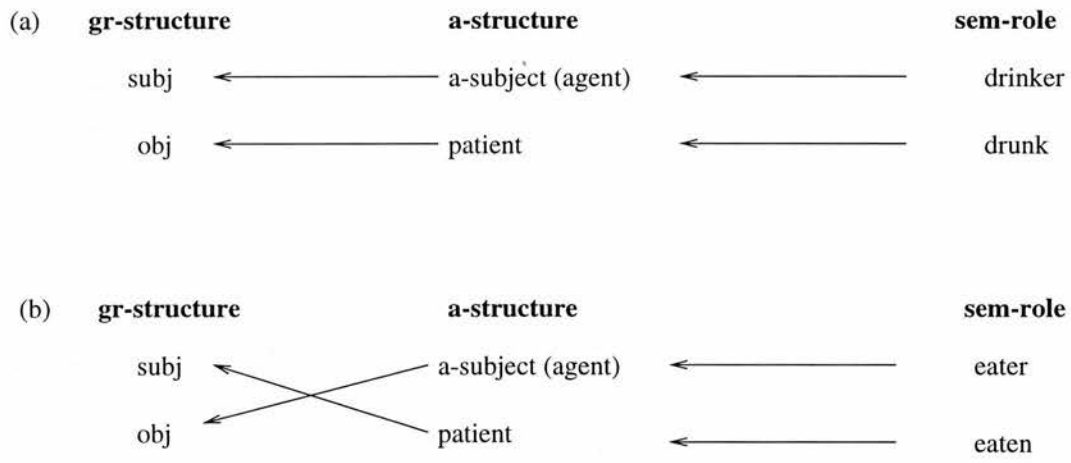


Figure 3.20: Syntax-*arg-s* - Semantics Mapping

syntax and semantics and to establish a link between them, and the ARG-S list is the locus of this mapping, as demonstrated by Manning & Sag (1999) and at length by Wechsler (1995). The fact that this list is also the locus of the HPSG binding theory in no way invalidates its crucial role at the syntax/semantics interface within lexical items. Manning (1994) claims that two principles govern the ordering of arguments on ARG-S:

1. "direct arguments precede obliques", and
2. "within each of the direct and oblique arguments, arguments are ordered according to thematic obliqueness. The thematic hierarchy of Bresnan & Kanerva (1989) will be sufficient:

   Ag ⟩ Ben ⟩ Recip/Goal/Exp ⟩ Instr ⟩ Thm/Pat ⟩ Loc

   Manning (1994):pp.42-43

Manning's first principle is a reflection of the traditional obliqueness hierarchy and his second principle is a bow in the direction of traditional thematic principles à la Gruber (1976), Jackendoff (1976), Grimshaw (1990). He reports these two principles as being motivated by Hellan (1988) and supported by Dalrymple (1993).

It is notable that Bresnan & Kanerva (1989) predates both Wechsler's and Manning's theories of role ordering, and yet they draw different inferences from it. While Manning retains the primacy of the obliqueness ordering, because his concerns include the binding theory, Wechsler focuses on doing away with thematic hierarchies and replacing them with a minimal set of principles, the Notion, Nuclear

Role and Part rules. Assuming that he does not seek to discard his ordering rules, what Wechsler seems to be doing with this move is ignoring the SUBCAT ordering of Pollard & Sag (1994) which was designed to work with the binding theory. Manning respects this ordering but reintroduces the thematic hierarchy which Wechsler had carefully argued away. In order to respect both syntactic role ordering, for the binding theory, and semantic role ordering, including the minimized set of semantic roles, a synthesis of the two theories of role ordering is called for, which requires:

1. that direct arguments precede obliques (following Manning (1994)),
2. that arguments are ordered within these two sublists according to the Notion, Nuclear Role and Part rules (following Wechsler (1995)).

A third theory of linking, and that which is adopted in current HPSG (see Sag (1997), reviewed below) is that developed in Davis (1995). Davis develops a rather highly constrained system using multiple inheritance, which constrains the relationships between semantic roles in the NUCLEUS and positions in argument structure (ARG-S): though the motivation for the semantic roles is ultimately based on Dowty's (1991) work on proto-roles.

3.4 Davis (1995) and Semantic Relations

Various authors since Gruber (1976) have tried to develop the notion that the arguments of a predicate have distinct semantic roles associated with them, and these roles have an order which does not always accord with the surface syntactic order. Jackendoff (1972), Foley & Van Valin (1984), Bresnan & Kanerva (1989) (as above, p.74) and Grimshaw (1990) all motivate (or 'assume') distinct roles and orderings, for example Grimshaw suggests a thematic hierarchy in which Agents are

most prominent, followed by Experiencers and the Locative roles of Goal or Source or Location, and finally Theme.

> Agent ⟩ Experiencer ⟩ Goal/Source/Loc ⟩ Theme
>
> Grimshaw (1990:p.8)

The idea is that whenever a predicate has any two or three of these roles, the role which is more prominent (or less oblique) in the hierarchy is lexicalized as the subject, and the more oblique roles are lexicalized as objects. A predicate such as 'murder' has Agent and Theme roles, and the Agent is subject while Theme is direct object. The trouble is that there has been little agreement of the numbers, types and orderings of roles in the Thematic Hierarchy, and in his paper on semantic roles Dowty (1991) develops an alternative in which what he calls the *proto-agent* and *proto-patient* roles are determined by summing sets of entailments and the associated properties for each proto-role are:

> **Proto Agent Properties**
>
> (a) Volitional involvement in the event or state,
> (b) Sentience (and/or perception),
> (c) Causing an event or change of state in another participant,
> (d) Movement (relative to the position of another participant),
> (e) (Exists independently of the event named by the verb).
>
> **Proto Patient Properties**
>
> (a) Undergoes change of state,
> (b) Incremental Theme,
> (c) Causally affected by another participant,
> (d) Stationary relative to movement of another participant,
> (e) (Does not exist independently of the event, or not at all).
>
> Dowty (1991:p.572)

For a predicate, the argument with the greatest number of proto agent entailments is lexicalized as the subject, and the argument with the greatest number of proto patient entailments is lexicalized as object. The predicate 'murder', for example, has arguments one of which is (a) volitionally involved, (b) sentient and (c) causing a change of state in the other, and this 'murderer' role is lexicalized as the subject. The other role (a) undergoes a change of state, and (c) is causally affected by the other participant, and so is lexicalized as the object. Not all examples are quite so clear cut, and Dowty concentrates on transitive verbs, and moreover he counsels

against reifying the entailments in a fixed set of labelled roles, claiming that no such set of roles exists. Notwithstanding Dowty's counsel, Davis's (1995) identifies roles ACT and UND (and other role labels) based on Dowty's proto-roles. His starting point is to note that the relations used by Pollard & Sag (1994) fail to constrain the roles that can appear in a relation, so there is nothing to stop the transitive relation 'devour' from containing DEVOURER, DEVOURED and SLEEPER roles: in short, the content type hierarchy misses a major generalization in failing to identify intransitive, transitive and ditransitive verbs as separate relations with one, two or three roles, respectively. In the previous section, the review of Wechsler (1995) shows the development of a linking theory based on just three primitive semantic rules, the Notion, Nuclear Role and Part rules. Davis' criticism is that Wechsler's system does not provide universal coverage of all predicates, and moreover he is only concerned with the ordering of 'direct' NP arguments, and the status of obliques is not addressed. Davis' solution is to lexicalize the theory of Dowty (1991), using Foley & Van Valin's (1984) macrorole labels of ACT and UND as realizations of proto agent and proto patient roles, and a small set of extensions to realize oblique roles.

Under this model, unergative verbs such as 'dance' have a single argument with the role ACT and unaccusatives such as 'die' have a single argument with the role UND. A transitive verb inherits both roles ACT and UND. Davis makes no distinction between Activities and Accomplishments, and the undergoer role is used for the object in both cases. The possible co-occurrences of these roles in a semantic relation under CONT are organized in a relation type hierarchy such as Figure 3.21, which constrains the values of the CONT feature (Davis ignores quantificational matters).

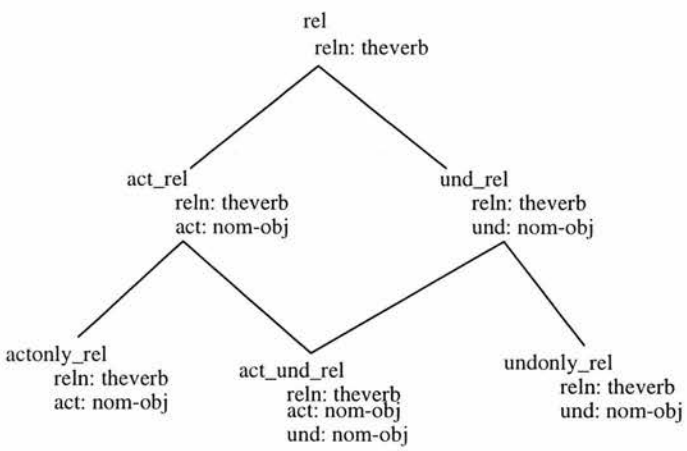


Figure 3.21: Davis' (1995) Relation Hierarchy

The *rel* type hierarchy introduces the relation (RELN) feature with a value which is the specification of the unique verb. Subtypes of *rel* inherit this feature and the *act-rel* introduces an actor (ACT) feature identifying a nominal-object, while *und-rel* introduces the undergoer (UND) feature. Unergatives are identified as *actonly-rel* and are a subtype of *act-rel* and unaccusatives are identified as *undonly-rel* a subtype of *und-rel*, and neither of these sorts adds any new features; *act-und-rel* is the sort which inherits features from both its parents, thus RELN and ACT from *act-rel* and (the same) RELN and UND from *und-rel*. The roles in these relations are linked together with the elements of ARG-S in a linking sort.

(a):
$$\begin{bmatrix} \text{ARG-S} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2} \rangle \\ \text{CONT} & \begin{bmatrix} \text{RELN} & \textit{the.verb} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \end{bmatrix} \\ & \textit{act-und-rel} \end{bmatrix}$$

(b):
$$\begin{bmatrix} \text{ARG-S} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2} \rangle \\ \text{CONT} & \begin{bmatrix} \text{NUC} & \begin{bmatrix} \text{RELN} & \textit{the.verb} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.22: (a) *act-und-ls*, and (b) in the *nucleus*

Davis defines a set of linking sorts containing the macroroles, and in which the mapping between the roles and the ordered elements of the ARG-S list are constrained by inheritance relations. The linking sort *act-ls*, for instance, links the ACT role with the top-level element of the ARG-S, while the linking sort *und-ls* links the UND role with *any* element on the ARG-S. The sort *act-und-ls*, illustrated in Figure 3.22(a), which relates to transitive verbs is a subtype of both *act-ls* and *und-ls* and is therefore tightly constrained as to the linking of its roles. Davis' theory makes no use of quantificational content, and his entire concern with the CONTENT field (of verbs) is with the features of the relational NUCLEUS. In fact, since Davis' concern for the CONTENT value is entirely devoted to the subfeatures of the NUCLEUS, he omits the intervening NUCLEUS label and defines RELN, ACT and UND directly as features of CONTENT. Figure 3.22(b) shows the *act-und* relation under NUCLEUS rather than CONTENT. In common with a tradition in HPSG exposition, Davis' omits higher levels of feature path in his explanation, and while ARG-S is shown independently in the figure, it is actually embedded under CAT, so for every linking sort which creates a constraint on ARG-S, that sort must also contain a unique subtype of *cat* with respect to the relation types under CONT. This dependence is broken in Sag & Wasow (1999), where CAT, ARG-S and CONT features are separated out under SYNSEM, and I adopt their convention.

In Davis' linking theory a ditransitive relation involves not only the addition of an argument linked to a third element of ARG-S, but it creates a relation between the

two oblique arguments, with a variety of structures possible including *cause-effect*, *cause-means* and *figure-ground* relations. As an example, the *cause-effect* relation in Figure 3.23 represents the semantics of example (33) in which the relation of actor and undergoer exists between Tanko and her bracelet, and the relation of figure and ground exists between the bracelet and Yasuko, and these are related together as the value of the EFFECT role in a ditransitive, containing figure FIG and ground GRND roles, using the sense of figure and ground as introduced by Talmy (1978).

(3.33) Tanko showed Yasuko her bracelet.

$$\begin{bmatrix} \text{ARG-S} & \langle \text{NP}:\boxed{1}, \text{NP}:\boxed{3}, \text{NP}:\boxed{2} \rangle \\ \text{CONT} & \begin{bmatrix} \textit{ditrans-rel} & \begin{bmatrix} \textit{cause-rel} & \begin{bmatrix} \text{RELN} & \textit{ditrans.verb} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \\ \text{EFFECT} & \begin{bmatrix} \text{FIG} & \boxed{2} \\ \text{GRND} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.23: Ditransitive Relation inside *Nucleus*

Rather than presume that such a relation exists for the Japanese verbs under study, I will use Pollard & Sag's (1994) earlier characterization of ditransitive relations, which assumes that the three roles are all equally features of the NUCLEUS. Relation types are developed for the range of verb frames given in the last chapter. Davis (1995) lexicalizes Dowty's (1991) proto role entailments to identify the linking between semantic roles and syntactic arguments, and application of the same principles to these Japanese verb frames leads to the argument-role mapping summarized in Table 3.2. These verbs are organized into a semantic relations type hierarchy, with *rel* types representing the contents of the associated verbal nuclei, shown in Figure 3.25.

With a transitive relation, the actor is linked to subject, or nominative, and the undergoer is linked to object or accusative argument, in the relation *act-und-rel.* In the case of a one place relation, the role may be actor (in unergative Activities), in the *act-rel* relation, or undergoer (in unaccusative Achievements), in the *und-rel* relation, and this argument is in either case always linked to the subject role. The locative in intransitive relations such as States or Achievements, or ditransitive relations such as Accomplishments, is linked to the dative role.

Aspect	Case Marking	Syntactic Frame	Semantic Roles
States			
	ni,ga	dat,nom	loc,und
	ga,ga	nom,nom	loc,und
Achievements			
	ga	nom	und
	ga,ni	nom,dat	und,loc
Activities			
	ga	nom	act
	ga,o	nom,acc	act,und
Accomplishments			
	ga,ni	nom,loc	act,loc
	ga,o	nom,acc	act,und
	ga,o,ni	nom,acc,loc	act,und,loc

Table 3.2: Role-Argument Mappings

(a): $\begin{bmatrix} \text{ARG-S} & \langle \text{NP}:\boxed{1} \rangle \\ \text{CONT } \textit{index}[\text{ROLE } \boxed{1}] \end{bmatrix}$ (b): $\begin{bmatrix} \text{ARG-S} & \langle \text{NP}:\boxed{1} \rangle \\ \text{CONT } \textit{nom-obj}[\text{ROLE } \boxed{1}] \end{bmatrix}$

Figure 3.24: Linking Nominal Arguments in (a) Pollard and Sag and (b) Davis.

A crucial change which Davis makes is to structure share nominal objects by their entire semantic content values, rather than an INDEX linked to agreement features, which is Pollard and Sag's approach. The difference between the approaches is illustrated in Figure 3.24, with the INDEX linked approach in 3.24(a) and the *nom-obj* linking approach in 3.24(b). While *nom-obj* is a subtype of content, INDEX is a feature of *nom-obj*, in English, containing the agreement features. The ROLE feature represents either ACT, UND or LOC.

Other structures of Pollard & Sag (1994) are already structure shared by CONTENT value, for example the SOA-ARG in a control verb such as 'persuade' or 'want' has a SOA-ARG value which is a non-finite VP. Davis offers three reasons for rationalizing the linking of nominals: (i) the fact mentioned above that this brings the subcategorization of NPs, CPs and VPs into line; (ii) reflexives are semantically monadic in languages such as Dutch, but dyadic in others, such as English, and sharing of content values allows this distinction to be drawn more simply; and (iii) an account of role-argument uniqueness becomes possible. Davis' proposal is in line with several others, including that of Dowty & Jacobson (1989). Pollard & Sag (1994) on the other hand are proposing a unique argument to explain the facts of English agreement; one thing that their system does not allow, though, is any notion of

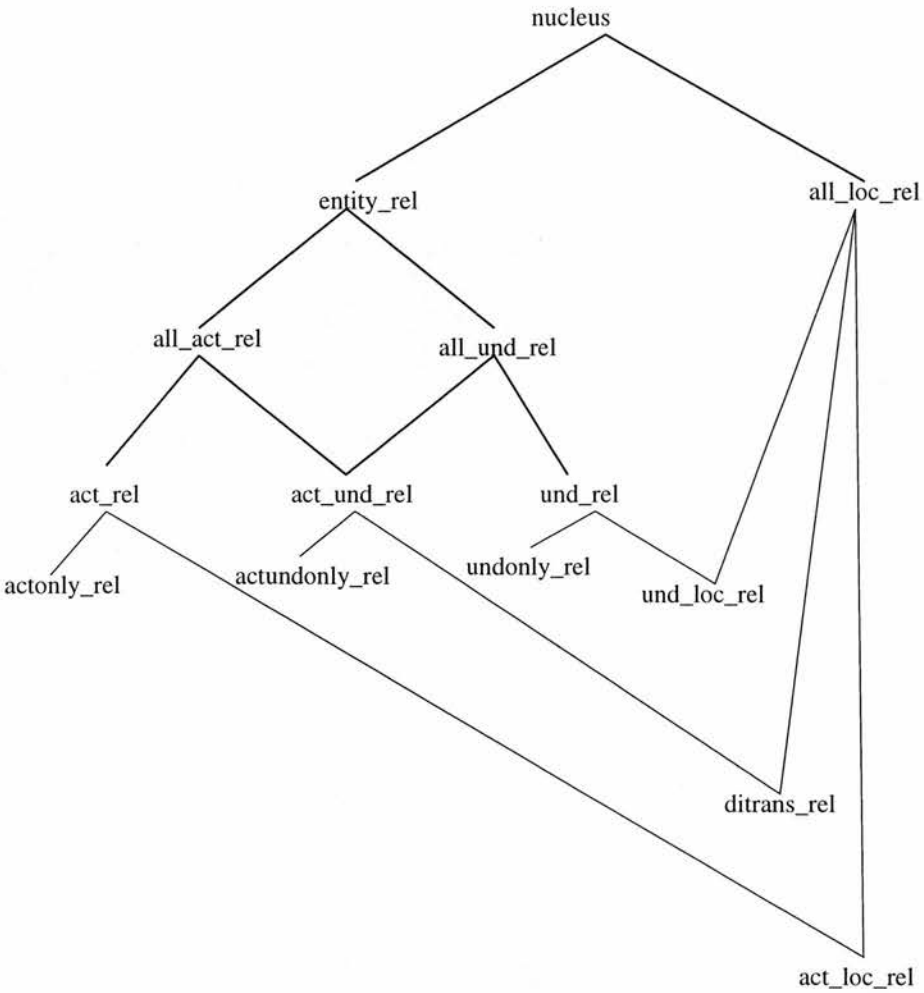


Figure 3.25: The Relation Hierarchy based on (Davis 1995)

polysemy. If a verb such as 'want' takes an NP 'wantee' argument, or in its control form a SOA-ARG, then there are necessarily two separate lexical entries. Pollard & Sag's (1994) analysis of English control verbs is reviewed in the next chapter. By allowing the subcategorization of content underspecified for phrasal type under a single common role label such as UND, Davis' proposal opens the door to a polysemous account of control verbs, and this is necessary for the present account.

Pollard & Sag (1994) distinguish phrasal signs from lexical signs by their Head and Non-Head Daughter features, but subsequent work in HPSG has masked this phrasal configuration by continuing to use tree-based depictions of a parse. Sag (1997) further develops the phrasal type hierarchy in such a way that universal principles constrain phrasal features through the type hierarchy. This approach 'loses information' in a tree-based representation, so the complex phrasal sign is urged as the preferred method.

3.5 Sag (1997) and Phrasal Structure

In Sag (1997) a phrasal type hierarchy is developed in which phrases are cross-classified for clausality and headedness, Head Phrases include the Subject-Head, Head-Complement, Head-Specifier, Adjunct-Head and Filler-Head phrases. Clausal types for this data include (English) declarative clauses, relative clauses and interrogative clauses, all varieties of finite-Subject-Head phrase. The English data is not interesting for the current thesis, but the reorganization of headed phrases, and the formalization of phrasal types, is highly relevant to my analysis and it is principally this part of Sag's machinery I will concentrate on here.

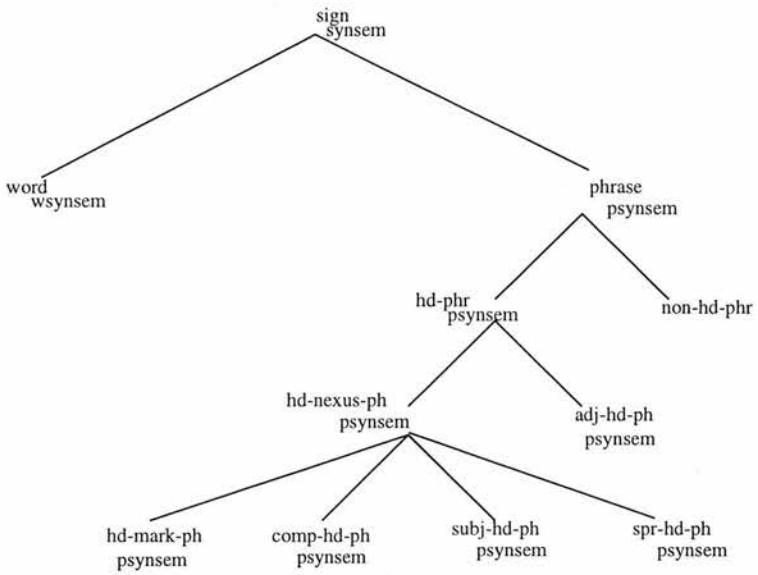


Figure 3.26: The *sign* Hierarchy (Sag 1997, amended)

3.5.1 Phrasal Signs

Phrasal signs are organized according to a hierarchy of features and so are also amenable to the multiple inheritance techniques used for structuring lexical signs. Sag's hierarchy of phrasal types is given in Figure 3.26, adapted for Japanese linear precedence and for SYNSEM type. With this hierarchical organization, various principles and constraints previously stipulated as part of the theory are now expressed as features of a supertype, inherited by its subtypes. Phrases are divided into headed and non-headed phrasal types. There is no current application for non-headed phrases in HPSG. Headed phrases include Head-Nexus phrases and Adjunct-Head phrases. We saw in Section 3.2 that phrasal types are subject to a set of universal principles such as the *Head Feature Principle*, the *Semantics Principle*, the *Valence Principle*, the *Empty Complements Constraint* and the *Marking*

Principle; these principles are represented in Sag's hierarchy as constraints between the phrasal SYNSEM value and the value of the appropriate daughter feature. All headed phrases are subject to the *Head Feature Principle*, for instance, which requires that the HEAD features of the phrasal sign and the HEAD features of the Head-Daughter are structure-shared. This is illustrated by co-indexation of the heads ([1] - [1]) in Figure 3.27. Of the remaining principles, the *Semantics Principle* requires that the CONTENT value of a Head-Nexus phrase be structure-shared with the Head Daughter (coindexed [2] - [2] in the figure), whereas the CONTENT value of an Adjunct-Head phrase is structure-shared with that of the Adjunct Daughter (not depicted). The Head-Nexus phrase has subtypes Complement-Head, Subject-Head, Specifier-Head and Filler-Head phrases. Of these, the Complement-Head phrase and Subject-head phrase have already been shown to be applicable to Japanese.

The *Valence Principle* requires the saturation of subject, complement and specifier as the result of application of their respective phrasal signs, while the *Empty Comps Constraint* in essence requires that complement saturation takes precedence over any other phrasal operation. As I will show later, in Chapter 5, the ECC is subject to reformulation to account for the data of phrasal conjunctions involving verbs which have been hitherto labelled "auxiliaries". I have made two additions to this type hierarchy: Sag's analysis makes no use of the Head-Marker phrase, though this is needed for the analysis of case markers and verb markers. This is attached as a subtype of Head-Nexus phrase, not directly subject to the ECC. Phrases are a subtype of *sign*, as are words, and I have indicated this division in the figure. The

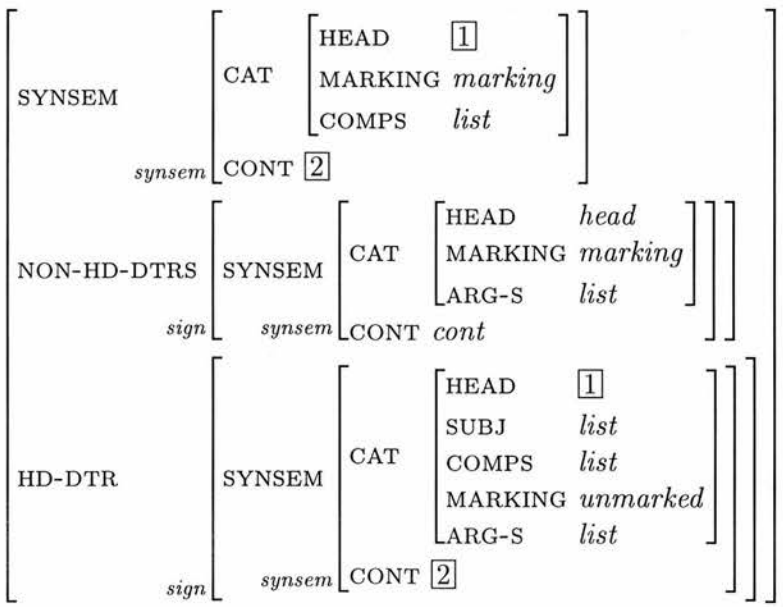


Figure 3.27: Head-Nexus Phrase (Totally Underspecified)

synsem value is partitioned into *wsynsem* for words with ARG-S, and *psynsem*, for phrases without ARG-S.

3.5.2 Lexical Configuration and ARG-S

Word signs contain the SYNSEM value shown in Figure 3.28 (a) and phrasal signs contain the SYNSEM value shown in 3.28 (b).

(a):

$$\begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{head} \\ \text{SUBJ} & \textit{list} \\ \text{COMPS} & \textit{list} \\ \text{ARG-S} & \textit{list} \end{bmatrix} \\ \text{CONT} & \text{CONT} \end{bmatrix} \end{bmatrix}$$

(b):

$$\begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{head} \\ \text{SUBJ} & \textit{list} \\ \text{COMPS} & \textit{list} \end{bmatrix} \\ \text{CONT} & \text{CONT} \end{bmatrix} \end{bmatrix}$$

Figure 3.28: Word and Phrasal *synsem*

If ARG-S remains embedded as a feature under CAT then the complexity of the type hierarchy is needlessly increased because in addition to the articulation of *synsem* into word and phrasal types, the feature CAT must also be articulated into word and phrasal types. Yet the function of ARG-S is not simply a syntactic one, it also plays an important role as the interface between syntax (CAT) and semantics (CONT) and is therefore in some sense independent of both. This follows the discussion of Davis (1995) in Section 3.4.

If the ARG-S list is separated out from both CAT and CONT in lexical signs, as in Figure 3.29(a), then the binding possibilities are not damaged, and in addition the ARG-S position as the locus of linking is made independent[8]. Moreover, the distinction between the SYNSEM values of lexical and phrasal signs can be straightforwardly made in the *synsem* type hierarchy, as shown in Figure 3.29(b), where CAT and CONT are features of *synsem* inherited by both *wsynsem* and *psynsem*, and ARG-S is a feature introduced in *wsynsem*. Finally, lexical signs include the *wsynsem* specification while the *synsem* value of phrasal signs is *psynsem*[9].

3.5.3 Phrasal Feature Geometry

The application of Sag's phrasal principles to example (1) yields a complex phrasal feature structure as illustrated in Figure 3.30. The individual lexical items contain

[8]The separation of ARG-S from CAT is also advocated in Sag & Wasow (1999).

[9]Phrasal signs also include daughters values. See Figure 3.30.

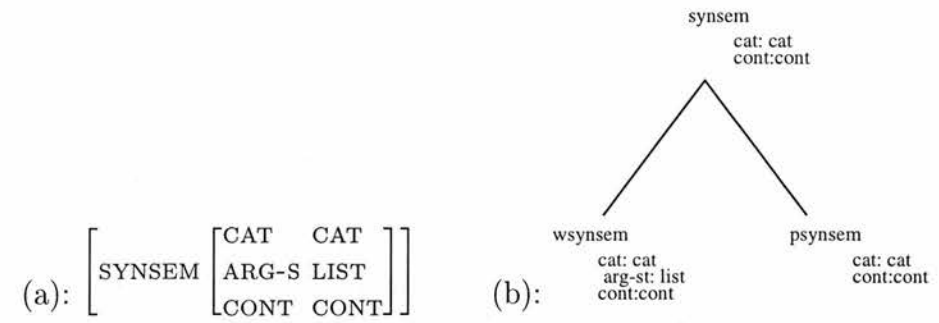


Figure 3.29: (a) *Arg-s* Repositioned, (b) The *synsem* Hierarchy

wsynsem structures of which only the verb *nonda* has a non-trivial argument structure, representing its SUBJ (1) and COMPS (3) arguments, and these are linked to semantic roles *a* and *c*, which are of content type *nom-obj*, following Davis (1995). *Oguchi ga* and *biiru o* are Head-Marker phrases indexed 1 and 3, with nominal heads, and with marking values of *nom* and *acc*, respectively; these have SYNSEM values of type *psynsem* so they carry no argument structure ARG-S feature. The complement value of *nonda* (drank) is coindexed with the CAT value 3 of *biiru o*, and the undergoer (UND) value is coindexed *c*, with the *nom-obj* value of the complement's content. Similarly, the subject value of the verb is coindexed with the *cat* (1) of *Oguchi ga*, and the actor (ACT) with the content value of the SUBJ, indexed *a*. The Valence Principle ensures that the SUBJ and COMPS values of the verb are cancelled off, and the Empty Complements Constraint ensures that the SUBJ value cannot be cancelled off if the COMPS value is not. The Head Feature Principle ensures that the HEAD values of the verb are structure-shared with the phrasal mother, and the Semantics Principles guarantees that the CONTENT value of the phrasal mother is structure-shared with that of the verb.

3.6 Summary

In this chapter I have introduced the aspects of HPSG relevant to the analysis of the fragment of Japanese used in this thesis. The data addressed is Japanese verbal alternations, and the lexical components include verbs, nouns and case markers; the associated phrasal structures are Head-Marker, Complement-Head and Subject-Head, and these structures reflect Japanese linear precedence ordering, with markers following heads, and both complements and subjects preceding heads.

Argument structure, and the linking between syntactic arguments and semantic roles, are important to the development of this account. Existing accounts of argument structure, exemplified by Wechsler (1995) and Manning (1994) create some

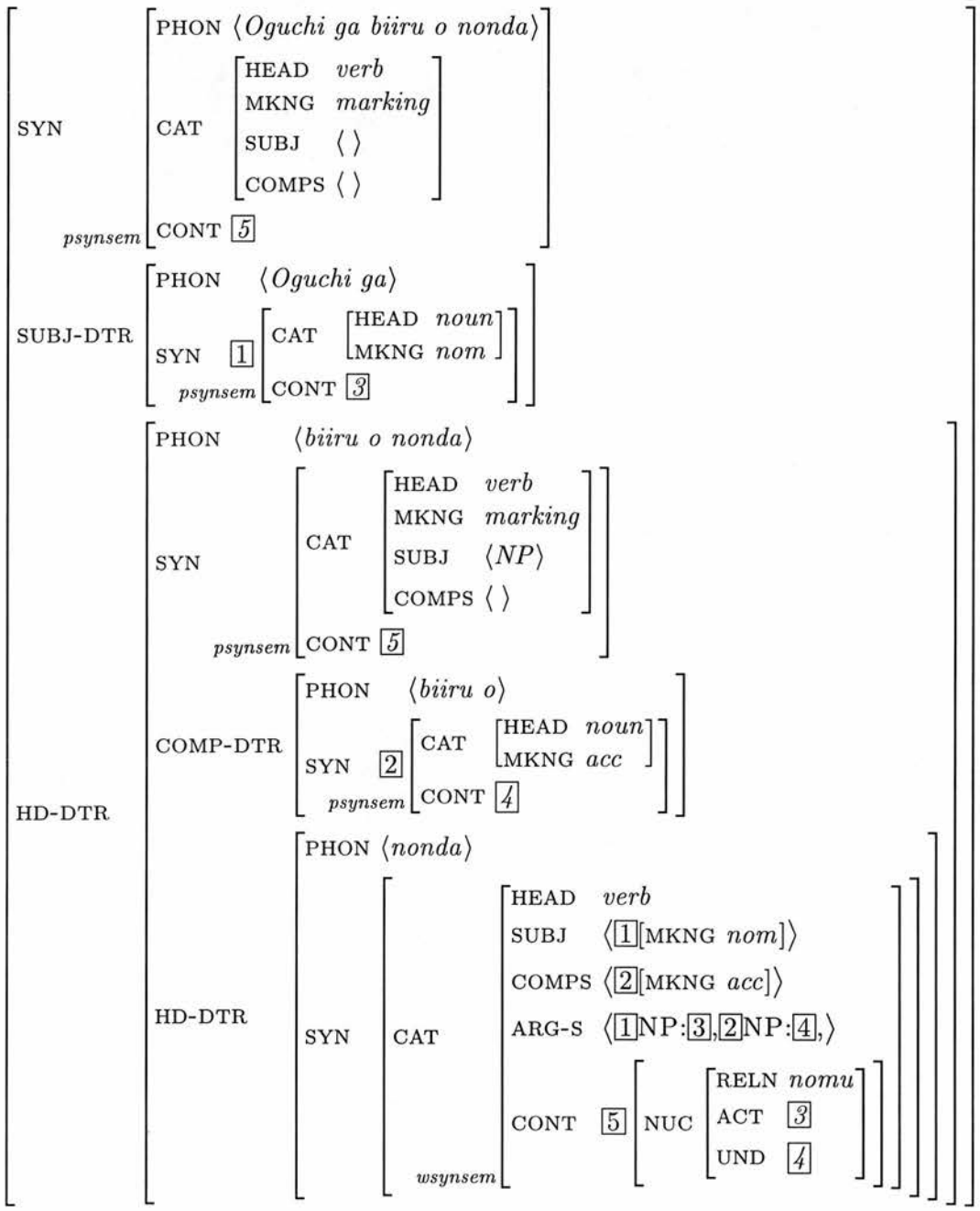


Figure 3.30: Phrasal Sign for "Oguchi-kun ga ..."

ambiguity in what is meant, in HPSG, by argument structure,but later HPSG, such as Sag (1997), adopts the linking theory of Davis (1995), which puts role mapping and ordering on a more strongly constrained basis. Although the current theory insists that argument structure is strictly a feature of lexical types, not shared by phrases, no systematic account of the consequences of this stipulation on the feature geometry, have been developed. Articulating the *synsem* type hierarchy into word synsem (*wsynsem*) containing ARG-S and phrasal synsem (*psynsem*) without it, seems a reasonable solution.

Chapter 4

Polysemy in Verb Frame Alternations

The use of Hasegawa's (1995) typology of Valence-Maintaining and Valence-Changing predicates, with respect to the first conjunct verb in a *te* conjunction, obscures an important generalization about the valence characteristics of the second conjunct among many of the verbs which exhibit the *te* conjunctive form as an alternate verb frame. This is that, if the *te* marked conjunct is taken as a VP, this VP alternates with an NP as complement of the second conjunct verb in an apparently polysemous construction. This is an important consideration in lexicalist theory where the objective is to find the most general structure for the lexicon, which is done in HPSG by articulating type hierarchies using single and multiple inheritance of lexical properties (or features). The development of a lexical account of the various forms of polysemy is a way to capture some aspects of this generality. Only some of the verbal data presented in Chapter 2 suggests a polysemous analysis, while the remaining verbs must be homonymously listed in the lexicon, as their different simplex and conjunct forms only really show phonological similarity. There is a possible explanation for the apparent distinction in the various verbal constructions in the evidence of grammaticalization (Hopper & Traugott (1993)). According to this theory, evidence of the evolution of various language phenomena is present in synchronic data in the appearance of clines of grammaticalization. In this case it seems plausible that a cline of grammaticalization of valence proceeds through a period of polysemy towards the complete separation of verb forms. In the next two chapters I will discuss the verbal data in its synchronic state, concentrating mostly on the lexicalist analysis of apparently polysemous forms. I do not go into a diachronic analysis of the possible processes of grammaticalization as such.

The evidence for a polysemous analysis is most clear in the cases of *miru* (see), *morau* (receive) and *ageru* (give). Examples (1) and (2) show the alternations of *miru*, with an *o* marked accusative complement *keeki o* (a cake) in (1), and a *te* marked VP, *keeki o tukutte* (baking a cake) in (2). If it can be shown that there is a correspondence between these two complement types, and if their selectional restrictions are compatible, then *miru* can be encoded as a single entry in the lexicon.

(4.1) Sugiyama-san ga keeki o miru.
Mr Sugiyama-NOM cake-ACC see
"Mr Sugiyama sees the cake".

(4.2) Sugiyama-san ga keeki o tukutte miru.
Mr Sugiyama-NOM cake-ACC bake-TE try
"Mr Sugiyama tries baking a cake".

The selectional restrictions on the nominal complement require that it be physically or metaphorically visible. The VP complement is further restricted by aspectual class. The *te* marked verb may be an Activity or an Accomplishment, but it may not be a State or an Achievement. Hitherto, there has been no provision in HPSG for identifying the aspectual class of a verb, and the CONTENT must be extended to account for it. For this reason, Pustejovsky (1995) is reviewed in Section 4.1 with a view to incorporating the aspectual machinery. The analysis of polysemous *miru* follows in Section 4.2. Among the literature on *te* conjunctions, Ishikawa (1985) and Matsumoto (1996) have looked at the *te morau* construction and developed a lexical analysis using Lexical Functional Grammar (Bresnan (1982)). In Matsumoto's formulation of that analysis, the assumption is that one of the expressed arguments of the head verb is overtly coindexed with the unexpressed subject of the embedded verb in a *control* construction. The Ishikawa/Matsumoto analysis is introduced in Section 4.3, and motivates a general analysis of *te* conjunctions as control constructions. The development of phrasal structures for *morau* and *ageru* follows in 4.4, and this chapter is summarized in Section 4.5.

4.1 Pustejovsky on Polysemy

Pustejovsky's inquiry into lexical semantics is associated with an interest in computational linguistics, and so he develops computationally tractable representations of the lexicon, which capture the polysemous behaviour of words. The simplest form of lexicon is to list every word sense and every subcategorization frame as a separate

entry, and Pustejovsky calls this a *Sense Enumerative Lexicon.* Using this model, the different contrastive senses of a word such as 'line', from Examples (3) and (4) are separately listed in the lexicon, and selected during a parse according to their sense. The *GENUS* field in the examples in Figure 4.1 is intended to characterize the semantic interpretation of the different forms of the noun 'line', including line_1 as written material, related to a letter, and line_2 as a sort of connector.

(4.3) I'll drop you a line next week sometime.

(4.4) The line went dead, so I can't phone you.

(a): $\begin{bmatrix} \text{PHON} & \textit{line}_1 \\ \text{CAT} & \textit{noun} \\ \text{GENUS} & \textit{written-material} \end{bmatrix}$ (b): $\begin{bmatrix} \text{PHON} & \textit{line}_2 \\ \text{CAT} & \textit{noun} \\ \text{GENUS} & \textit{connector} \end{bmatrix}$

Figure 4.1: SEL Entries for **line**.

Complementary ambiguities such as the figure/ground alternation of 'door' in examples (5) and (6) are also separately listed, as are the different subcategorized forms of 'begin' which are shown here in Figure 4.2 as $\textit{begin}_1$ $\textit{begin}_2$ and $\textit{begin}_3$, differentiated by their argument structures[1] and semantic relations.

(4.5) Pablo painted the door blue again.

(4.6) The cat shot straight out through the door.

A Sense Enumerative Lexicon has traditionally offered the advantage of providing a lexicon relatively independent of syntactic and semantic theories. The price paid, however, is in the redundant listing of senses and substructures, and the consequent limited possibilities offered for syntactic and semantic generalization. Pustejovsky's program enriches the semantic representations of words listed, by giving the lexicon greater internal structure, using the mechanisms of *argument structure*, *event structure* and *qualia structure*, bound together through a Lexical Inheritance Hierarchy, which uses the techniques of hierarchical typing and multiple inheritance found in HPSG and formally defined in Carpenter (1992). The various components of this framework are introduced in Subsections 4.1.1 through 4.1.4, and following some correction for defects in Pustejovsky's aspectual analysis, these are integrated into the HPSG framework in Subsection 4.1.5.

[1]Pustejovsky's (1995) definition of argument structure is very different from the recent HPSG tradition, resembling more the listing of roles in the NUCLEUS than the separate ARG-S list which is the locus of linking and binding. The difference is more comprehensively addressed in 4.1.1.1.

$$\text{(a):}\begin{bmatrix}\text{PHON} & \textit{begin}_1\\ \text{CAT} & \textit{verb}\\ \text{SEM} & R_1(\theta_1,\theta_2)\\ \text{ARGSTR} & \begin{bmatrix}\text{ARG1}\ \textit{NP}\\ \text{ARG2}\ \textit{VP}(+\textit{inf})\end{bmatrix}\end{bmatrix}\qquad \text{(a):}\begin{bmatrix}\text{PHON} & \textit{begin}_2\\ \text{CAT} & \textit{verb}\\ \text{SEM} & R_2(\theta_1,\theta_2)\\ \text{ARGSTR} & \begin{bmatrix}\text{ARG1}\ \textit{NP}\\ \text{ARG2}\ \textit{VP}(+\textit{prog})\end{bmatrix}\end{bmatrix}$$

$$\text{(c):}\begin{bmatrix}\text{PHON} & \textit{begin}_3\\ \text{CAT} & \textit{verb}\\ \text{SEM} & R_3(\theta_1,\theta_2)\\ \text{ARGSTR} & \begin{bmatrix}\text{ARG1}\ \textit{NP}\\ \text{ARG2}\ \textit{NP}\end{bmatrix}\end{bmatrix}$$

Figure 4.2: SEL Entries for **begin**.

4.1.1 Argument, Event and Qualia Structures

Among the problems arising from reliance on a Sense Enumerative Lexicon are the multiple, redundant listing of related lexical items, and the duplication of information common to classes of item. For these reasons linguistic research focuses on ways to organize the lexicon, and in HPSG this research leads in the direction of using inheritance techniques in a *hierarchical lexicon* (cf Flickinger (1987), Davis (1995)) with the ultimate goal of expressing generalizations applicable to classes of words only once in the lexicon. Pustejovsky addresses deeper semantic problems by organizing the semantics of the lexicon using a set of structures, and a set of generative devices to relate them. His representational levels are *Argument Structure*, *Event Structure*, *Qualia Structure* and *Lexical Inheritance Structure* and these are exploited using the devices of *Type Coercion*, *Selective Binding* and *Co-Composition*. **Type Coercion** occurs when a subordinate lexical structure is coerced to a different semantic representation, whereas **Selective Binding** occurs when a lexical structure operates on a specific substructure, without changing its type. In **Co-Composition**, multiple elements in a phrase combine to generate novel senses in composition.

4.1.2 Argument Structure

While HPSG argument structure, through the ARG-S list, provides the locus of binding theory, and of linking between valences and content, Pustejovsky's argument structure more closely resembles the NUCLEUS feature complex of HPSG, where the

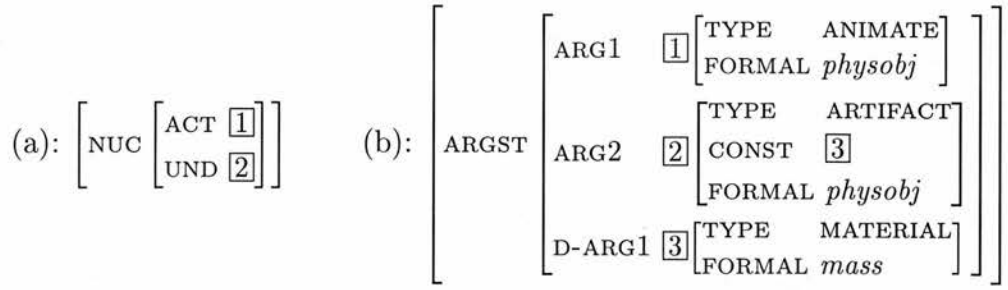


Figure 4.3: HPSG *nucleus* and GL *argument structure*

semantic specifications of arguments are listed. Generative Lexicon argument structure makes reference to 4 different kinds of arguments: *true*, *default* and *shadow arguments*, and *true adjuncts*.

True arguments are direct arguments, always expressed and assigned a semantic role by the verb, and so they are analogous to the ACT, UND and LOC in HPSG projected to ARG-S;

Default arguments are logical arguments of the qualia, not necessarily expressed in syntax; these have different representations in HPSG, sometimes appearing in ARG-S as pronominal elements (Manning *et al.* (1999)), and sometimes appearing as adjuncts, as for instance the representation of the *actor* in a passive verb, which in English surfaces as an optional prepositional 'by' phrase, and in Japanese surfaces as an optional postpositional *ni* phrase - and we have already seen in Manning (1994) that prepositional adjuncts are ordered as the oblique elements of ARG-S.

In example (7), while 'John' and 'house' are true arguments directly assigned a semantic role by the verb 'build', 'bricks' is a default argument, denoting the material of which the house is constituted. In Pustejovsky's argument structure, this default argument is related to the direct argument, while current HPSG treats it as an adjunct, not directly represented in the semantics of the verb. The difference between the HPSG nucleus for the verb 'build' and Pustejovsky's argument structure, is illustrated in Figures 4.3(a) and (b). The HPSG nucleus represents 'John' and 'house' as ACT and UND respectively, while the Generative Lexicon argument structure represents these direct arguments, and the constitutive relation between 'house' and 'bricks'.

Shadow arguments include things like the particular type of 'dance' involved in a verb like 'to dance': not normally expressed in the syntax as in (8), but it may be given to add specificity, (10). These arguments are semantically incorporated into

the lexical item and are lexically represented in Qualia Structure, as will be shown in Subsection 4.1.1.3 below.

True adjuncts are typically temporal, spatial or perhaps modal modifiers, as with the (spatial) locative "in Boston" in (11).

(4.7) John built a house out of bricks.
(4.8) Mr Sugiyama danced.
(4.9) ??Mr Sugiyama danced a dance.
(4.10) Mr Sugiyama danced the paso doble.
(4.11) Mary saw Bill in Boston.

The status of these different types of argument is mediated by Qualia Structure, which is in turn coerced into surface expression through Event Structure, introduced next.

4.1.3 Event Structure

Pustejovsky develops a model of Event Structure based on the work of Moens & Steedman (1988) and he argues that verbal events comprise processes, states and transitions, and that (some of) these events have subeventual structure, as in Figure 4.4.

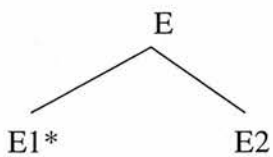


Figure 4.4: Subeventual Structure

In the general case a subevent can be any of process, state or transition and these can occur sequentially (E1 then E2) or overlapping (E2 starts before E1 finishes). An Accomplishment such as "John built the house" is an example of an event comprising a process and a state, in strict partial order: process = 'John builds', followed by result = "a house, built'. An example of a process followed by a transition is a causative: "His wife made John build their house", where "His wife" is the causing agent in an E1 headed event, and "John build their house" constitutes the event at E2. Subevent headedness, illustrated by starring the headed subevent in 4.4, allows a way to emphasize one or other event : an Accomplishment is focused on the initial event, so is E1 headed, while an Achievement emphasizes the resultant

state, according to Pustejovsky, so is E2 headed. As will be shown in the discussion of Qualia Structure next, the practical consequence of this "headedness" is in its effective licensing of the arguments to be projected.

$$\begin{bmatrix} \text{EVENTSTR} & \begin{bmatrix} \text{E1 } \textit{process} \\ \text{E2 } \textit{state} \\ \text{RESTR } \langle \\ \text{HEADED E1} \end{bmatrix} \end{bmatrix}$$

Figure 4.5: An Event with Subeventual Structure.

The ordering relations for bieventual structures include "partial order" ($\langle$) in which E2 strictly follows E1 (E1 $\langle$ E2); "overlap" ($\circ$) in which E1 and E2 occur simultaneously (E1 $\circ$ E2); and "partially ordered overlap" ($\langle\circ$) where E1 starts first, then E2 starts before E1 finishes (E1 $\langle\circ$ E2). The number of possible combinations of events, orderings and headedness relations is (3x3) x 3 x 4, for a total of 108, though Pustejovsky (1995) suggests examples of 12 of these orderings.[2]

. There is no direct mapping between these event pairs and the Vendler-Dowty classes. Pustejovsky does not offer a structure for States, but a bi-eventual structure is not needed.

The clearest example of bi-eventual ordering is given in the mapping to an Accomplishment "build", which is an E1 headed event, with strict partial ordering (E1* $\langle$ E2). Figure 4.5 shows a feature structure representation for an Accomplishment, and in this case a state (E2) follows a process (E1) in strict partial order ($\langle$) with the first subevent (starred) as head. Event headedness interacts with Qualia Structure to determine which arguments are projected, and which arguments are defaulted, or shadowed. This compares with Dowty's logical structure given in Table 4.1(d), where event1 ϕ causes event2 ψ, and ϕ and ψ can be any well-formed logical structure. An LS includes a predicate (=state), do (=process), or become(predicate) (=transition), or indeed ϕ cause ψ. Both of these latter cases are equivalent to uncaused, and caused transitions, respectively. Dowty's *cause* and *become* are variants of Pustejovsky's partial order, with the extra notion of agentivity. An Achievement in Pustejovsky's system is an E2 head, partially ordered event (E1 $\langle$ E2*), and he

[2]Pustejovsky's (1995) orderings are based on 2 variables, E1 and E2; three ordering relations among these variables: partial order ($\langle$), overlap ($\circ$) and partially ordered overlap ($\langle\circ$); and four headedness relations, E1 head, E2 head, both headed and unheaded. This gives 3 X 4 = 12 combinations. As examples, the overlapped relation $[e1_* \circ e2]$, is given as the representation of 'buy', while 'sell' is an E2 headed overlapped relation, $[e1 \circ e2_*]$. The Activity 'walk' is given as an E1 headed, partially ordered overlapping relation, $[e1_*\langle\circ e2]$, but with a goal as in 'walk home', it is given as an E2 headed, partially ordered overlapping relation $[e1\langle\circ e2_*]$.

No.	**Verb Class**	**Logical Structure**	**Event Structure**
(a)	State	predicate'(x) or predicate'(x,y).	e1:state, hdd:e1
(b)	Activity	do'(x,[predicate'(x)]) or do'(x,[predicate'(x,y)]).	e1:process, hdd:e1
(c)	Achievement	become[State LS].	e1:process, e2:state, restr:⟨, hdd:e2
(d)	Accomplishment	ϕ cause ψ where ϕ and ψ are any well-formed LS. However ϕ is typically an activity and ψ an achievement.	e1:process, e2:state, restr:⟨, hdd:e1

Table 4.1: Logical Structures and Event Structures

claims that in this case the action is focused on the result. For Dowty, an Achievement is an uncaused transition *become(predicate)*, so again he makes the point about agentivity, that this is a non-agentive transition. For Activities, Dowty offers the *do* operator, equivalent to a *process*; Pustejovsky claims that an Activity such as 'walk' is an E1 headed bi-eventual structure which is "partially ordered, overlapping". While Dowty is not completely satisfied with the use of the *do* operator to define Activities, his dissatisfaction stems from the fact that some Activities are non-agentive (buzz, roll, non-agentive motion verbs), and *do* cannot successfully account for these. Pustejovsky's bi-eventual structure adds extra machinery and still does not account for differences in agentivity. Finally, in Dowty's scheme, a State is mapped onto a *predicate*, while Pustejovsky offers no logical structure for States.

In constructing a system to represent aspectual classes later, it will be necessary to remodel the Event Structure to reflect Dowty's findings more closely. Before that, a discussion of the Qualia Structure is in order.

4.1.4 Qualia Structure

The *Generative Lexicon* theory involves the interaction of word meaning and compositionality to generate a larger set of word senses than the traditional homonymous

listing allows. The mechanism which allows each word to potentially express multiple meanings, depending on its mode of combination, is Qualia Structure. Every contentive lexical item has up to four different types of relation, in composition, and these are Qualia:

- **Constitutive**: the relation between an object and its constituent parts,
- **Formal**: that which distinguishes it within a larger domain,
- **Telic**: its purpose and function,
- **Agentive**: factors involved in its origin or "bringing it about".

Pustejovsky (1995:p.76)

For example a noun such as 'novel' expresses each of these qualia: its constitution is that of a "narrative", formally it is a "book", its telic role is "reading", and its agentive role one of "writing". When it combines with a verb such as 'begin' in example (12), the telic meaning of *reading* or the Agentive meaning of *writing* are possible interpretations of the composed verb phrase ' "begin a novel", so 'begin' and 'novel' coerce these possible meanings, and not the Constitutive or Formal Qualia. If we have the world knowledge that Mary is a novelist then the possibilities include "Mary writing a novel".

(4.12) Mary began a novel.

Constitutive and Telic Qualia seem to be relevant to the interpretation of nouns only, with Constitutive, for example, involving a static *part-of* relation, and while Formal and Agentive Qualia are encoded in other lexical categories, they are interpreted differently for verbs: a verb such as 'bake' has both Formal and Agentive Qualia, and this allows the projection of one or both "true" arguments, with the Agentive Quale projecting both *actor* and *undergoer* arguments in example (13), and the Formal Quale projecting the *undergoer* in (14), where the *actor* is shadowed.

(4.13) Keiko baked a cake.
(4.14) The cake baked.

The selectional restrictions on the arguments include that the first (actually the *actor*) argument is animate, and the second (actually the *undergoer*) is a physobj. The *actor* engages in the process of 'baking' and the object ends up in a result state, 'baked'. The item is underspecified for headedness: in the case of E1 headedness the agentive *bake-act* arguments are projected, with the animate argument as

$$\begin{bmatrix} \text{ARG-S} & \langle \boxed{1}\text{ANIMATE}, \boxed{2}\text{PHYSOBJ} \rangle \\ \text{EVENTSTR} & \begin{bmatrix} \text{E1} & \textit{process} \\ \text{E2} & \textit{state} \\ \text{RESTR} & \langle \\ \text{HDD} & \text{EITHER} \end{bmatrix} \\ \text{QUALIA} & \begin{bmatrix} \text{FORMAL} & \textit{bake-result}(\textit{e2}, \boxed{2}) \\ \text{AGENTIVE} & \textit{bake-act}(\textit{e1}, \boxed{1}, \boxed{2}) \end{bmatrix} \end{bmatrix}$$

Figure 4.6: Event and Qualia Structure for *bake*.

subject and the physobj as accusative object (syntax not specified in this figure); in the case of E2 headedness, the formal quale projects the arguments associated with the formal *bake-result* predicate so the physobj (undergoer) is projected to subject and the agent is *existentially quantified over*. The verb 'bake' is a member of a class of verbs which undergo causative/unaccusative alternation, where the causative form is transitive, projecting the actor/animate argument to subject and the undergoer/physobj argument to object. The unaccusative form is intransitive, projecting the undergoer argument to subject and 'existentially quantifying' over the actor, so it is not expressed, and is in Pustejovsky's terminology, "shadowed".

While the set of causative/unaccusative verbs is underspecified for headedness, and contains agentive and formal qualia both potentially expressible, a pure unaccusative such as *die* only has a Formal Quale expressible, as in (15) and so it is obligatorily E2 headed. The agent of death is retrievable as an adjunct (16), but is listed in the *argument structure* as a *default* argument, so it cannot be directly expressed as an agent, as illustrated by example (17). An agentive/animate argument is, then, encoded in the lexical semantics of a pure unaccusative verb, and may be expressed obliquely, but never directly.

(4.15) The cow died.
(4.16) The cow died from mad cow disease.
(4.17) *Mad cow disease died the cow.

There are 4 levels of semantic structure in Pustejovsky's scheme and these are argument, event and qualia structure, reviewed above, and also lexical inheritance structure. The first three levels interact by means of the operations of type coercion, selective binding and co-composition. A verb like 'want' takes three different types of complement, an NP in (18), an infinitival VP in (19) and an infinitival sentence in (20).

(4.18) Thabo wants an ice cream.

(4.19) Nelson wants to kiss Hardy.

(4.20) Scott wants Oates to step outside for a while.

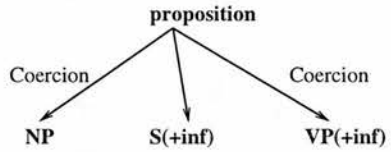


Figure 4.7: Complement Coercion

Pustejovsky notes that to accommodate the different complement types selected by control verbs such as 'want', one can either enumerate the senses, as suggested by Dowty (1985), or select for a single type of complement, and have this complement undergo a type-shifting operation when it combines with the verb. This is illustrated in Figure 4.7, where 'want' selects for a type *proposition* and the complement is coerced, or type-shifted, under lexical governance from the verb. Thus, the NP complement of (18), the VP complement of (19) and the S(inf) complement of (20) are type-shifted to match the *proposition* type required. My interest is in integrating Pustejovsky's scheme into HPSG as a way to explain polysemous phenomena which cannot otherwise be represented. HPSG has two possible ways to realize the type-shifting operation inherent in these constructions, and these are *default* and *underspecification.* For example:

(a) the default type is S(inf) and it can be overridden by a different type such as NP or VP; or

(b) *NP*, *S(inf)* and *VP* are subtypes of a supertype *proposition*, which the verb selects for: that is, complement selection is underspecified for the subtypes of the *proposition* type hierarchy.

The first case can be understood in the terms of a system of default unification such as that of Lascarides & Copestake (1999), in which the non-default value of a specific feature value overrides the default value on the supertype. The second case is directly compatible with an HPSG type hierarchy in which the subtypes of *proposition* are admitted in virtue of the verb's selection of *proposition* as complement. In the subsequent analysis of Japanese verbs, I maintain the HPSG underspecification analysis.

One consequence of the view that NP, VP and S(inf) are to be treated the same is the question of how such an underspecified complement can be described and

selected for in a lexicalist analysis. In the traditional HPSG analysis of 'want' there are homonymous lexical forms which select for an NP complement, an infinitival VP or an infinitival sentence. These arguments must be underspecified categorially in the COMPS of 'want' to account for the alternation. This can be done by regarding all subcategorized elements as nominal forms, or by creation of a supertype for nouns and verbs, or, more likely, by some combination of the two techniques. This particular question, of how to treat seemingly incompatible complements to yield a polysemous analysis, is deferred to the next section, and examined there in the light of the Japanese data which is the subject of this thesis. In the next subsection, Pustejovsky's Argument, Event and Qualia structures are made compatible with the HPSG framework, so that lexical entries with a fully specified syntactic projection can result.

4.1.5 Integrating Pustejovsky (1995) into HPSG

The functionality which Pustejovsky's framework provides, which is not currently available in HPSG, includes the ability to express aspectual classes and the ability to express polysemy in lexical entries through alternations in qualia structure. There are, however, defects in the framework which prevent a direct and seamless integration. For example, Pustejovsky's model of aspect develops a structure for Activities, Achievements and Accomplishments, but not for States. Moreover, while Achievements and Accomplishments are telic events involving a change of state, and therefore encoded as bi-eventual, Activities are also encoded as bi-eventual. For example, Pustejovsky characterizes the verb *walk* as a bi-eventual structure with overlapping subevents, and justifies this as an encoding of "the efficient motion of the legs leading to the final motion of the body". Dowty's (1979) view of Activities (notwithstanding his reservations about **do**) is more consistent and well worked out, and I propose to encode Activities in the aspectual hierarchy as mono-eventual, following this model. Dowty also develops an analysis of Stative verbs in which they are encoded as a single predicate, and these States must also be encoded in the aspect hierarchy. By adapting Pustejovsky's Event Structure framework we can encode the four Vendler-Dowty aspectual classes in a type hierarchy. This, moreover, splits along the lines of the *agentivity*, *telicity* and *stativity* features identified in Chapter 1. An Event Structure which shows mono-eventual States and Activities, and bi-eventual Achievements and Accomplishments, is depicted in Figure 4.8. This represents the possible values of the Event Structure (EVSTR) feature under *content*.

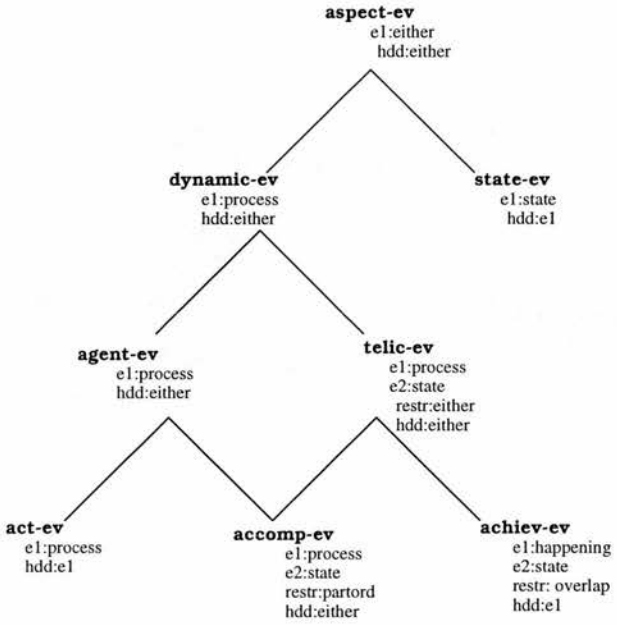


Figure 4.8: Aspect Sensitive Event Structure

All verbs are associated with at least one event variable, encoded as feature E1, and *dynamic* events differ from *state* events in the value of this variable: States have an E1 value of *state*, Dynamic verbs have an E1 value of *process*. Dynamic verbs are further subdivided according to *agentivity* and *telicity*. *Agentive* events add no new features; *telic* events are bi-eventual, adding an E2 subevent with the value *state*. *Telic* events are associated with an ordering restriction, and Pustejovsky argues that both Accomplishments and Achievements are part-ordered (i.e. subevent E1 occurs before subevent E2), with complementary headedness values. If this were true, the Achievements with an E2:*state* "headedness" value would seem to be simply stative predicates. In chapter 5, I argue from the Japanese data that the *process/state* headedness alternation is actually a characteristic of Accomplishments alone, and this is encoded in the figure with a headedness value underspecified for either of E1 or E2 subevents. So what then is the event structure of an Achievement?

The distinction between the two forms of 'bake' given in examples (13) and (14) is that in the Accomplishment case, 'Keiko' is the causative agent, while in the Achievement case, no cause is specified. The outcome in both cases is a Stative "the cake is baked". It seems, then, that there may be a distinction between the forms of *process*, which can be the value of E1. If the general case of a caused or uncaused process can be called a *culmination*, to borrow a term from Moens & Steedman (1988), then *process* can be the label for a caused process, and perhaps

happening can be the label for an uncaused process. The subevents E1 and E2 are, then, features with a type *subevent* with values as depicted in Figure 4.9(a).

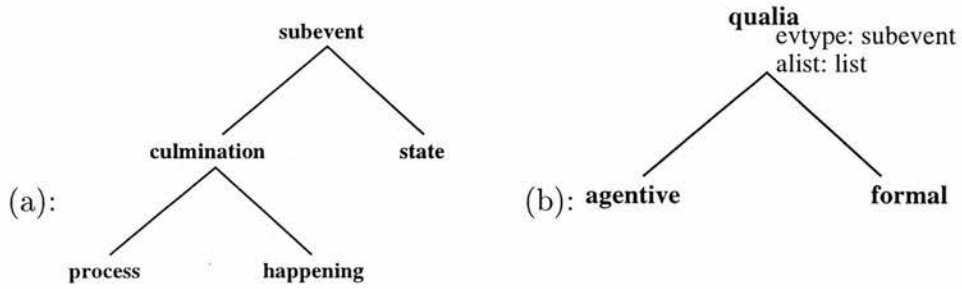


Figure 4.9: Type Hierarchies for (a) Subevents and (b) Qualia

The Event Structure value of an Achievement is, then, led by an E1:*happening* feature, as shown in Figure 4.8.

A critical way in which Pustejovsky's framework does not map directly into HPSG is in the relation of headedness, qualia structure and projection to syntax. Under the *Default Causative Paradigm*, the semantic structure of a verb like 'bake' given in Figure 4.6, earlier, is underspecified for the Accomplishment reading, with *animate* and *physobj* arguments, and the Achievement reading with *physobj* only. The projection to syntax of either of these two alternatives is controlled by the headedness value: the Accomplishment form of 'bake' is associated with the *agentive* quale as head, and *animate* is projected to SUBJ, *physobj* to COMPS; the Achievement form of 'bake' is associated with the *formal* quale as head, so the *physobj* is projected to SUBJ and the *animate* is 'shadowed' and therefore not projected. What this implies in terms of a mapping to HPSG structure is that Accomplishments and Achievements have a common supertype, and the Accomplishment projection establishes the relationship between category (CAT) and content (CONT) in which the SUBJ and COMPS are linked with *animate* and *physobj* arguments respectively. The achievement projection establishes the relationship in which SUBJ and *physobj* are linked. Headedness, then, is a principle in which one or other of the qualia is expressed as the maximal subtype encoded in the lexicon. The Qualia Structures which appear in Pustejovsky's diagrams, will be rendered differently in an HPSG lexical entry. In an Accomplishment, the *agentive* quale is represented, and in an Achievement, the *formal* quale is given. Pustejovsky shows these qualia as containing a subevent variable and an ordered list of roles. These are represented in a lexical entry as values of features EVTYPE and ALIST. The qualia hierarchy is shown in Figure 4.9(b) with both *agentive* and *formal* subtypes having features EVTYPE and ALIST. The EVTYPE gives the headedness value determining which quale is activated. The list

of roles on the ALIST seems to be compatible with the 'direct' arguments (cf discussions of Wechsler (1995) and of Manning (1994) in the previous chapter) which are elements of the ARG-S list. Further work on the relationship between qualia and role ordering in ARG-S would be fruitful, perhaps leading to a more streamlined lexical geometry.

The combination of standard HPSG with Pustejovsky's lexical semantics is given in Figure 4.10, with the Accomplishment and Achievement projections of the verb 'bake'. In 4.10(a), the event structure is that of a Accomplishment (*accomp-ev*) with bi-eventual structure, headed by an E1 event of type process. The ALIST value of qualia lists both *animate* and *physobj* roles, and these are linked with the elements of ARG-S. The Achievement form in 4.10(b) has an event structure, *achiev-ev*, which is also E1 headed, but in this case the value is that of *happening*, reflecting the fact that an Achievement is an uncaused process. The UND/*physobj* argument is linked to the ARG-S role. Projection from ARG-S to valence in both figures follows standard HPSG conventions.

Following this integration of Pustejovsky (1995) and HPSG, the means are available to analyse the simplex and conjunctive forms of Japanese verbs, beginning with the analysis of *miru* (see) and its conjunct form *V-te miru*.

(a):

$$\begin{bmatrix}\text{PHON } \langle bake \rangle \\ \text{SYN} \begin{bmatrix}\text{CAT} \begin{bmatrix}\text{HEAD}_{verb}[\text{VFORM } fin] \\ \text{SUBJ } \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS } \langle \boxed{2}[\text{MKNG ACC}] \rangle \\ \text{ARG-S } \langle \boxed{1}_{NP}:\boxed{3}, \boxed{2}_{NP}:\boxed{4} \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix}\text{NUC}_{act\text{-}und\text{-}rel} \begin{bmatrix}\text{ACT } \boxed{3}_{nom-obj} \\ \text{UND } \boxed{4}_{nom-obj}\end{bmatrix} \\ \text{EVSTR } \boxed{5} \begin{bmatrix}\text{E1} & process \\ \text{E2} & state \\ \text{RESTR} & \langle \\ \text{HDD} & e1 \end{bmatrix}_{accomp\text{-}ev} \\ \text{QUAL}_{agentive} \begin{bmatrix}\text{EVTYPE } e1 \\ \text{ALIST } \langle \boxed{3}, \boxed{4} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

(b):

$$\begin{bmatrix}\text{PHON } \langle bake \rangle \\ \text{SYN} \begin{bmatrix}\text{CAT} \begin{bmatrix}\text{HEAD}_{verb}[\text{VFORM } fin] \\ \text{SUBJ } \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS } \langle\,\rangle \\ \text{ARG-S } \langle \boxed{1}_{NP}:\boxed{3} \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix}\text{NUC}_{act\text{-}und\text{-}rel} \begin{bmatrix}\text{ACT } animate_{nom-obj} \\ \text{UND } \boxed{3}_{nom-obj}\end{bmatrix} \\ \text{EVSTR } \boxed{4} \begin{bmatrix}\text{E1} & happening \\ \text{E2} & state \\ \text{RESTR} & \langle \\ \text{HDD} & e1 \end{bmatrix}_{achiev\text{-}ev} \\ \text{QUAL}_{formal} \begin{bmatrix}\text{EVTYPE } e1 \\ \text{ALIST } \langle \boxed{3} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 4.10: Lexical Entries for (a) Accomplishment and (b) Achievement *bake*

4.2 The Polysemous Analysis of *miru*

One of the forms of verbal polysemy discussed by Pustejovsky involves the selection of alternative complements. He cites 'begin' as an example, as it takes either an NP, an infinitive VP or a gerundive VP as complement as shown in examples (21) to (23).

(4.21) Napoleon's army began the long march.

(4.22) The men began to desert.

(4.23) The stragglers began dying of the cold.

The Japanese verb *miru* (see) shows similarities in that it takes an accusative NP or a *te* marked VP as complement, as shown in examples (1) and (2) repeated here as (24) and (25) and I call these simplex and conjunct forms respectively.

(4.24) Sugiyama-san ga keeki o miru.
Mr Sugiyama-NOM cake-ACC see
"Mr Sugiyama sees the cake".

(4.25) Sugiyama-san ga keeki o tukutte miru.
Mr Sugiyama-NOM cake-ACC bake-TE try
"Mr Sugiyama tries baking a cake".

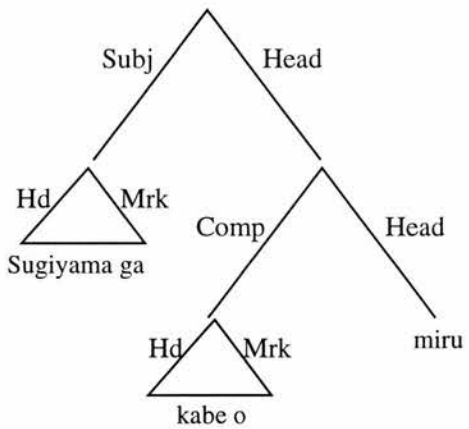
In previous sections I have introduced both Pustejovsky's semantic content, and Davis' semantic relations, which are based on a synthesis of Dowty's (1991) proto-roles and Foley & Van Valin's (1984) macro-roles of *actor* and *undergoer*. I have also introduced the HPSG argument structure feature, ARG-S, which acts as a sort of interface between CATEGORY and CONTENT fields, differing from Pustejovsky's conception of argument structure which defines semantic roles and is therefore equivalent to the HPSG NUCLEUS. The lexical entry for simplex *miru* is developed in Section 4.2.1, based on the lexical entry for transitive 'bake', given in the previous section. The challenge in the analysis of conjunct *miru* involves representing the selectional restrictions for accusative NP and *te* marked VP complements in a single underspecified form; this is shown in Subsection 4.2.2, and the integration of Pustejovsky into HPSG, together with the encoding of the polysemy analysis, are developed in 4.2.3.

4.2.1 Simplex miru

The simplex form of *miru* combines with an atelic adjunct *1-jikan* (for an hour) in (26), but not with a telic adjunct *1-jikan de* (in an hour), so *miru* is classified as an Activity.

(4.26) Sugiyama-san ga 1-jikan/*1-jikan de kabe o miru.
Mr Sugiyama-NOM for an hour/*in an hour wall-ACC look at
"Mr Sugiyama looks at the wall for an hour/*in an hour".

Figure 4.11: Constituent Structure of *miru*



The constituent structure of this verb is shown in Figure 4.11, in which *miru* projects a *ga* marked NP as subject and an *o* marked NP as complement. Before considering extensions for aspect, or for the conjunct analysis, the lexical entry for *miru*, given in Figure 4.12, is that of a finite (VFORM *fin*), transitive verb, with actor (ACT) and undergoer (UND) arguments projected to SUBJ and COMPS, following Davis's (1995) linking theory. Both ACT and UND roles are nominal objects, in a NUCLEUS of type *act-und-rel.* Japanese has explicit case marking, and the SUBJ is an NP marked nominative, while the COMPS contains an accusative marked NP. The verb subcategorizes for complete *synsem* objects, indexed $\boxed{1}$ and $\boxed{2}$; the argument structure (ARG-S) feature is the locus of linking between CAT and CONT, and the first element is linked to the ACT role ($\boxed{3}$), while the second element is linked to UND ($\boxed{4}$).

Combination with the Generative Lexicon framework allows for aspectual constraints to be realized and this can be done by extending CONTENT to contain *event structure* and *qualia structure*, and to map the *argument structure* field into the HPSG NUCLEUS. Aspect is encoded in Pustejovsky's Event Structure (as modified for compatibility with Dowty's model) and this is introduced as feature EVSTR

Figure 4.12: *miru à la* Davis (1995)

$$\begin{bmatrix} \text{PHON} & \langle miru \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD}\ _{verb}[\text{VFORM}\ fin] \\ \text{SUBJ} \quad \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS} \quad \langle \boxed{2}[\text{MKNG ACC}] \rangle \end{bmatrix} \\ \text{ARG-S} & \langle \boxed{1}\text{NP}{:}\boxed{3},\ \boxed{2}\text{NP}{:}\boxed{4} \rangle \\ \text{CONT} & \begin{bmatrix} \text{NUC} & \begin{bmatrix} \text{RELN}\ miru \\ \text{ACT}\ \boxed{3}_{nom-obj} \\ \text{UND}\ \boxed{4}_{nom-obj} \end{bmatrix} \\ act\text{-}und\text{-}rel & \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

under CONTENT, at the same level as the *nucleus* (NUC). As *miru* is an Activity, its Event Structure is that of a mono-eventual, agentive verb, of aspectual subtype *act-ev*. The E1 event represents a *process*. The Qualia Structure of an Activity includes an *agentive* quale, which selects for an *e1* headed event, and the list of projected roles, encoded here as ALIST, with roles structure shared with ACT and UND, licensing the projection of SUBJ and COMPS. The *animate* and *physobj* labels are mapped into ACT and UND roles as selectional restrictions on the arguments.

$$\begin{bmatrix} \text{PHON} & \langle miru \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD}\ _{verb}[\text{VFORM}\ fin] \\ \text{SUBJ} \quad \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS} \quad \langle \boxed{2}[\text{MKNG ACC}] \rangle \end{bmatrix} \\ \text{ARG-S} & \langle \boxed{1}\text{NP}{:}\boxed{3},\ \boxed{2}\text{NP}{:}\boxed{4} \rangle \\ \text{CONT} & \begin{bmatrix} \text{NUC} & \begin{bmatrix} \text{RELN} \quad miru \\ \text{ACT}\ \boxed{3}[\text{SEL-R}\ animate] \\ \text{UND}\ \boxed{4}[\text{SEL-R}\ physobj] \end{bmatrix} \\ act\text{-}und\text{-}rel & \\ \text{EVSTR} & \begin{bmatrix} \text{E1} \quad process \\ \text{HDD}\ e1 \end{bmatrix} \\ act\text{-}ev & \\ \text{QUAL} & \begin{bmatrix} \text{EVTYPE}\ e1 \\ \text{ALIST} \quad \langle \boxed{3},\ \boxed{4} \rangle \end{bmatrix} \\ agentive & \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 4.13: *miru à* la Pustejovsky (modified)

The lexical entry for *miru*, extended for aspectual considerations, is given in Figure 4.13. This differs from the Accomplishment lexical entry for English 'bake' in its event structure: *act-ev*, with a single *process* subevent, versus the bi-eventual Accomplishment; and in the fact that Japanese has explicit case marking on NPs, while (apart from pronouns) English does not.

4.2.2 The Conjunctive Form of miru

The data of *miru* shows that its complement can be an accusative nominal, as above, or a *te* marked agentive verb. Given that the semantic undergoer in Davis' scheme can be underspecified for *content* type, it seems possible that *miru* is polysemous in the sense of Levin (1993), on the basis of its argument taking behaviour. However, in order to justify such an analysis within the HPSG framework, the accusative nominal complement, and the *te* marked verb, must be selectable as complements of *miru* on a common basis. Following a reprise of the conjunct data, I argue for the viability of a polysemous approach.

In its conjunctive form *miru* selects a verbal argument which is agentive. That is to say, it combines with an Activity (27) or an Accomplishment (28), but not with an Achievement (29), or a State. In conjunct form, *miru* translates as something like "to sample something and see", although it is often glossed as "try", a usage which is deprecated in Martin (1975). There is a rough parallel between simplex and control forms of the verb 'see' in English and the simplex and conjunct forms of *miru*, because one can "see to doing something" or "see about doing something", but the event ontology is reversed, and conjunct 'see' indicates intention to do, while conjunct *miru* denotes completion and checking. Although it is not a perfect translation, I will continue to use the gloss 'try V-ing' for *te miru*.

(4.27) Sugiyama-san ga odotte miru.
Mr Sugiyama-NOM dance-TE try
"Mr Sugiyama tries dancing".

(4.28) Sugiyama-san ga keeki o tukutte miru.
Mr Sugiyama-NOM cake-ACC bake-TE try
"Mr Sugiyama tries baking a cake".

(4.29) *Mishima-san ga tennosama ni natte miru.
Mr Mishima-NOM emperor-DAT become-TE try
"*Mr Mishima tries becoming the emperor".

In combination with an Accomplishment, as in (30), the *te miru* form is compatible with a telic, not an atelic adjunct, suggesting that the adjunct is compatible with the first conjunct[3]. Adjuncts scramble freely anywhere before the verb. Example (31) shows that it is not actually possible to get a reading where the adjunct is compatible with the second conjunct, as the conjunction cannot be 'split'[4].

[3]In combination with an Activity, the *te miru* form is compatible with an atelic adjunct, but this is not evidence for anything, since both verbs are Activities.

[4]Though *morau* is an exception with respect to adjuncts, see Matsumoto (1996).

(4.30) Keiko ga *1-jikan/1-jikan de keeki o tukutte miru.
Keiko-NOM *for an hour/in an hour cake-ACC bake-TE try
"Keiko tries baking a cake in an hour".

(4.31) *Keiko ga keeki o tukutte 1-jikan/1-jikan de miru.
Keiko-NOM cake-ACC bake-TE *for an hour/*in an hour tries

Conventionally, the finite verb heads the sentence, the non-finite verb is subordinate, and the question of the missing subject of the subordinate verb arises. In an LFG account of *te morau* and other related constructions, Matsumoto (1996) shows that these are *control* phenomena, and the embedded verb is an XCOMP construction which has semantic *actor* and *undergoer* arguments (in the transitive case), with the *undergoer* expressed as accusative and the *actor* linked to the controlling subject. The *control* aspects of this phenomenon are reviewed in the next Section, while the lexical structure of the embedded *te* form is given here.

If indeed the *te miru* constructions in (27) and (28) are control constructions, then the *te* forms of *odotte* (dancing) and *tukutte* (baking) are verb phrases with an unsaturated subject: the finite form *odoru* (dance) projects a single argument linking the actor (ACT) role to SUBJ; nonfinite *odotte*, shown in Figure 4.14(a), retains the actor as a semantic role, but this is not projected. Thus, the argument structure (ARG-S) list of *odotte* contains a single pronominal, and the SUBJ value is a non-empty list. As a transitive Accomplishment, *tukuru* projects its actor to subject, and its undergoer to complement; non-finite *tukutte*, shown in 4.14(b), retains the saturated complement, and the ACT role is not projected. The ARG-S list of *tukutte* contains a pronominal element, and an NP linking undergoer (UND) to COMPS. The arguments in these verbs are licensed by Qualia Structure: in addition to the linking of nuclear roles to syntactic arguments, *odotte* as an Activity has mono-eventual Event Structure of type *process*, an *agentive* quale which licenses *e1* as the headed event, and an argument list (ALIST) containing the single *animate* role. As an Accomplishment, *tukutte* has bi-eventual Event Structure, with *agentive* and *formal* qualia. This form is headed *e1*, with the *agentive* quale having an ALIST containing both *animate* and *physobj* roles.

4.2.3 Selecting for Polysemous COMPS

From the data, we can see that *miru* is associated with an accusative marked NP and a *teform* VP as possible complements. In the previous section I showed that *miru* subcategorizes for its nominal complement on the basis of nounhood and a marking value of accusative; the valences of this subcategorized NP are both saturated (SUBJ

(a) *odotte*:

$$\begin{bmatrix}\text{PHON} & \langle odotte \rangle \\ \text{SYNSEM} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD } teform[\text{VFORM } teverb] \\ \text{SUBJ} & \boxed{1}pro \\ \text{COMPS} & \langle\,\rangle\end{bmatrix} \\ \text{ARG-S} & \langle \boxed{1}{:}pro \rangle \\ \text{CONT} & \begin{bmatrix}\text{NUC} & [\text{ACT } animate] \\ \text{EVENTSTR}_{act\text{-}ev} & \begin{bmatrix}\text{E1} & process \\ \text{HDD} & e1\end{bmatrix} \\ \text{QUALIA} & \begin{bmatrix}\text{EVTYPE } e1 \\ \text{ALIST} \quad \langle animate \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

(b) *tukutte*:

$$\begin{bmatrix}\text{PHON} & \langle tukutte \rangle \\ \text{SYNSEM} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD } teform[\text{VFORM } teverb] \\ \text{SUBJ} & \boxed{1}pro \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG } acc] \rangle\end{bmatrix} \\ \text{ARG-S} & \langle \boxed{1}{:}pro,\ \boxed{2}\text{NP}{:}\boxed{3} \rangle \\ \text{CONT} & \begin{bmatrix}\text{NUC} & \begin{bmatrix}\text{ACT } animate \\ \text{UND } \boxed{3}physobj\end{bmatrix} \\ \text{EVENTSTR}_{accomp\text{-}ev} & \begin{bmatrix}\text{E1} & process \\ \text{E2} & state \\ \text{RESTR} & partord \\ \text{HDD} & e1\end{bmatrix} \\ \text{QUALIA} & \begin{bmatrix}\text{EVTYPE } e1 \\ \text{ALIST} \quad \langle animate,\ \boxed{3} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 4.14: Lexical Entries for (a) *odotte* and (b) *tukutte*

$\langle\rangle$, COMPS $\langle\rangle$). By the *Semantics Principle* of HPSG, the undergoer (UND) role is structured with the semantic content of the noun, which is a nominal object (*nom-obj*). The subcategorized verb (*odotte* (dancing), or *keeki o tukutte* (baking a cake)) projects its complements: none in the case of unergative *odotte*, *keeki o* (a cake) in the case of *tukutte*, so the COMPS value is saturated. The subject is not projected, however, and therefore remains unsaturated (SUBJ *ne-list*). One function of the *te* affix is to suppress the projection of the subject[5]. If the *te* form is a nominalized VP, then a further function of *te* is to provide a marking value, equivalent to accusative case. In fact I will show in the next chapter that the *te* form is underspecified for

[5]Alternatively, since infinitival verbs (*odori*, *tukuri*) are also subject unsaturated, it can be claimed that one function of finite inflection (*odoru*, *tukuru*) is to license the subject projection.

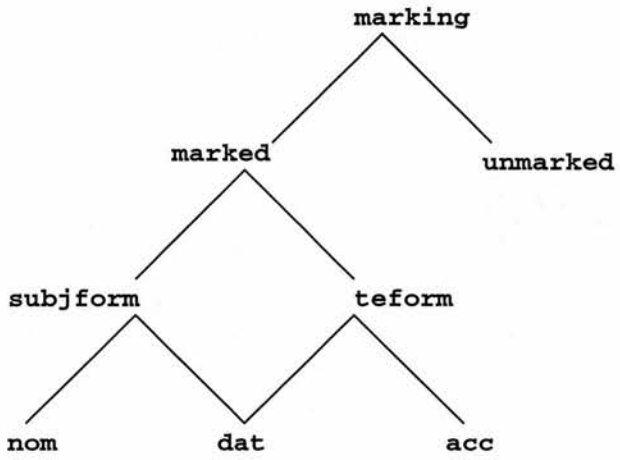


Figure 4.15: The *marking* Hierarchy

both accusative and dative cases, in virtue of its participation in the *V-te iru* and *V-te aru* constructions. Also in virtue of the fact that subjects of *iru* (animate be) and *aru* (inanimate be) can take either *ni* (dative) or *ga* (nominative) marking, *subjform* is the supertype of *nom* and *dat* types. Following these considerations, the marking hierarchy is given in Figure 4.15.

Semantically, the *te* form VP is a parameterized state of affairs (*psoa*), giving features which differ from a nominal object. The specific differences between noun (or NP) and VP are as follows:

Category	noun	verb
Marking	acc	teform
Comps	saturated	saturated
Subj	saturated	unsaturated
Content	nom-obj	psoa

The fact that the *te* form verb has (for transitive verbs at least) COMPS arguments with conventional case marking suggests that it is indeed of category verb. However the fact that it alternates with NP as a complement of *miru* suggests that it might be, at least, a nominalized VP. For such a nominalization to occur, there must be some commonality between nouns and verbs which distinguish them from other substantives. Suppose, then, that there is a category VN which is the supertype of nouns and verbs. How would such an analysis be technically possible?

The sort hierarchy for type *head* given in Pollard and Sag (1994, p. 396) includes:

Partition of *head*: *subst*, *func*

Partition of *subst*: *noun*, *verb*, ... *prep*, ...

Partition of *func*: *mark*, *det*

Allowing for the fact that Japanese has postpositions, not prepositions, and only case markers as functional categories, then if nouns and verbs are a class with a common parent, the type hierarchy is:

Partition of *head*: *subst*, *mark*

Partition of *subst*: *vn*, *postp*

Partition of *vn*: *allnom*, *allverbal*

The common parent of nouns and verbs is a category *vn*. While this might seem to set up the possible unrestrained subcategorization of one verb by another, this is in practice constrained by the selectional restrictions on complements, which require features inherited by nominals only. If the *teform* is a nominalized VP then it inherits features from both nominal and verbal parents, thus:

Partition of *allnom*: *noun*, *teform*

Partition of *allverbal*: *teform*, *verb*

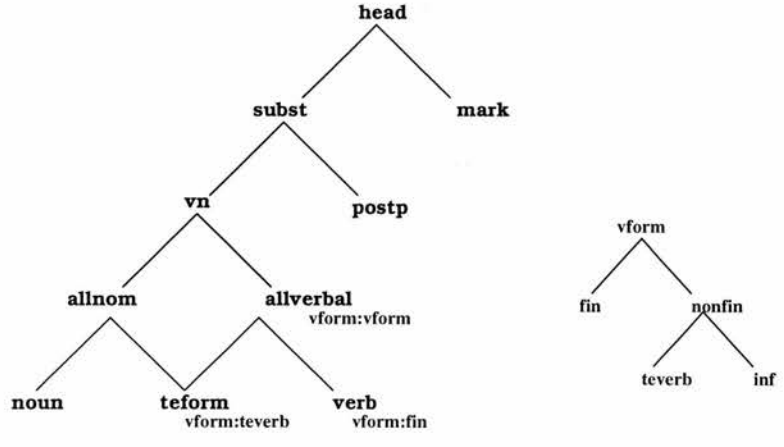


Figure 4.16: Category Hierarchy

The category hierarchy is amended as in Figure 4.16: *vn* is a direct subtype of *subst* and a supertype of *allnom* and *allverbal*; verbals have feature VFORM, which has values finite (*fin*) and non-finite (*nonfin*). The *teform* inherits from *allnom* and *allverbal*, and the value of VFORM inherited by the *teform* is restricted to *teverb*. The marking value of nouns and teforms is a feature of the complex type *cat*: nouns

have a marking value of *none* but in combination with the Head-Marker phrase and marker *o*, the CAT value of the phrasal mother (NP) takes on accusative marking[6]. The lexical entry for the *te* form is also associated with accusative marking.

Semantically, *keeki o* in (32) is a nominal object, and *keeki o tukutte* in (33) is a parameterized state of affairs (*psoa*).

(4.32) Sugiyama-san ga keeki o miru.
Mr Sugiyama-NOM cake-ACC sees
"Mr Sugiyama sees the cake".

(4.33) Sugiyama-san ga keeki o tukutte miru.
Mr Sugiyama-NOM cake-ACC bake-TE try
"Mr Sugiyama tries baking a cake".

Pollard and Sag's (1994, p. 397) type hierarchy for *content* types includes:

Partition of *content*: *psoa*, *nom-obj*, ...

Again, if the *teform* can inherit from both nominal and verbal parents, then the undergoer of *miru* can select for the common supertype, thus:

Partition of *content*: *eventuality*, *allpsoa*

Partition of *eventuality*: *nom-obj*, *tepsoa*

Partition of *allpsoa*: *tepsoa*, *psoa*

For this to work, *miru* must select for an underspecified set of selectional restrictions compatible with the features of a *nom-obj* and those of a *psoa*, and these features must therefore be present in the *content* type hierarchy, under *eventuality* (See Figure 4.18).

When *miru* selects for a nominal object there is a selectional restriction associated with the role: the actor is an *animate* form and the undergoer a *physobj* - or indeed some metaphorically 'visible' object. When *miru* selects for a *te* form, the selectional restriction is associated with the aspectual class of the subcategorized verb. For *miru* to select either one, then the same set of selectional restrictions must select for either the nominal or the *te* form argument, so their *content* types must be

[6]Given that postpositional case markers in Japanese are associated with a MARKING value under CAT, following Tomabechi (1989), it is not therefore necessary for nouns to have a *head* feature, CASE.

compatible and the lexical structures of each must contain both the *animate/physobj* type of selectional restriction (let us call this the *phenomena* type) as well as the Event Structure type.

The *phenomena* hierarchy includes nominal and verbal phenomena, and the nominals include *physobj* and *animate*, with the latter as a subtype of the former. The hierarchy of 'things' includes not only physical, but conceptual objects: a 'dream' or a 'collection', for example, so these are all subtypes of *realization*. Phenomena also include *events*, so a rudimentary hierarchy for all *phenomena*, looks like Figure 4.17.

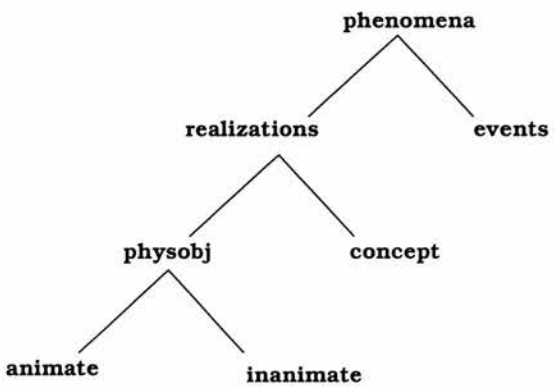


Figure 4.17: Hierarchy of Selectional Restrictions for *phenomena*

When *miru* selects for a nominal object, the *phenomena* is specified and the event structure is underspecified. The Event Structure hierarchy is articulated in Figure 4.9 in the previous section, with examples of *odoru/odotte* (dance/dancing) as an Activity, typed *unerg-ev*, and *tukuru/tukutte* (bake/baking) as an Accomplishment, type *accomp-ev*. The maximal supertype is *aspect-rel*, covering all aspect classes.

On the one hand Pollard & Sag (1994) defines the *content* type of a nominal to be of type *nom-obj*; on the other hand, the *content* type of a verb includes a NUCLEUS with the relations between the arguments and the predicate[7]. If polysemous verbs subcategorize for a common semantic object, then the alternate content forms of noun and verb must be unified in the *content* type hierarchy: verbs and their *te* forms are grouped together as *psoas* containing selectional restrictions (SEL-R), event structure (EVSTR), NUCLEUS and qualia structure (QUALIA); *teforms* and *nom-objs* are grouped together as subcategorizands containing SEL-R and EVSTR, and in this latter case they are subtypes of a common type, let us call it *eventuality*,

[7]Again, this is omitting the quantificational content of both nouns and verbs.

from which they inherit the features constituting selectional restrictions: so the *tepsoa* inherits from both *eventuality* and *psoa*, as shown in Figure 4.18.

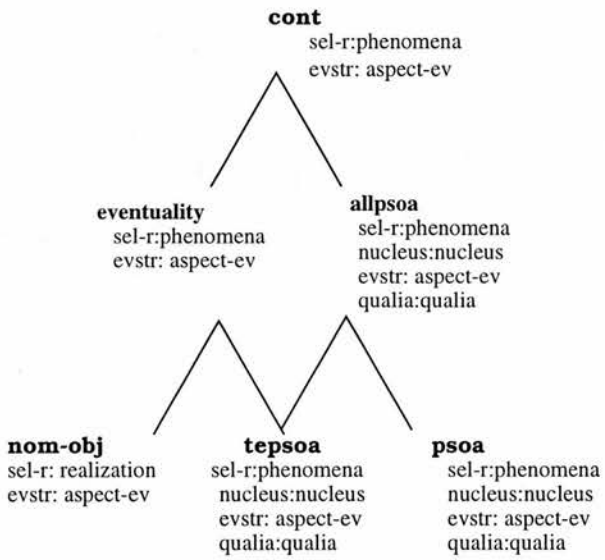


Figure 4.18: The *cont* Hierarchy

With the scheme outlined above, the semantic content of the subcategorized elements is represented as in Figure 4.19, with selectional restrictions (SEL-R) and event structure (EVSTR) as features under CONTENT[8].

(a):
$$\begin{bmatrix} \text{PHON} & \textit{miru} \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{CONT} & \begin{bmatrix} \text{SEL-R} & \textit{phenom} \\ \text{EVSTR} & \textit{aspect-ev} \end{bmatrix}_{\textit{eventuality}} \end{bmatrix} \right\rangle \end{bmatrix}$$

(b):
$$\begin{bmatrix} \text{PHON} & \textit{keeki} \\ \text{CONT} & \begin{bmatrix} \text{SEL-R} & \textit{physobj} \\ \text{EVSTR} & \textit{aspect-ev} \end{bmatrix}_{\textit{nom-obj}} \end{bmatrix}$$

(c):
$$\begin{bmatrix} \text{PHON} & \textit{tukutte} \\ \text{CONT} & \begin{bmatrix} \text{SEL-R} & \textit{events} \\ \text{EVSTR} & \begin{bmatrix} \text{E1} & \textit{process} \\ \text{E2} & \textit{state} \\ \text{RESTR} & \langle \\ \text{HDD} & \textit{e1} \end{bmatrix}_{\textit{accomp-ev}} \end{bmatrix}_{\textit{tepsoa}} \end{bmatrix}$$

Figure 4.19: Underspecified *eventuality* on the Head Verb (a) , selecting for Selectional Restrictions on (b) a Noun and (c) a *te* Form Verb

The underspecified CONTENT value of *eventuality* is shown in Figure 4.19(a). Both *nom-obj* and *tepsoa* contain more restricted values of this feature complex. A *nom-obj* subcategorized as complement has a SEL-R value of *physobj*, and a maximally

[8]It seems likely that the *phenomena* values of SEL-R are actually a contextual background feature, rather than CONTENT. However for geometric convenience I will continue to place this feature under CONTENT, for the purposes of this thesis.

underspecified EVSTR value of *aspect-ev*, and this is shown in Figure 4.19(b). Figure 4.19(c) gives the content value of a *tepsoa*, with SEL-R value of *events* and an EVSTR value specifying a bi-eventual structure, consistent with that of an Accomplishment.

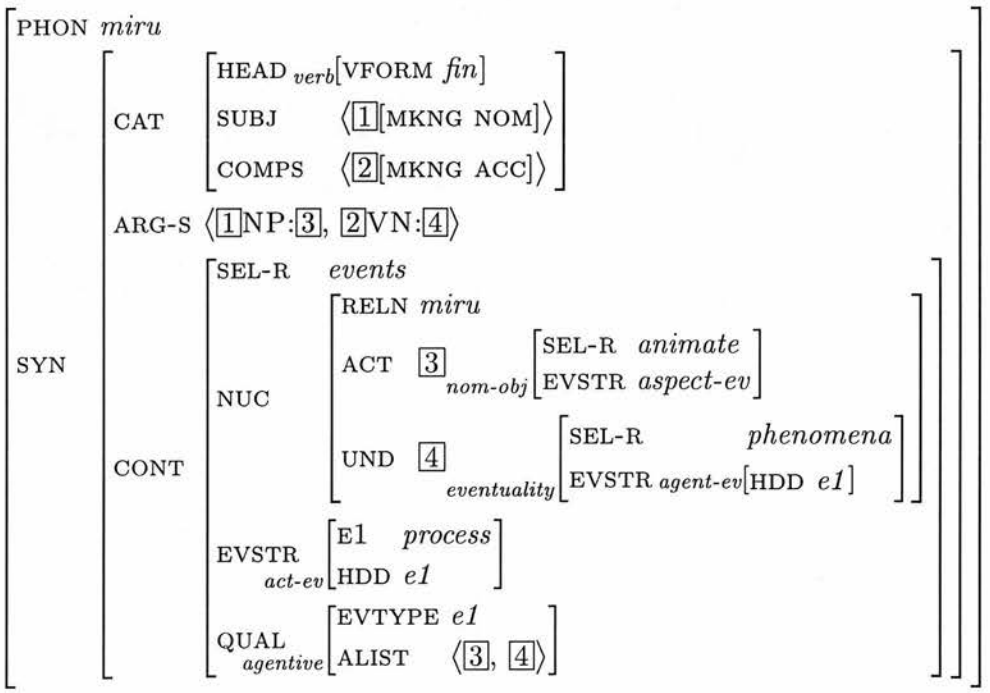


Figure 4.20: Lexical Entry for *miru*

The lexical entry for *miru* is amended in Figure 4.20 to account for its polysemous combination with these lexical phenomena. The undergoer (UND) role is of type *eventuality*, selecting for an underspecified SEL-R of type *phenomena*, which captures either *physobj* or *events* (see Figure 4.18) and an EVSTR of type *agent-ev*, capturing either an Accomplishment (*accomp-ev*) or an Activity (*act-ev*), or alternatively unifying with the maximally underspecified *aspect-ev*, specified on a *nom-obj* (See Figure 4.9).

There is a single remaining problem before the polysemy analysis is complete: when *miru* subcategorizes for an NP complement, the subject value of that NP is saturated; however when *miru* subcategorizes for a *teform* VP, the subject value is unsaturated. If *miru* leaves the subject value of its *comps* argument underspecified (PHON *miru*, COMPS ⟨VN[SUBJ *list*]⟩), then in addition to accepting the correct form, the incorrect (34) is also accepted, in which *teform* (*tukutte*) and finite verb (*miru*) license separate subjects.

(4.34) *Keiko ga Maiko ga keeki o tukutte miru.
Keiko-NOM Maiko-NOM cake-ACC bake-TE tries

The non-finite form, under Nuclear Conjunction, never licenses a saturated subject. One solution which enables this projection to be blocked is if the Subject-Head rule is admitted only when the head daughter has VFORM *fin*.[9]

4.2.4 Phrasal Structures for miru

The phrasal structures resulting from the simplex and conjunct verb phrases of examples (32) and (33), respectively, are developed in Figures 4.21 and 4.22. In Figure 4.21, the selectional restrictions of *miru* unify with those of the nominal complement *keeki o* (a cake), with CONTENT indexed [4]. The SEL-R value of *physobj* on the noun is a subtype of *phenomena* selected by *miru*, and the event structure *aspect-ev* is a supertype of *agent-ev*, so again they unify.

The projection of the VP complement is given in the complex phrasal structure of Figure 4.22. The CONTENT of the head verb *miru* is indexed [5], and that of its complement VP, *tukutte*, is indexed [4], the latter coindexed with the *undergoer* of *miru*. In the selectional restrictions hierarchy, *miru* selects for type *phenomena*, and this is matched by the SEL-R value of *tukutte*; *miru* also selects for an event structure of type *agentive-ev*, headed E1, and this unifies with the more specific *accomp-ev*, E1 headed event structure of *tukutte*.

4.2.5 Summary of miru Related Phenomena

In this section I have shown that the polysemous behaviour of the Japanese verb *miru* (see), with simplex and *te* marked conjunct forms, can be captured in a single

[9]The Core and Clausal constructions (examples 2.1 and 2.2) involve a *teform* which is a subordinate/non-main clause, which attaches as an adjunct of the main clause, and in this way they differ from Nuclear Conjunction. In order to allow for the analysis of these constructions, the Subject-Head rule is associated with a goal as follows:

> **Nonfinite Subject Licensing Principle**
>
> Finite verbs combine with a subject, and they are never adjuncts, sothey do not have a MOD value, formally: If the value of VFORM on the Head-Daughter is *fin* then its MOD value is *none*;
>
> If the *teverb* is a complement, it does not have a MOD value, otherwise, the *teverb* is an ajunct, bearing a MOD value in which it modifies a VP or an S. Formally, if the value of VFORM on the Head-Daughter is *teverb* then its MOD value is V[SUBJ *list*, COMPS *elist*].
>
> The Complement-Head rule is also associated with a restriction, that complements do not bear a MOD value.
>
> In this way, the Nuclear Conjunct form can never be analysed as an adjunct of the head verb, since it is prevented from attaching to a verb with unsaturated complement.

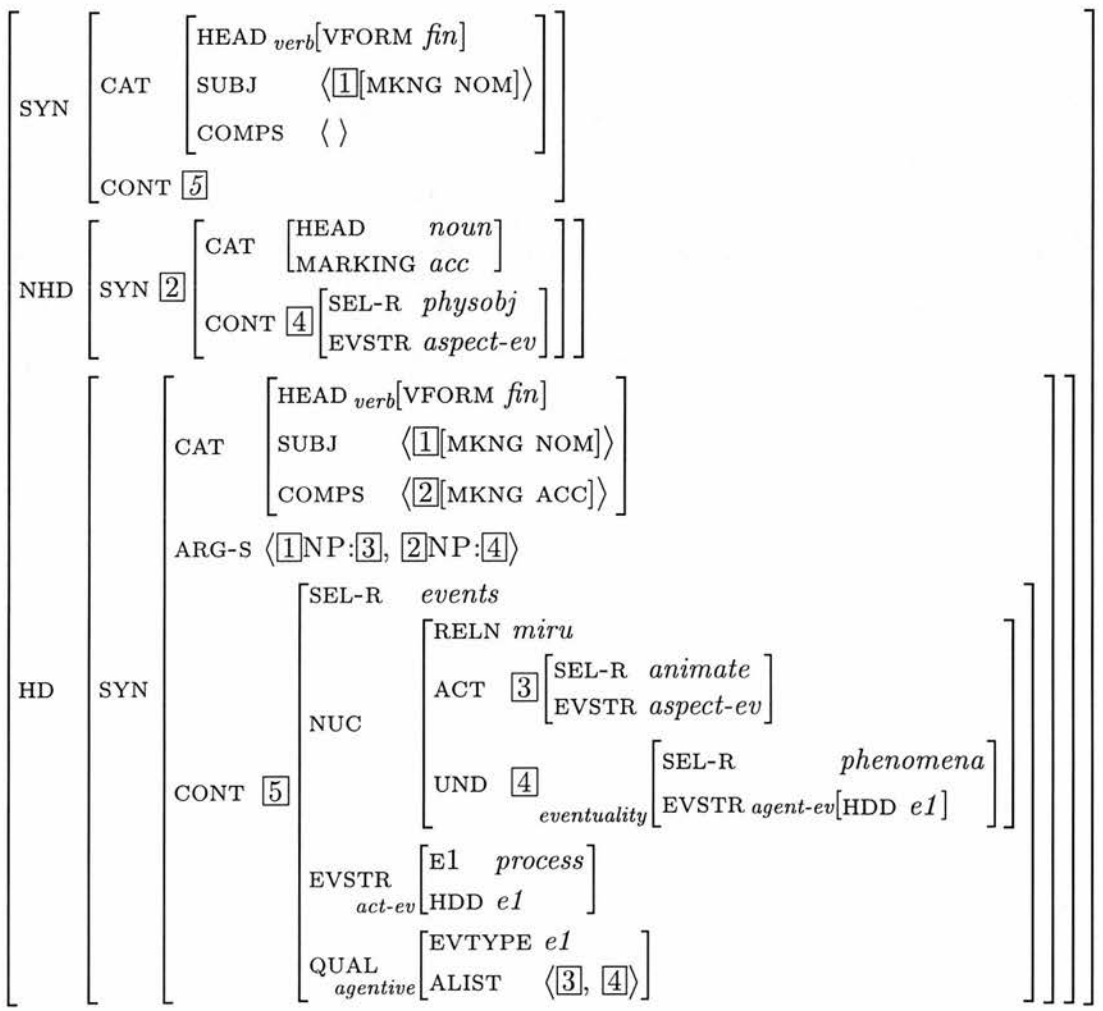


Figure 4.21: keeki o miru

lexical entry, by underspecifying the complement, and its undergoer semantic role. The phenomena of *te miru* and its related constructions *te morau* and *te ageru* exhibit some of the behaviour of *control* constructions, in which there is a debate about whether the unexpressed subject of the embedded verb is, or is not, overtly linked with some argument of the higher predicate. The outcome of this debate has a critical bearing on whether the *te* constructions are, or can be, polysemous. If there is an overt link in the conjuncts of *miru* et al, and no such link can occur in the simplex forms, then the polysemy analysis is called into question. A review of the debate on *control* is called for, and this is prefigured by the introduction of the Ishikawa/Matsumoto analysis of *te morau*.

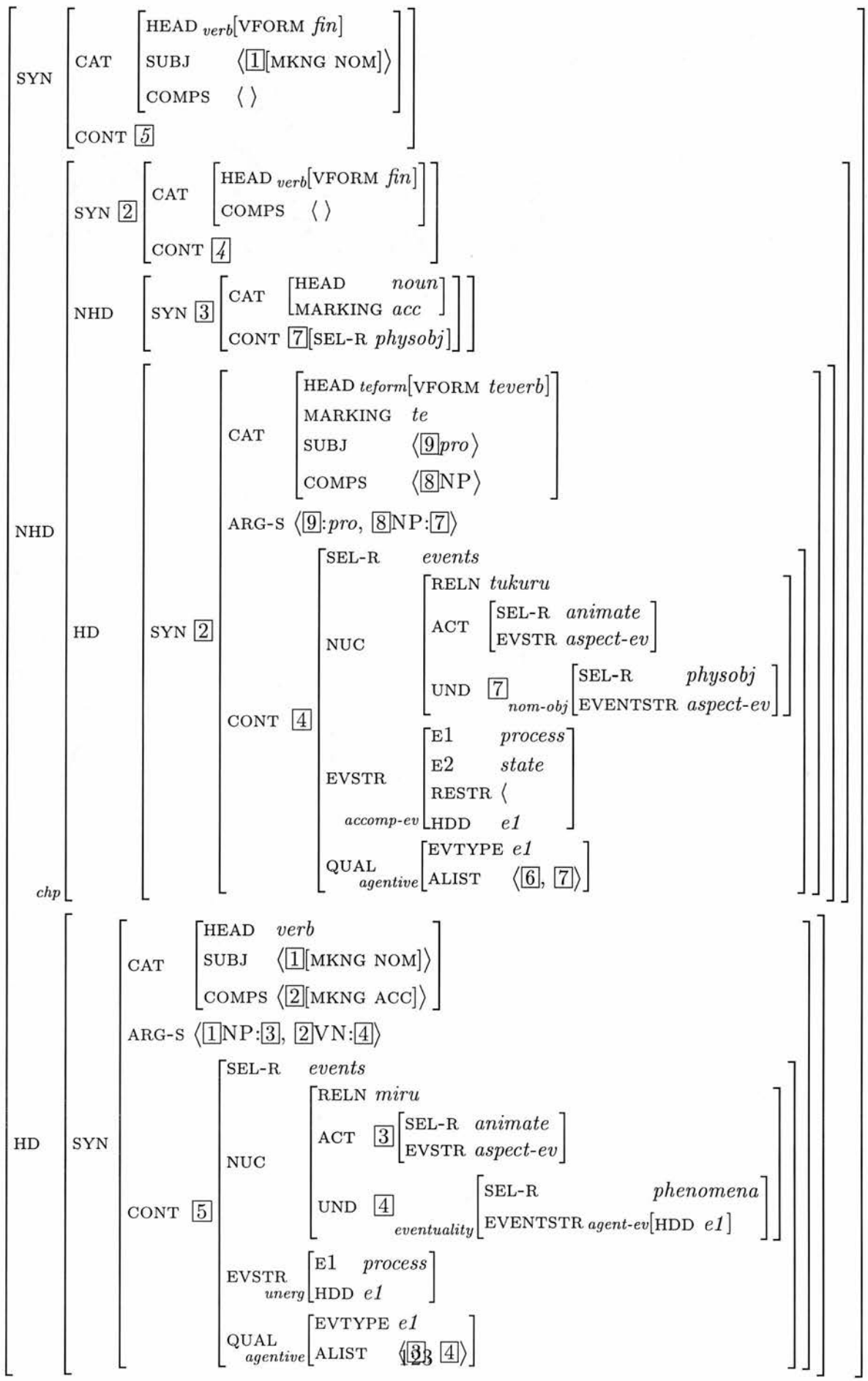


Figure 4.22: keeki o tukutte miru

4.3 Control Analyses

When verbs are combined together in a conjunctive form, the subject of the embedded verb is suppressed, but the semantic role is usually in some way compatible with the expressed subject of the higher verb[10]. The question arises of what, if anything, is the relationship between the subject of, say, *miru* and the subject of the subcategorized V1 verb. These *te* conjunctions, in fact, actually bear more than a superficial similarity to a class of linguistic phenomena which has been called 'Control Constructions', following Rosenbaum (1967), and these are exemplified by the verbs 'try' and 'persuade'. In example (35), 'try' takes an argument which may be gerund or infinitival, and the outstanding question is whether the unexpressed subject of the complement verb is overtly structure-shared with the controlling subject in the case of 'try' (35), or object in the case of 'persuade' (36). Pollard & Sag (1994) follow the view of Jackendoff (1972) and Jackendoff (1974) that the unexpressed subject is coindexed with the controlling subject in what they term the "traditional" view of control.

(4.35) I tried to eat haggis/eating haggis

(4.36) I persuaded you to go home.

In a previous analysis of the *te morau* construction in Lexical Functional Grammar, Ishikawa (1985) repeated by Matsumoto (1996), claims that the embedded verb is an *XCOMP* construction[11] and its subject is structure shared with the oblique argument of *morau*, called an OBJSRC to indicate that it is the source of the transfer. Although Ishikawa looks at the *te ageru* construction, he produces no model theoretic analysis, nor does Matsumoto develop one. Matsumoto's (1996) structure for the example sentence in (37) is given in Figure 4.23, showing 'read' as an embedded predicate of *morau* and indicating the structure sharing between the Objsrc of *morau* and the subject of 'read'.

(4.37) John ga Mary ni hon o yonde morau
John-NOM Mary-DAT book-ACC read-TE receive
"Mary reads the book for John" (Matsumoto (1996)).

[10]This is so for Nuclear Conjunctions and also for Core Conjunctions. Clausal Conjunctions combine with all their arguments expressed, and are therefore their analyses are more related to coordinate structures.

[11]An XCOMP is an LFG (Bresnan (1982)) category with a 'missing' subject. The subject is controlled by an argument of an upper clause, and so *te morau* and *te ageru* (and all other *te* conjunctions) are control constructions.

$$\begin{bmatrix} \text{PRED} & \text{RECEIVE}\langle subj, objsrc, xcomp\rangle \\ \text{SUBJ} & [\text{PRED } John] \\ \text{OBJSRC} & [\text{PRED } \boxed{1}Mary] \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{READ}\langle subj, obj\rangle \\ \text{SUBJ} & \boxed{1} \\ \text{OBJ} & [\text{PRED } book] \end{bmatrix} \end{bmatrix}$$

Figure 4.23: Matsumoto/Ishikawa, LFG View of *morau*

Matsumoto's thesis is to identify structures which are morphologically and functionally words, and he concludes that *te morau* is a functionally biclausal, syntactic construction. Neither Ishikawa nor Matsumoto makes a connection between simplex and conjunct forms, and the question of polysemy never arises. Both authors presume that there is an overt link between the controlling and embedded predicates, but no strong argument is given for that position.

These conjunctive forms have another interpretation which has been available since the early days of research on Montague Grammar(Montague (1974)): English *control* verbs such as *try* are functions taking one place properties as arguments, and that the phenomenon of control is part of the entailments that a verb can carry. The evidence for this view is propounded in Chierchia (1984). In developing the Nuclear Conjunctive analyses I will situate the Japanese phenomena with respect to these traditional control constructions. There seems to be, intuitively, a rough correspondence between the *te miru* construction and the 'try V-ing' construction, although *te miru* does not translate directly as 'try V-ing', but as something more like "do V (and see what happens)". Thus, *te miru* is completive, where 'try V-ing' is not necessarily so. A similar correspondence can be drawn between 'want' and the desiderative form of *te morau*, which is *te moraitai* (want to receive), or indeed with the related *te hoshii* (want) construction. So in examples (38) and (39) the verbal *moraitai/hoshii* has a nominative subject, and dative object, and a *te* marked VP argument in the same pattern as 'want'.

(4.38) Midori ga Yasuko ni sara o katazukete moraitai.
Midori-NOM Yasuko-DAT dishes-ACC tidy up-TE receive-DESID
"Midori wants Yasuko to tidy away the dishes".

(4.39) Midori ga Yasuko ni sara o katazukete hoshii.
Midori-NOM Yasuko-DAT dishes-ACC tidy up-TE want
"Midori wants Yasuko to tidy away the dishes".

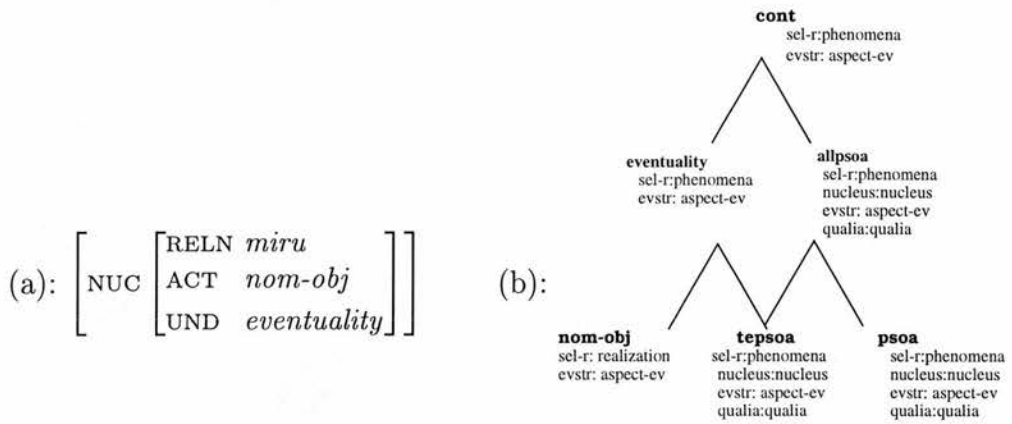


Figure 4.24: HPSG *nucleus* and Content Type Hierarchy

Because the analysis of *miru* in the last chapter shows it to be polysemous, the NUCLEUS field contains an ACTOR argument which is a *nominal-object*, projected to subject, and an UNDERGOER argument underspecified for its *content* type, and this *content* can be either a *nom-obj* (following Davis (1995)), projected to complement as an accusative object, or (following Pollard & Sag (1994)) or a *psoa* of an agentive verb projected as a *te* marked VP, following the last chapter. This NUCLEUS is shown in Figure 4.24(a). Since the *te* form of a verb and the nominal object are related constructions, and since *miru* (and any other V2 verb) selects either one polysemously, both forms must be related in the semantic content type hierarchy, and this is shown in Figure 4.24(b), with *tepsoa* as a subtype of both *nom-obj* and *psoa*, inheriting constraints from both types. This is in contrast with Pollard & Sag's (1994) theory of control, where there are three types of control verbs, defined by a semantic relation under the *content* field of *control-relation*, of which 'try' and ditransitive 'want' are examples of *commitment* and *orientation* types respectively. The third type of control relation is that of *influence*, with 'persuade' and 'force' as examples of this type. *Commitment*, *orientation* and *influence* types of control relation are exemplified in (40) to (42) respectively. In example (42), the commitment verb 'try' is subject controlled, with *Mr Sugiyama* as the subject of both 'try' and 'dancing'; in (41) the *orientation* verb 'want' is subject controlled, with *Shunsuke* as subject of 'want' and 'eat sushi'; and in (42) 'persuade' is a ditransitive object controlled verb of type *influence*, with *Yasuko* as complement of 'persuade', coindexed with the unexpressed subject of *dance*.

(4.40) Mr Sugiyama$_i$ tried ($_i$) to dance with Mai.

(4.41) Shunsuke$_i$ wants ($_i$) to eat sushi.

(4.42) Mr Sugiyama$_i$ persuaded Yasuko$_j$ ($_j$) to dance the foxtrot with him$_i$.

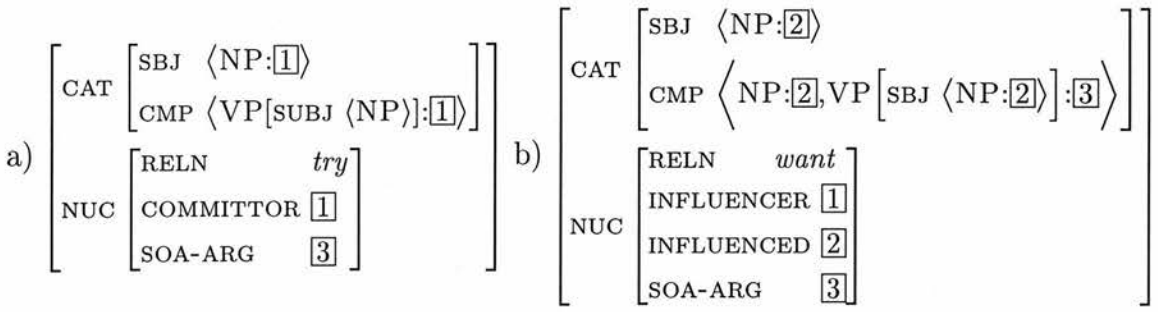


Figure 4.25: Control Relations for Commitment and Influence Types

While Pollard & Sag (1994) give a long list of control verbs of the *orientation* type, including *want*, *hate*, *expect*, *need*, *fancy*, *ache*, *itch*, *thirst* and *yearn*, what they don't say is that many of these verbs have a ditransitive form which is object-controlled, so they pattern after *commitment* verbs in their transitive forms (example (53)(a)-(i)), but after *influence* type verbs in their ditransitive forms (Example (43(j)-(r)).

(4.43)

(a) I want to go.	(j) I want you to go.
(b) I hate to go.	(k) I hate you to go.
(c) I expect to go.	(l) I expect you to go.
(d) I need to go.	(m) I need you to go.
(e) I fancy going.	(n) I fancy your going there.
(f) I ache to go.	(o) *I ache you to go.
(g) I itch to go.	(p) *I itch you to go.
(h) I thirst to go.	(q) *I thirst you to go.
(i) I yearn to go.	(r) *I yearn you to go.

The two control forms of *commitment* and *influence* types are associated with separate semantic structures. The unsaturated complement of a *commitment* verb such as 'try' in example (40) is represented lexically as SOA-ARG and its subject is coindexed with the COMMITTOR as shown in Figure (4.25(a)). For the *influence* relation, the unsaturated subject of the SOA-ARG is coindexed with the INFLUENCED complement of a verb such as 'persuade' in (42), so the NP complement is coindexed as the subject of the VP complement (e.g. *dance*) as shown in Figure (4.25(b)).

There are differences between the Japanese and the English control constructions, and these seem to flow partly from the analysis of the Japanese verbs as polysemous, and partly from differences in argument patterns, particularly with *te ageru* and related constructions. Under the polysemy analysis, *miru* has ACTOR and UNDERGOER arguments with the UNDERGOER underspecified for nominal and verbal

arguments, whereas the equivalent *commitment* control construction 'try' has COMMITTOR and SOA-ARG arguments, the latter strictly verbal, and the COMMITTOR is coindexed with the subject of the SOA-ARG. While it may be possible to argue for a similar coindexation between the ACTOR of *miru* and the subject of the verbal UNDERGOER in the conjunct case, no such link can be made in the simplex case, because the nominal UNDERGOER has no subject. In example (44), there is no suggested relation between *Midori* and any argument of *keeki*, and indeed it is *miru* which provides the direct relation between them, through the RELN field under NUCLEUS (Figure 4.20).

(4.44) Midori ga keeki o miru.
Midori-NOM cake-ACC see.
"Midori sees the cake".

While both 'try' and 'want' have broader argument taking behaviours than just the control readings, there is no suggestion in Pollard & Sag (1994) that any polysemy may be involved, and this can be inferred from the following examples. Example (45) is similar to (44) in that there is no suggested relation between 'Mr Sugiyama' and any argument of 'foxtrot', and so because the control form of 'try' exemplified in (40) analyses with a SOA-ARG, it is mutually homonymous with the form exemplified in (45), which analyses as an UNDERGOER. A similar argument applies to the transitive and ditransitive forms of 'want' in (46) and (47), except that while the simplex transitive 'want' in (46) is homonymous with respect to the control form in (52), there is no simplex ditransitive 'want', and (47) is ill-formed.

(4.45) Mr Sugiyama tried the foxtrot.
(4.46) Shunsuke wanted a beer.
(4.47) *Shunsuke wanted you sushi.

As I showed previously, Pustejovsky (1995) treats 'want' and also 'begin' as polysemous, positing that their arguments are coerced into the type expected by the verb, but he has no systematic treatment of syntactic projection, and therefore has no need to take a position on whether any arguments are overtly coindexed. It seems to be a reasonable consequence of Pustejovsky's position that relations between the arguments of a polysemous predicate are consistent, with no overt linkage, as above, and this suggests that Chierchia's (1984) account of control constructions supports the polysemy hypothesis.

4.3.1 Control as Entailment

In Chierchia (1984), the semantics of control constructions are considered as an extension to the semantics of infinitives and gerunds, and he reviews a number of hypotheses, including the view which derives from Montague Grammar that a verb such as 'try' which takes a VP argument is a function which takes a 1-place predicate, and this is the view which he supports. The idea is that a sentence such as (48(a)), has semantics as in 48(b)), and this is *Hypothesis A*.

(4.48) (a) $[\text{John}_{NP}\ [\text{tries}\ [\text{to run}]_{VP}]_{VP}]_{S}$.
(b) try'(run')(j)

Chierchia reviews a range of alternative views, including Bach & Partee (1980) (Infinitives and gerunds are semantically associated with propositions, and are semantically clausal); Klein & Sag (1982) (The unexpressed subject of the embedded nonfinite verb phrase is linked using a semantic rule); and (Chierchia's interpretation of) the standard GB position (an infinitive is a sentence with a pronominal element which is coindexed with the controlling subject). In all these cases the semantics of example (48) is a proposition in which the embedded predicate takes an argument identical with that of the controlling predicate, viz: try'(run'(j))(j), and I will collectively identify their common semantic position as *Hypothesis* Ω. The problem with this approach is that it leads to manifestly unintuitive results under discourse entailments such as the following:

(4.49) (a) Nathalie tries everything Nadia tries.
(b) Nadia tries to kiss the Dalai Lama.
(c) Nathalie tries to kiss the Dalai Lama.

Under Hypothesis A the semantics of this syllogism model as (50) and 'try' is semantically a relation between individuals (e.g. Nadia) and properties (e.g kissing the Dalai Lama):

(4.50) (a) $\forall\alpha try'(\alpha)(Nad) \longrightarrow try'(\alpha)(Nat)$.
(b) try'(kissing DL)(Nad).
(c) try'(kissing DL)(Nat).

Under *Hypothesis* Ω however, the semantics of (50(b)) are a relation between Nadia the subject of 'trying' and the proposition that "Nadia kissed the Dalai Lama"

and this entails that (50(c)) is interpreted as (51), which is absurd, and therefore Hypothesis Ω cannot be correct:

(4.51) (b) try'(kissing DL(Nad))(Nad).
"Nadia tries for Nadia to kiss the Dalai Lama".

and consequently (c) is:

(4.52) (c) try'(kissing DL(Nad))(Nat)
"Nathalie tries for Nadia to kiss the Dalai Lama".

It is therefore this *reductio ad absurdum* which Chierchia claims as support for Hypothesis A, with the consequence that there is no syntactic or semantic rule linking the unexpressed subject with anything else. Pollard & Sag (1994), however, claim that the facts of agreement provide a compelling counterargument over a broad range of languages, including English, French and German which have strong agreement morphology, and Japanese and Korean, which have honorific agreement morphology. The evidence from binding is offered as support for this view, because in example (53) the controlling subject 'Ollie' agrees with the complement of the embedded predicate 'perjure', and according to Principle A of the HPSG Binding Theory "a locally o-commanded anaphor must be locally o-bound" ((Pollard & Sag 1994)). Because 'Ollie' and 'himself' are in separate argument structure domains, and 'himself' is an anaphor on the lower argument structure, it follows that 'himself' must be anteceded by a local SYNSEM object with which it agrees, and this is offered as demonstration of the syntactic reality of a linked subject.

(4.53) Ollie$_i$ tried ($_i$) to perjure himself$_i$

This phenomenon can indeed be demonstrated for the Japanese *te miru* construction, thus in example (54) *Naoko* drives a car, and in (55) we see that it is her own car, as *Naoko* is coindexed with *jibun* (self).

(4.54) Naoko ga kuruma o unten shite miru.
Naoko-NOM car-ACC drive-TE see about
"Naoko sees about driving the car".

(4.55) Naoko$_i$ ga jibun$_i$ no kuruma o unten shite miru.
Naoko-NOM own-GEN car-ACC drive-TE see about
"Naoko$_i$ sees about driving her$_i$ own car".

This looks like possible support for a coindexing analysis, although the possibility of constructing a semantic argument to account for agreement exists, following Dowty & Jacobson (1989). On balance, however, the evidence from the *te morau* and the *te ageru* constructions also seems to refute neither of the two positions. Example (56), *te morau* patterns after a control verb of type *influence*, although the roles INFLUENCER and INFLUENCED are not particularly perspicuous for what is, here, a transfer relation with *Keiko* as the *source*, and the possible *agent* of the embedded verb; the *te ageru* construction, however, differs from all of the English control constructions discussed by Pollard & Sag (1994)[12] in that it is a ditransitive verb and yet it is subject-controlled, so in (57) *Oguchi* is the actor of *ageru* and the implied agent of *ie o tatete* (building the house), while the dative *Keiko* is a goal. Since the dative argument has subject like properties in the *te morau* case, a test for anaphoric reference might be expected to show that *jibun* (self) has two antecedents, while in the *te ageru* case only the controlling subject is expected as antecedent, so this case would not provide a good binding test. However the test in (58) shows that *jibun* in fact coindexes only with *Oguchi*, the controlling subject, and not with *Keiko*, the entailed subject of the embedded verb, so this test in fact offers no support for the structure-sharing account.

(4.56) Oguchi ga Keiko ni ie o tatete morau.
Oguchi-NOM Keiko-DAT house-ACC build-TE receive
"Keiko builds the house for Oguchi".

(4.57) Oguchi ga Keiko ni ie o tatete ageru.
Oguchi-NOM Keiko-DAT house-ACC build-TE give
"Oguchi builds the house for Keiko".

(4.58) Oguchi$_i$ ga Keiko ni ($_i$) jibun$_i$ no ie o tatete morau.
Oguchi-NOM Keiko-DAT own-GEN house-ACC build-TE receive
"Keiko builds his own house for Oguchi".

The foregoing discussion treats the conjunct forms of English, 'try', 'persuade', et al, and Japanese *miru*, and the outcome seems to be that the position which entails structure-sharing of the controlling argument and the controlled subject is not proven. When the hypothesis of polysemy is factored in, the overt linkage position in conjuncts requires a similar link between the nominal arguments of simplex *miru*, and such a link is unmotivated. The polysemy hypothesis, then, provides further impetus for the argument that there is no overt link in *control* constructions and the correlation between the subjecthood of the embedded verb

[12]Though there is an English dialect in which promise takes the same pattern.

and the nominative, or dative expressed argument, is simply an entailment of the interpretation of the second conjunct (head) verb.

4.4 Analyses of *morau* and *ageru*

The analyses of *morau* (receive) and *ageru* (give) pattern with those of simplex and conjunct *miru* in that the accusative object alternates with a *te* marked verbal conjunct. Both are ditransitive though, with *morau* denoting a receiving relation (59) and *ageru* denoting a giving relation, (60).

(4.59) Shunsuke ga Maiko ni hon o moratta.
Shunsuke-NOM Maiko-DAT book-ACC receive-PAST
"Shunsuke got a book from Maiko".

(4.60) Shunsuke ga Maiko ni hon o ageta.
Shunsuke-NOM Maiko-DAT book-ACC give-PAST
"Shunsuke gave Maiko a book".

Both verbs associate with *telic* adjuncts as with *1-ppun go ni* (after 1 minute), in examples (61) and (62).

(4.61) Shunsuke ga 1-ppun go ni Maiko ni hon o moratta.
Shunsuke-NOM after 1 minute Maiko-DAT book-ACC receive-PAST
"Shunsuke got a book from Maiko after a minute".

(4.62) Shunsuke ga 1-ppun go ni Maiko ni hon o ageta.
Shunsuke-NOM after 1 minute Maiko-DAT book-ACC give-PAST
"Shunsuke gave a book to Maiko after a minute".

These two verbs also associate with purposive adjuncts such as *shinchyou ni* (deliberately), indicating agentivity, in (63) and (64), so they are both Accomplishments.

(4.63) Shunsuke ga shinchyou ni Maiko ni hon o moratta.
Shunsuke-NOM deliberately Maiko-DAT book-ACC receive-PAST
"Shunsuke deliberately got a book from Maiko".

(4.64) Shunsuke ga shinchyou ni Maiko ni hon o ageta.
Shunsuke-NOM deliberately Maiko-DAT book-ACC give-PAST
"Shunsuke deliberately gave a book to Maiko".

The conjunctive forms of these verbs pattern a little differently, as *morau* combines with a *te* marked verb of any class, thus the Stative conjunction in (65), the Achievement in (66), the Activity in (67) and the Accomplishment in (68).

(4.65) John wa Mary ni Nihonjin de ite moratta
John-TOP Mary-DAT Japanese COP be-TE receive-PAST
"John received from Mary the favour of (her) remaining a Japanese."
((Ishikawa 1985))

(4.66) John ga Mary ni kokage ni kakurete moratta
I-TOP he-DAT shade-DAT hide-TE receive-PAST
"Mary was concealed in the shade for John".

(4.67) John ga Mary ni odotte moratta
John-TOP Mary-DAT dance-TE receive-PAST
"Mary danced for John".

(4.68) o-kaasan wa Miyuki-san ni shujin ni shashin o misete morau
Mother-TOP Miyuki-san-DAT husband-DAT photo-ACC show-TE receive
"Miyuki shows her husband the photograph, for her mother".

Examples (69) and (70) show that *ageru* combines with an Activity or an Accomplishment, but not with an Achievement, so (71) is unacceptable.

(4.69) John ga Mary ni odotte ageta
John-TOP Mary-DAT dance-TE give-PAST
"John danced for Mary".

(4.70) John ga Mary ni hon o yonde ageta
John-NOM Mary-DAT book-ACC read-TE give-PAST
"(lit) John gave Mary a reading of the book".

(4.71) *Saigo ga gunjin ni shinde ageta
Saigo-NOM soldier-DAT die-TE give-PAST

Another difference between *morau* and *ageru* is that they denote opposite transfer relations: in examples (65) to (69) *John* is the syntactic subject of *morau*, but *Mary* is the understood subject of the first conjunct verb; in (69) and (70) *John* is both the syntactic subject and the understood subject of *ageru*. Overall, both *te morau* and *te ageru* follow *te miru* in that the accusative nominal alternates with the *te* form in the direct object position, and the analysis also patterns after *miru*.

The lexical entry for ditransitive *morau* and its VP projection are given in Figures 4.26 and 4.27. The lexical entry for *morau* in 4.26 is ditransitive and has an extra, *locative*, role (6) linked to a dative marked NP (3). The ⟨ *nom,dat,acc* ⟩ ordering on the argument structure list follows the principles of linking proposed by Davis

(1995). He shows that in a transitive relation, the ACTOR links to the first position on ARG-S and the UNDERGOER links to the last among the subcategorized arguments[13].

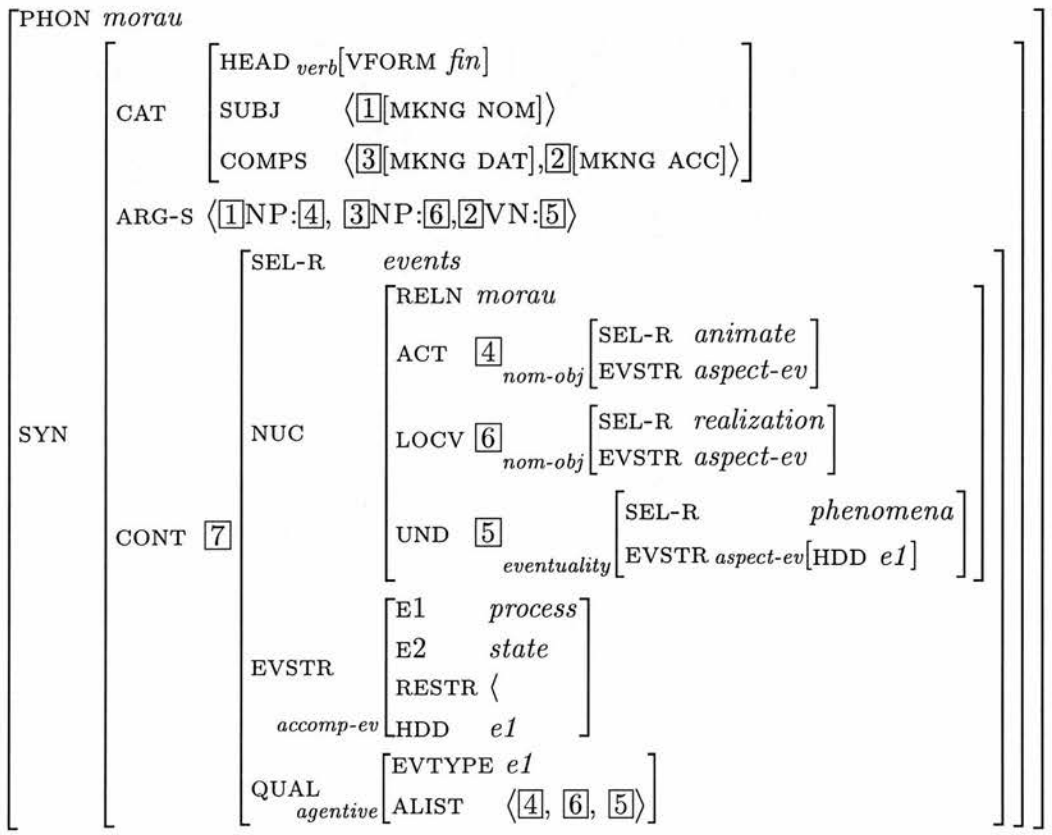


Figure 4.26: Lexical Entry for *morau*

Similar to the lexical entry for *miru* (see), the UNDERGOER role is underspecified to structure share with either a nominal object (*nom-obj*) or a *psoa*, each of which is a subtype of *eventuality* in the *content* type hierarchy. Both the simplex in (72) and the conjunct form in example (73) are admissible to this formulation of *morau*. The nominal *keeki o* (a cake) is admitted because it has a SEL-R value of *physobj*, subtype of *phenomena*, and an event structure of *aspect-rel*, maximally underspecified. The VP *keeki o tukutte* is admitted because it has a SEL-R value of *phenomena* and

[13]Davis represents the position of UNDERGOER on ARG-S as ARG-S ⟨ ..., *NP(und)*, ... ⟩ with the presumption that other, unspecified, arguments may follow. So if the UNDERGOER is the only role, as in an unaccusative relation, it takes first position and is linked to subject; if there are two roles, it takes second position, and if there are three roles it takes third position. The LOCATIVE role in a ditransitive relation takes second position in ARG-S, then. Davis does not show what happens in a Stative predicate with a LOCATIVE relation, but the consequence of his system is that LOCATIVES take first position in ARG-S and are therefore lexicalized as subject. I discuss evidence from Japanese in Chapter 2 and Chapter 4 of this thesis and show that LOCATIVES are not always subjects in Stative predicates. Clearly, more work is needed on the argument structure of States.

an event structure of *accomp-rel*, and is underspecified for headedness. Because the selectional restriction on event structure of the UNDERGOER of *morau* is more permissive than that of *miru*, the restriction *EVSTR [HDD e1]*$_{aspect-rel}$ admits a verb of any aspectual class, provided it has an E1 headed event structure.

(4.72) (Maiko ga) Naoko ni keeki o morau.
(Maiko-NOM) Naoko-DAT cake-ACC receive
"Maiko receives a cake from Naoko".

(4.73) (Maiko ga) Naoko ni keeki o tukutte morau.
(Maiko-NOM) Naoko-DAT cake-ACC bake-TE receive
"Naoko bakes a cake for Maiko".

The phrasal structure for *Naoko ni keeki o tukutte morau* given in Figure 4.27 is complex, so I have included a key to assist in reading it (Table 4.2).

Abbrev.	**Field**	**Explanation.**
HD1	Head Daughter	The head of the Comp-Head phrase dominating *te*, the lexical entry for the verb *morau.*
HD2	Head Daughter	The head of the Comp-Head phrase *keeki o tukutte.*
NHD1	Non-Head Daughter	The dative marked noun phrase *Naoko ni.*
NHD2	Non-Head Daughter	The *te* marked verb phrase *keeki o tukutte.*
NHD3	Non-Head Daughter	The accusative marked complement daughter of *tukutte*, *keeki o.*
SYN1	synsem	The *synsem* value of the phrasal mother of the Comp-Head phrase headed by *morau.*
NUC	nucleus	The *nucleus* in the *content* of *tukuru.*
EVSTR	Event Structure	The Event Structure in the *content* of *tukuru.*
QUAL	qualia	The Qualia Structure in the *content* of *tukuru.*
chp1	Complement-Head Phrase	The phrasal sign for *Naoko ni keeki o tukutte morau.*
chp2	Complement-Head Phrase	The phrasal sign for *keeki o tukutte.*
hmp1	Head-Marker Phrase	The phrasal sign for *Naoko ni.*
hmp2	Head-Marker Phrase	The phrasal sign for *keeki o.*

Table 4.2: Key to Figure 4.27

The phrasal structure which is a Comp-Head phrase (*chp1*) has two non-head daughters, a head daughter and a SYNSEM value for the phrasal mother. The head daughter (HD1) is the saturated form of the verb *morau* given in Figure 4.26. The non-head daughters are the complements NHD1, the dative nominal *naoko ni* and NHD2, the *te* marked VP with accusative marking value, *keeki o tukutte.* The phrasal mother has saturated complements (COMPS ⟨⟩), unsaturated subject (COMPS ⟨$\boxed{1}$⟩)

and no argument structure. The contents of the accusative marked non-head daughter (NHD2) are identical with the same VP argument of *miru*, given in Figure 4.20. However, the entailed subject of the embedded VP in the case of *te morau* is the dative marked argument NHD1. There is no overt reference to this in the ARG-S of *tukutte*, following the arguments of Chierchia (1984).

The lexical entry for *ageru* (give) is given in Figure 4.28, and this differs from *morau* only in the selectional restrictions of the undergoer. Where *morau* selects for an Event Structure (EVSTR) of type *aspect*, allowing any aspectual class, *ageru* selects only for an agentive verb. Of course the real semantic distinction between them is that *morau* represents a receiving relation in which the LOCATIVE role is the source, while *ageru* represents a giving relation in which the ACTOR is the source. This may well be captured by articulating the relation (RELN) field to specify a logical relationship between the roles.

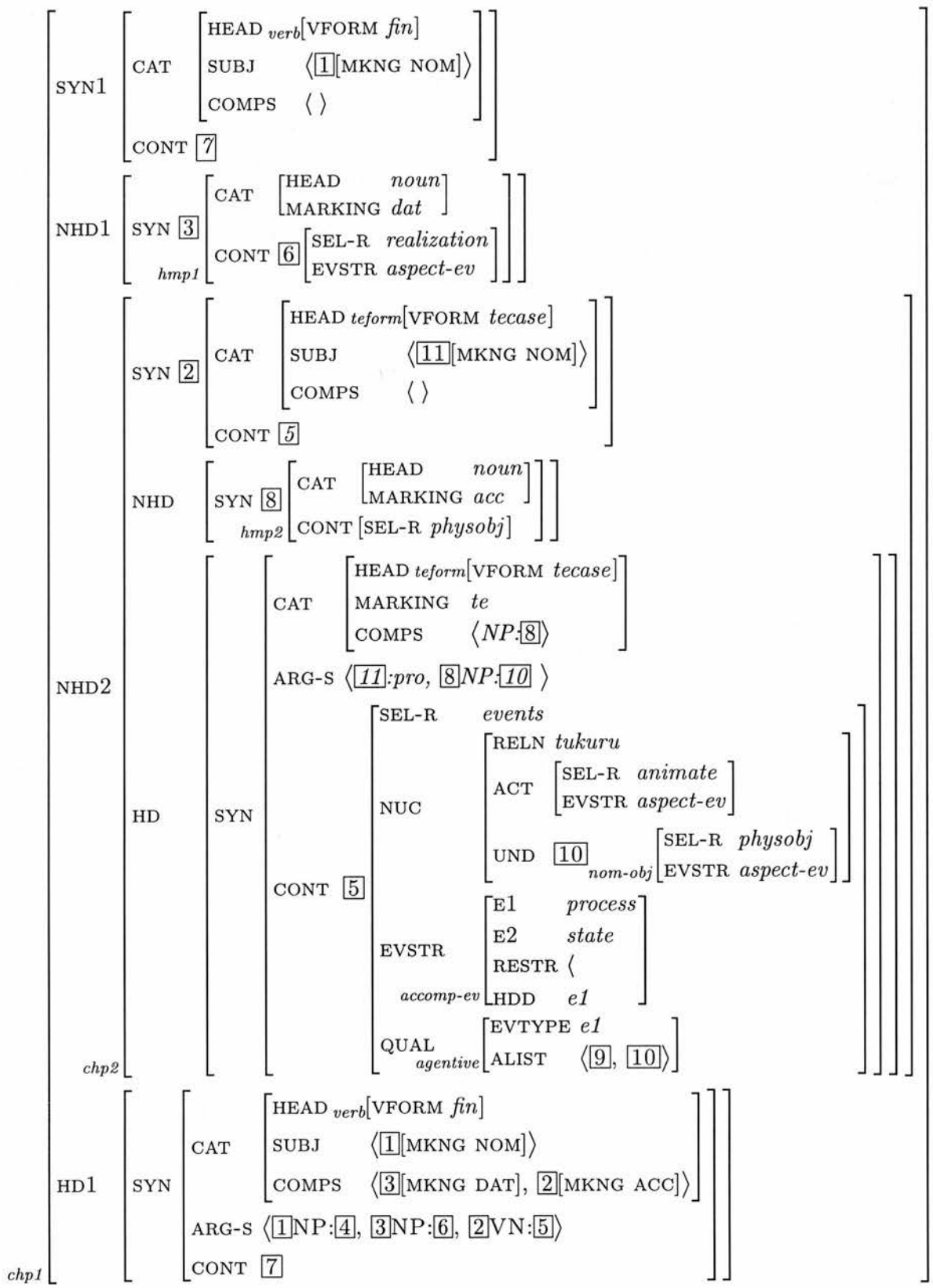


Figure 4.27: keeki o tukutte morau

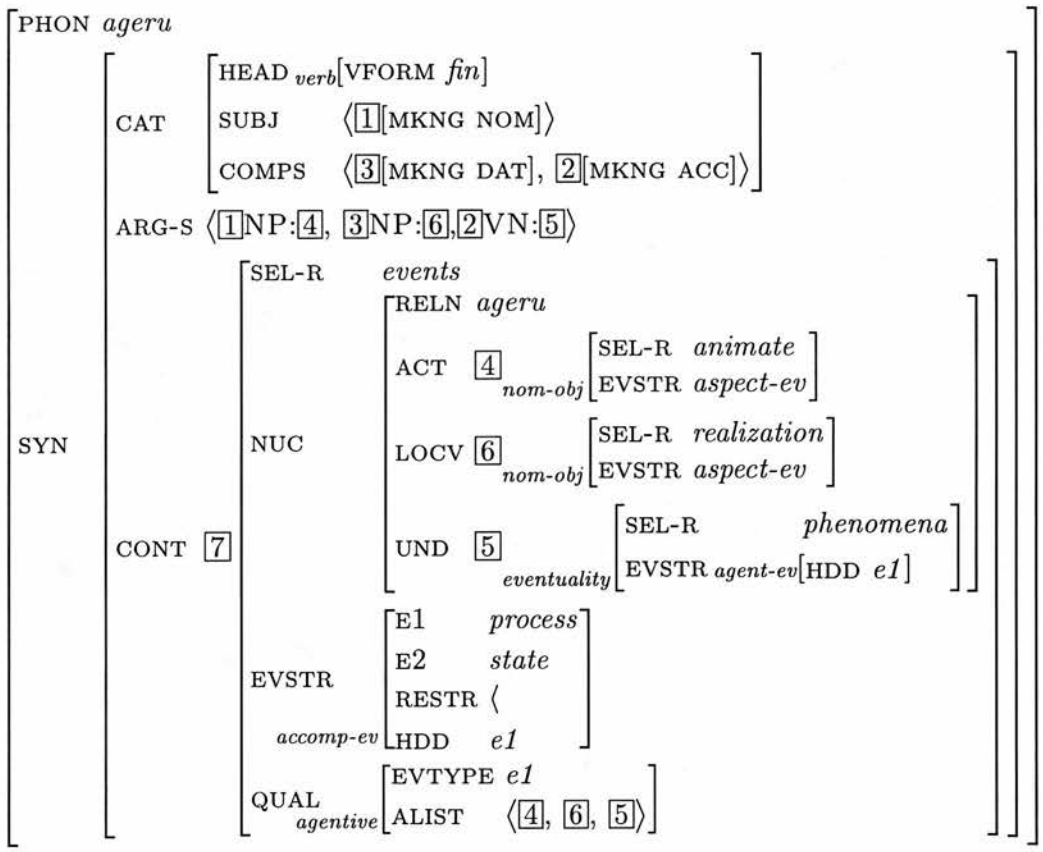


Figure 4.28: Lexical Entry for *ageru*

4.5 Summary

The simplex and conjunctive phenomena of the verbs *miru*, *morau* and *ageru* show a regular alternation between the accusative complement and the *te* marked V1 conjunct. On the basis of such verb frame alternations, these verbs are polysemous (according to Pustejovsky (1995), following Levin (1993)). In this chapter I have shown that this polysemy can be captured by underspecifying the complements field, and its linked undergoer role. The selectional restrictions on verbal complements are sensitive to aspect, but HPSG does not currently differentiate aspectual phenomena. If, however, the CONTENT field is extended to incorporate Pustejovsky's (amended) Event Structure and Qualia Structure, then aspectual features can be captured by articulating *content* into a type hierarchy. The *content* types developed by Davis (1995) are put back into the NUCLEUS, and subordinated to aspect. In order to maintain consistency between the selectional restrictions on verbal and nominal arguments respectively, nominals must also account for aspect. If nominals are maximally underspecified for aspect, then a verb which selects for any particular aspectual class (Activity, say, or Achievement) or feature (agentivity or telicity) will automatically select any nominal complement, provided that other selectional restrictions are compatible.

In the next chapter, a different set of Nuclear Conjunctive phenomena is investigated, and the analyses are not so clearly polysemous. A question which then arises is whether any generalization can be made to distinguish those verbs which are polysemous, from those verbs which are not.

Chapter 5

Sense Enumerativity and Homonymy

In the last chapter, the analysis of *miru*, *morau* and *ageru* exploits the fact that the machinery of HPSG can be used to underspecify syntactic arguments and semantic roles, and thereby provide a polysemous account of the valence alternations observed in the uses of those verbs.

Of the data presented in Chapter 2, none of the remaining verbs exhibit such clearly polysemous behaviour. The Stative verbs *iru* (animate be) and *aru* (inanimate be) show multiple simplex projections; *iru* has extra complexity in the semantics of its conjunct forms; and *aru* shows two distinct conjunct projections. The complexity of these phenomena stretch the capabilities of the HPSG framework for developing unitary lexical entries. Pustejovsky (1995) claims that traditional frameworks for describing lexical items are deficient in the machinery which will allow the unit representation of polysemous items. The consequence is that the separate senses of words which are, in fact, polysemous, must be separately listed in the lexicon. Pustejovsky calls this the *Sense Enumerative Lexicon*. In his development of lexical semantics, he propounds a mechanism which allows more of the polysemous behaviour to be described, resulting in more complex lexical entries, in a more compact lexicon. I have incorporated part of this mechanism within the framework of HPSG, to capture the expression of verbal aspect. Together with the type hierarchy of HPSG, this machinery explains the polysemy of *miru*, *morau* and *ageru*, and the selectional restrictions on the aspectual class of the embedded conjuncts.

In this chapter I reiterate the data of *iru* and *aru*, in their simplex and conjunct forms, and explore analyses which show the limitations of the HPSG framework. The simplex lexical forms of *iru* and *aru* are developed in Subsections 5.1 and 5.2, respectively.

The conjunct form of *aru* is the more complex among the conjunctive forms of these verbs, and has been analysed previously, particularly by Hasegawa (1995). Her analysis is reviewed in Section 5.3. The following section looks at the lexical construction of the verbs which can appear as first conjunct, and I develop an account which modifies Pustejovsky's (1995) Event Structure framework to describe the various aspectual forms. A question which has arisen often in the study of Japanese is "what is the *te* morpheme'?" (Hasegawa gives a comprehensive bibliography). Following my review of Hasegawa, and analysis of the complex lexical form of Accomplishments, I take up this question and propose a solution consistent with the phenomena of verbal conjunctions. In the next two sections, I return to the lexical analysis of the verbs *iru* and *aru*, after having considered their various conjunct forms, and develop lexical entries in Sections 5.5. and 5.6.

The remaining verbs which I presented in Chapter 2 show further lexical diversity, and I report on their analyses in Section 5.7. The chapter is concluded in 5.8.

5.1 Simplex *iru*

Simplex *iru* (animate be) is a Stative predicate with two arguments as shown in example (1): the *ga* marked argument is invariably animate, and the *ni* marked argument is a location. The entire predicate denotes a locative relation: the fact that the *ni* marked argument is a location suggests that its semantic role is that of a locative. The *ga* marked argument, although animate, is not agentive, as demonstrated by the fact that it does not happily combine with the agentive adjunct *otonashiku* (quietly).

(5.1) Taroo ga ie ni shibaraku mae kara iru.
Taroo-NOM house-DAT since a while ago be
"Taroo has been at home since a while ago".

(5.2) *Taroo wa ie ni otonashiku iru.
*Taroo-TOP house-DAT quietly be
"?Taroo is quietly at home".

Apart from that of ACTOR, the only available role in Davis's (1995) linking system is the UNDERGOER, so the non-locative argument of *iru* in particular (and Stative verbs in general) is linked to this role.

One particular problem with Stative verbs in Japanese is how to determine which of the two arguments, LOCATIVE or UNDERGOER, is projected to subject and it is

important to resolve this problem because in the HPSG framework, the subject is projected as a separate valence. On the one hand Tateishi (1994) suggests that the alternation of *ni/ga* marking is indicative of subjecthood, and on the other hand, coindexing with *jibun* (self) denotes antecedence, and in a two-place relation this is indicative of subjecthood[1]. In example (3) *Taroo* is coindexed with *jibun* and is therefore subject, and moreover the LOCATIVE *ie* does not undergo *ni/ga* alternation, so (4) is bad.

(5.3) Taroo ga jibun no ie ni iru.
Taroo-NOM self-GEN house-DAT be
"Taroo is in his own house".

(5.4) *ie ga Taroo ga iru.
ie-NOM Taroo-NOM be

There seems to be a distinction in forms, or levels, of animacy, giving rise to differences in which kinds of arguments can undergo *ni/ga* alternation: if the general animate *hito* (people) replaces the named individual *Taroo*, then the locative *ie* (house) does undergo *ni/ga* alternation, in examples (5) and (6).

(5.5) hito ga ie ni iru.
people-NOM house-DAT be
"People are in the house".

(5.6) hito ga ie ga iru.
people-NOM house-NOM be
"It is the house where people are'.

In testing for *jibun* (self) binding, while the pure *ga* marked *hito* (people) coindexes with *jibun* in (7), indicating that *hito ga* is subject, *it ie* cannot bind *jibun* in example (8) and cannot therefore be subject in this sentence.

(5.7) hito ga jibun no ie ni takusan iru.
people-NOM self-GEN house-DAT many be
"There are many people in their own houses".

(5.8) *ie ga jibun no hito ga iru.
house-NOM self-GEN people-NOM be

[1]Iida (1995) claims that there is no syntactic requirement that the antecedent of *jibun* be a subject, but that there is a discourse rule which blocks the oblique argument in a 3-place relation from coindexation. However since HPSG binding is consequent on obliqueness ordering in argument structure, the less oblique argument in a 2-place relation which binds the *jibun* adjunct of the more oblique argument is necessarily the subject.

The case where the locative argument is animate is indicative of a possessive relation, and in this case arguments can antecede *jibun*, as shown in (9); in its alternate *ga* marked form, the same argument can also antecede *jibun*, in (10).

(5.9) haha ni wa jibun no kodomo ga takusan iru.
mother-NOM-TOP self-GEN children-DAT many be
"As for the mothers, their own children are many".

(5.10) haha ga jibun no kodomo ga takusan iru.
mother-NOM self-GEN children-DAT many be
"There are mothers with many of their own children".

In summary, the behaviour of simplex *iru* is that the argument which is projected to subject depends on animacy, and there are cases where either the *ni/ga* marked argument, or the pure *ga* marked argument is the subject. Homonymous forms of *iru* can be straightforwardly derived with the two roles of UNDERGOER and LOCATIVE alternately projected to subject and complement. Figure 5.1(a) shows one alternative with the undergoer projected to subject with *nom* marking, and the locative projected to complement, underspecified for *dat* and *nom* marking (the supertype is *subjform* in the marking type hierarchy); 5.1(b) shows the alternate projection with the LOCATIVE projected to subject, with *subjform* marking.

(a):

$$\begin{bmatrix}\text{PHON } \langle iru \rangle \\ \text{SYN} \begin{bmatrix}\text{CAT} & \boxed{0}\begin{bmatrix}\text{HEAD} & _{verb}[\text{VFORM } fin] \\ \text{SUBJ} & \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG DAT}] \rangle \\ \text{MARKING} & unmarked\end{bmatrix} \\ \text{ARG-S} & \langle \boxed{1}\text{NP:}\boxed{3},\ \boxed{2}\text{NP:}\boxed{4} \rangle \\ \text{CONT} & \begin{bmatrix}\text{SEL-R} & events \\ \text{NUC} & \begin{bmatrix}\text{RELN } iru \\ \text{UND} & \boxed{3}\begin{bmatrix}\text{SEL-R} & animate \\ \text{EVSTR} & aspect\text{-}rel\end{bmatrix} \\ \text{LOC} & \boxed{4}\begin{bmatrix}\text{SEL-R} & realization \\ \text{EVSTR} & aspect\text{-}rel\end{bmatrix}\end{bmatrix} \\ \text{EVSTR} & _{state\text{-}ev}[\text{E1 } state] \\ \text{QUAL} & _{formal}\begin{bmatrix}\text{EVTYPE} & e1 \\ \text{ALIST} & \langle \boxed{3}, \boxed{4} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

(b):

$$\begin{bmatrix}\text{PHON } \langle iru \rangle \\ \text{SYN} \begin{bmatrix}\text{CAT} & \boxed{0'}\begin{bmatrix}\text{HEAD} & _{verb}[\text{VFORM } fin] \\ \text{SUBJ} & \langle \boxed{2}[\text{MKNG SUBJFORM}] \rangle \\ \text{COMPS} & \langle \boxed{1}[\text{MKNG DAT}] \rangle \\ \text{MARKING} & unmarked\end{bmatrix} \\ \text{ARG-S} & \langle \boxed{2}\text{NP:}\boxed{3},\ \boxed{1}\text{NP:}\boxed{4} \rangle \\ \text{CONT} & \begin{bmatrix}\text{SEL-R} & events \\ \text{NUC} & \begin{bmatrix}\text{RELN } iru \\ \text{UND} & \boxed{3}\begin{bmatrix}\text{SEL-R} & animate \\ \text{EVSTR} & aspect\text{-}rel\end{bmatrix} \\ \text{LOC} & \boxed{4}\begin{bmatrix}\text{SEL-R} & realization \\ \text{EVSTR} & aspect\text{-}rel\end{bmatrix}\end{bmatrix} \\ \text{EVSTR} & _{state\text{-}ev}[\text{E1 } state] \\ \text{QUAL} & _{formal}\begin{bmatrix}\text{EVTYPE} & e1 \\ \text{ALIST} & \langle \boxed{3}, \boxed{4} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 5.1: *iru* with Alternate Valence Projections

The fact that both alternative projections occur in linking semantic and syntactic arguments suggests that there may be a way to reconcile them in a single lexical entry. In Chapter 3 I discussed Manning's (1994) distinction between the mappings from ARG-S to valence in *accusative* and *ergative* languages respectively. This was illustrated in Figure 3.18, repeated here as Figure 5.2.

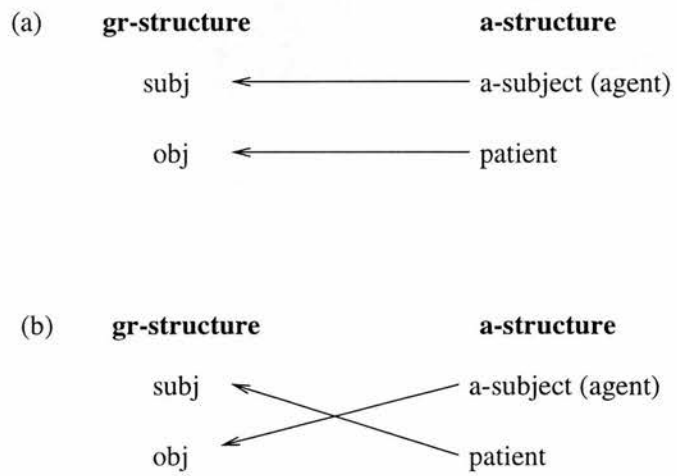


Figure 5.2: Accusative (a) and Ergative (b) Mappings

The alternative projections to subject can be represented by underspecifying the complex CAT feature, which requires that the indexed values [0] and [0'] in Figure 5.1 (a) and (b) respectively, inherit from a common supertype. There is a precedent for this in Davis's (1995) articulation of the CONTENT field into subtypes depending on the number and type of roles appearing in a verbal relation. For example, an unergative verb has a single actor (ACT) role, encoded in an unergative relation (*unerg-rel*), and a transitive verb has actor (ACT) and undergoer (UND) roles, encoded in an *act-und-rel*; these relations are all subtypes of type *relation* which defines all verbal content types. The category (*cat*) type hierarchy is undifferentiated in current HPSG and in particular the subject (SUBJ), complements (COMPS) and specifier (SPR) features are list types, unconstrained by their supertype (*cat*) as to the types and numbers of arguments appearing in a list. Manning's (1994) analysis of the complementary distribution of ARG-S to valence mappings between ergative and accusative languages suggests that *cat* might be differentiated into subtypes, with a major distinction being that between *ergative* and *accusative* mappings, with the valences complementarily articulated. I am not suggesting that *iru* is ergative, indeed, it is underspecified for whatever type controls the mapping to subject, and the binding evidence shows that the ordering of roles is also reversible. It is, perhaps, possible though that *iru* in particular, and perhaps other Stative verbs, is

representative of a *cat* type hierarchy with subtypes differentiated by their valence specifications.

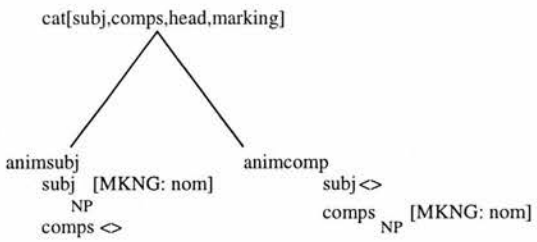


Figure 5.3: Articulation of the *cat* Type Hierarchy

The top of a possible type hierarchy for *cat* is sketched in Figure 5.3, with subtypes of *animsubj* and *animcomp*, and the animate *ga* argument is projected to subject in the *animsubj* subtype, and to complement in the *animcomp* subtype. The effect of this sort of scheme is to constrain projection from ARG-S to valence in the type hierarchy, and typing restrictions are placed on the feature values under *cat*, so in this case *animsubj* is differentiated from *cat* by detailing the value of SUBJ and constraining the development of lexical entries by restricting the types of some or all of the features under any supertype. For example, the supertype *cat* has features HEAD, SUBJ, COMPS and MARKING (SPR is unused in this formulation of Japanese). The features HEAD and MARKING are 'scalar' features, with values articulated in a type hierarchy, thus *head* has subtypes *subst* and *func*, while *marking* has subtypes *marked* and *unmarked*, and so on. In contrast, SUBJ and COMPS are list valued features, which may contain elements of any type, and this correspondingly lifts the constraints on a lexicon developer for populating lists, so, for example, the value of HEAD is constrained by the *head* type hierarchy to contain *noun*, *verb*, *mark*, *adjective* or *adverb*, but the value of SUBJ is a *list*, which may be empty or non-empty, and can contain any arrangement of the items under *bot*[2], not limited to lexical signs, but including also phrasal signs, content fields, boolean marking values, nuclei, or even other lists, and moreover one list may contain any combinations of these types of sign, as it is not context sensitive[3]. There is, however, a way to constrain the articulation of lists, because the "bracketed list" notation shown in feature structure diagrams is simply an informal notation for the convenience of the writer (and reader). In the ALE implementation (Carpenter & Penn (1998)), list

[2]In inheritance based polymorphic systems, there is a unique most general type. This is *bot*, and all other types in the system are subtypes of *bot*, or subtypes of subtypes.

[3]In the HPSG theory the is, however, a further constraint on the population of lists in that they are defined as, e.g. *list-of-synsem* or *list-of-sign*. In the articulation of HD and TL features of a valence (SUBJ or COMPS), the value of HD is, then, constrained to be of type *synsem*. This is illustrated in Figure 5.4.

valued items are formally articulated with empty-list and non-empty-list subtypes, and with features HD and TL, where TL is "recursively" defined to be of type *list*. In order to articulate the value of feature SUBJ under *animsubj*, then, it is necessary to specify that the list is of subtype *ne-list* with feature HD valued *synsem*, and with a head value of *noun* and a MARKING value *nom*. The list type hierarchy is shown in Figure 5.4.

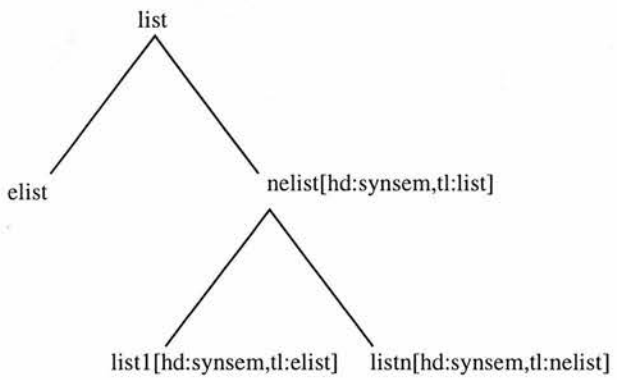
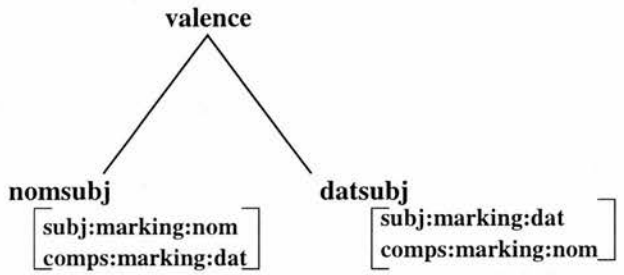


Figure 5.4: Articulation of the *list* Type Hierarchy

The problem with underspecifying list items in the type hierarchy is that it quickly leads to a combinatorial explosion: if it is necessary to underspecify lexical items by alternating their marking values in the subject (SUBJ) and complements (COMPS) lists, as I am proposing here, then the type hierarchy for *cat* must be articulated to specify every nested level of (simple or) complex type, from CAT[] to MARKING: CASE. For example if the required alternations are nominative (*nom*) and dative (*dat*) marking values, as in (11)(a) and (b), it is probably unrealistic to develop type hierarchies with this level of specification. Is there, then, a way to 'shorten' the specification?

(5.11) (a) **cat:subj:ne-list:hd**:[*sign*, **cat:marking**:*nom*],
(b) **cat:subj:ne-list:hd**:[*sign*, **cat:marking**:*dat*].

One way in which the HPSG formalism does not help is in requiring that valences be specified in a list, because as above, the list notation, (a) overcomplicates the specification of feature paths, but also (b) provides a discontinuity in the constraining mechanisms, by allowing the value of the head of the list (HD) to be any direct subtype of *bot*. Since a subject invariably comprises a list with a single *synsem* object, and complements comprise zero, one or two (in Japanese, as in English), then it may be possible to dispense with the list notation for valences.



b) $\left[\text{CAT}\left[\begin{array}{ll}\text{HEAD} & \textit{verb}[\text{VFORM } \textit{fin}] \\ \text{MKNG} & \textit{unmarked} \\ \text{VLNC} & \left[\begin{array}{ll}\text{SUBJ} & \textit{NP}[\text{MKNG } \textit{nom}] \\ \text{COMPS} & \textit{NP}[\text{MKNG } \textit{dat}]\end{array}\right]\end{array}\right]\right]$ c) $\left[\text{CAT}\left[\begin{array}{ll}\text{HEAD} & \textit{verb}[\text{VFORM } \textit{fin}] \\ \text{MARKING} & \textit{unmarked} \\ \text{VLNC} & \left[\begin{array}{ll}\text{SUBJ} & \textit{NP}[\text{MKNG } \textit{dat}] \\ \text{COMPS} & \textit{NP}[\text{MKNG } \textit{nom}]\end{array}\right]\end{array}\right]\right]$

d) $\left[\text{CAT}\left[\begin{array}{ll}\text{HEAD} & \textit{verb} \\ \text{MKNG} & \textit{unmarked} \\ \text{VLNC} & \textit{valence}\end{array}\right]\right]$

Figure 5.5: A Type Hierarchy for Valence Alternations.

Figure 5.5(a) sketches a type hierarchy for underspecifying valence values. The valence hierarchy underspecifies the marking value so that the subject takes either nominative (*nom*) marking in subtype *nomsubj*, or dative (*dat*) marking in *datsubj*[4]. The lexical realizations of these alternatives are given in 5.5(b) and (c) respectively. A lexical entry which is underspecified for valence projection, then, is given in Figure 5.5(d).

If this approach to lexical specification can be sustained and fully articulated, then another area of sense enumerativity can be alleviated, and verbs such as *iru* with subject marking alternations can be properly rendered as polysemous. There are still several problems to be solved in the feature geometry, though, before fully flexible lexical specifications can be realized. One such problem is in how to specify the link between valences and argument structure, and this includes developing a mechanism for coindexing SUBJ and COMPS, through their supertype, *valence* with positions in ARG-S.

[4]In the COMPS of nomsubj and the SUBJ of datsubj I have simply given *ni* or *ga* as the marking value. The actual possibilities are either *ni* or *ga* in both instances. This alternation is encoded in the marking type hierarchy as *subjform*, with subtypes *nom* and *dat*. I have not used the underspecified *allsubj* type in this discussion, for the sake of simplicity.

In summary[5], the issue is that the realization of the lexicon should be structured so as to maximize generalization and minimize the number of discrete lexical entries. While the formal structure of HPSG goes some way towards achieving this goal, the current formalism prevents specifying subject and complement alternations in the type hierarchy and they must therefore be listed separately in the lexicon. Although some linguistic phenomena (such as the alternations of the verb *shimau* (put, complete, regret) are homonymous in fact, and only phonologically related, the alternations of *iru* are semantically related and, properly, polysemous. The list notation on valences, however, prevents this fact from being encoded in the formal lexicon.

In developing the analyses of conjunct *iru* and simplex and conjunct *aru*, I will adhere to the current list-oriented formulation, rather than attempt to invent new mechanisms. Figures 5.1(a) and (b), then, enumerate the lexical forms of simplex *iru*, with the alternate projections to subject.

5.2 Simplex *aru*

The simplex form of *aru* resembles *iru* in that it denotes a locative relation in which the locative argument undergoes *ni/ga* alternation, as in examples (12) and (13).

(5.12) kuruma ni handoru ga aru.
car-DAT steering-wheel-NOM be
"In the car there is a steering wheel".

(5.13) kuruma ga handoru ga aru.
car-NOM steering-wheel-NOM be
"Only in the car is there a steering wheel".

However *jibun* binding does not apply to inanimate arguments, so there is no independent test for subjecthood in these cases. Martin (1975) does however point out some cases where animate forms can be used as arguments of *aru*, such as the interrogative possessive in (14), which can be antecedent of *jibun* in (15).

(5.14) dare ni kodomo ga arimasu ka.
who-DAT kids-NOM Be-POL QN
"Who has kids?".

[5]It has been pointed out to me that this analysis lacks a theory of how selectional restrictions work, particularly the *animacy* restriction on the nominative argument of *iru*. Following from the discussion of Pustejovsky (1995) in the previous chapter, the geometric placement of the *animate* (and *physobj*) restrictions is most conveniently located on the BACKGROUND feature.

(5.15) dare ni jibun no kodomo ga arimasu ka.
who-DAT self-GEN kids-NOM Be-POL QN
"Who has their own kids?".

This is indicative of subjecthood for the possessive case, but given that animacy[6] outranks inanimacy as a criterion for subjecthood, this still tells nothing about the inanimate cases. We will see in the conjunct cases in the next section that the pure *ga* marked UNDERGOER argument is projected to subject, as is the case with *iru*. It may be that *aru* patterns with *iru* and that in the pure inanimate cases (or the pure locative relation) the *ga* marked UNDERGOER is projected to subject, and so there are two lexical entries for simplex *aru* along the lines of 5.1 (a) and (b) for *iru*, encoded here as 5.6. While the evidence for whether there is, or is not, a projection of the UNDERGOER to *ga* marked subject is thin, in simplex *aru*, 5.6(a), this does allow a possible correspondence between the simplex and conjunct lexical forms, discussed below.

[6]The distinction between the undergoer arguments of *iru* and *aru* is actually more subtle and complex than a straight animacy/inanimacy distinction, but this is a topic which must to be reserved for future work.

(a):

$$\begin{bmatrix}\text{PHON } \langle aru \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } _{verb}[\text{VFORM } \textit{fin}] \\ \text{SUBJ } \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{COMPS } \langle \boxed{2}[\text{MKNG SUBJFORM}] \rangle \\ \text{MARKING } \textit{unmarked} \end{bmatrix} \\ \text{ARG-S } \langle \boxed{1}\text{NP:}\boxed{3}, \boxed{2}\text{NP:}\boxed{4} \rangle \\ \text{CONT} \begin{bmatrix} \text{NUC} \begin{bmatrix} \text{RELN } \textit{aru} \\ \text{UND } \boxed{3}\begin{bmatrix}\text{SEL-R } \textit{physobj} \\ \text{EVSTR } \textit{aspect-rel}\end{bmatrix} \\ \text{LOC } \boxed{4}\begin{bmatrix}\text{SEL-R } \textit{realization} \\ \text{EVSTR } \textit{aspect-rel}\end{bmatrix}\end{bmatrix} \\ \text{EVSTR } _{\textit{state-ev}}[\text{E1 } \textit{state}] \\ \text{QUAL } _{\textit{formal}}\begin{bmatrix}\text{EVTYPE } \textit{e1} \\ \text{ALIST } \langle \boxed{3}, \boxed{4} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

(b):

$$\begin{bmatrix}\text{PHON } \langle aru \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } _{verb}[\text{VFORM } \textit{fin}] \\ \text{SUBJ } \langle \boxed{2}[\text{MKNG SUBJFORM}] \rangle \\ \text{COMPS } \langle \boxed{1}[\text{MKNG NOM}] \rangle \\ \text{MARKING } \textit{unmarked} \end{bmatrix} \\ \text{ARG-S } \langle \boxed{2}\text{NP:}\boxed{4}, \boxed{1}\text{NP:}\boxed{3} \rangle \\ \text{CONT} \begin{bmatrix} \text{SEL-R } \textit{events} \\ \text{NUC} \begin{bmatrix} \text{RELN } \textit{aru} \\ \text{UND } \boxed{3}\begin{bmatrix}\text{SEL-R } \textit{physobj} \\ \text{EVSTR } \textit{aspect-rel}\end{bmatrix} \\ \text{LOC } \boxed{4}\begin{bmatrix}\text{SEL-R } \textit{realization} \\ \text{EVSTR } \textit{aspect-rel}\end{bmatrix}\end{bmatrix} \\ \text{EVSTR } _{\textit{state-ev}}[\text{E1 } \textit{state}] \\ \text{QUAL } _{\textit{formal}}\begin{bmatrix}\text{EVTYPE } \textit{e1} \\ \text{ALIST } \langle \boxed{3}, \boxed{4} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 5.6: *aru* with Alternate Valence Projections

5.3 Hasegawa's (1995) Analysis of *te aru*

Hasegawa (1995) develops a model-theoretic analysis of the *te aru* construction using Dowty's (1979) aspectual calculus, which I reviewed in Chapter 1, and Foley & Van Valin's (1984) Actor and Undergoer Hierarchy, utilizing the *macroroles* of Actor and Undergoer. A brief review of these concepts is appropriate, prior to reviewing Hasegawa.

In their development of Role and Reference Grammar, Foley & Van Valin (1984) define a thematic hierarchy with six roles, those of Agent, Effector, Experiencer, Locative, Theme and Patient, and the Locative includes one from Source, Path, Goal and Recipient. Every verb is semantically defined by a formula in the aspectual calculus, and the arguments in each formula are linked to thematic roles. The locative state *aru*, or 'be' is defined by the predicate *be-at(x,y)*, where x is mapped onto the locative and y is mapped onto the theme role. For an Accomplishment such as *kiru*, or 'chop', Dowty's formula is *do(x,ϕ) cause become pred'(y)*, and Foley and Van Valin identify x as the *agent* and y as the *patient*; more generally, they realized that once the roles are ordered in a hierarchy, in a 2-place relation the role higher in the hierarchy is always lexicalized as the subject and the lower role is lexicalized as object, independently of which particular pairs of roles are present in a predicate. The *macroroles* are given the general labels of *actor* and *undergoer*, and indeed these are the labels which Davis (1995) uses in defining the CONTENT relations in HPSG (See Chapter 3, Section 3.4).

In Hasegawa's analysis of the *te aru* sentence shown in example (16), she claims that the intransitivizing process is the result of an operation on the Logical Structures of the conjunct verbs, *kiru* (chop) and *aru* (be), an Accomplishment and a State respectively.

(5.16) yasai ga kitte aru
vegetables-NOM chop-TE be
"The vegetables are being chopped". (Hasegawa (1995))

These (Dowtian) logical structures are given, for *kiru* (chop) in (17)(a) and for *aru* (inanimate be) in (17)(b)[7].

[7]Hasegawa's (1995) thematic roles, *agent*, *patient*, *locative* and *theme* are based on Foley and Van Valin's (1984) Thematic Hierarchy. In the subsequent HPSG lexicalization, I use the Dowty/Davis model, including only *actor*, *locative* and *undergoer* roles. THus, Hasegawa's *agent* is mapped into an HPSG *actor*, her *patient* and *theme* are mapped into *undergoer*.

(5.17) (a) *kiru* (chop)
[**do'**(x, ϕ)]*cause*[*become***chopped'**(y)] x=agent, y=patient.
(b) *aru* (be)
be-at'(x,y) x=locative, y=theme.

According to Hasegawa, the combined conjunction has only one macrorole which is derived from the theme of the first conjunct *kiru* by an intransitivizing process she represents as (18).

(5.18) [**do'**(x,ϕ)] CAUSE [BECOME **chopped'**(y)] $\longrightarrow$ **chopped'**(y).

This is interesting, because it mirrors the distinction between Agentive and Formal Qualia in Pustejovsky (1995) as follows:

The Event Structure and Qualia Structure of *kiru* are given in Figure 5.7 for comparison. This is a bieventual structure with a *process* followed by a *state* in strict partial order. In the Agentive Quale, the ACTOR role is selected for animacy and the UNDERGOER role is selected as a *physobj*. In the Formal Quale, the UNDERGOER selected for *physobj* is the only argument projected. In the *kitte aru* conjunction, (19), this UNDERGOER argument is what Hasegawa claims as the single macrorole. The *do/cause* structure above is replaced in Pustejovsky by the E1 headed Agentive Quale, and the resultant predicate is replaced in the figure by the E2 headed Formal Quale. The intransitivizing process depicted in (18) is captured in Event Structure by underspecifying for headedness, therefore allowing either E1 or E2 to be projected.

$$\left[\begin{array}{l}\text{CONT}\left[\begin{array}{ll}\text{EVENTSTR} & \left[\begin{array}{ll}\text{E1} & \textit{process}\\ \text{E2} & \textit{state}\\ \text{RESTR} & \langle\\ \text{HDD} & \textit{e2}\end{array}\right]\\ \text{QUALIA} & \left[\begin{array}{ll}\text{EVTYPE} & \textit{e2}\\ \text{ALIST} & \langle \boxed{1}\textit{formal}\rangle\end{array}\right]\end{array}\right]\end{array}\right]$$

Figure 5.7: Semantic Content of *kitte* (Formal Projection)

Hasegawa goes on to say that the resultative *chopped'(y)* combines with the Logical Structure of *aru* to yield the structure for *kitte aru* in (19).

(5.19) *kitte aru*
be-at'(x, [chopped'(y)])

She also claims that x is the unexpressed locative of *aru*, so y is (only) an argument of *kiru*. Her claim is, then, that this structure is the realization of a separate lexical form for *aru* which specifically accounts for the conjunct cases. Her two logical structures for *aru* are given here in (20).

(5.20) *aru* (be)
(a) be-at'(x,y) x=locative, y=theme.
(b) be-at'(x,[LS]) x=locative.

So in Hasegawa's analysis of the sentence given here as example (16), *yasai* (vegetables) is the expressed *theme* of the first conjunct verb *kitte* (chopped) and the entire resultative predicate *yasai ga kitte* is the argument of *aru*, not identified by a macrorole; the locative role is unexpressed. This contrasts with Matsumoto's (1990) control analysis of *te* conjunctive forms, where the subject of the *te* form is unexpressed, and the expressed, controlling subject is that of the higher predicate, *aru* in this case. Matsumoto (1990) accords with my analysis of *control* constructions in the last chapter (essentially following Chierchia (1984)). Under Hasegawa's analysis, *yasai* (vegetables) is the expressed theme (UNDERGOER) of *aru*, and the LOCATIVE remains to be accounted for. I suggest that in (20)(b) x is the "theme" and the LS unifies with the LOCATIVE role. I also suggest, following the argument given above, that there is a simplex form of *aru* with *x=theme, y=locative*, using Hasegawa's formalism. The corrected forms of *aru* are given in (21).

(5.21) *aru* (be)
(a) be-at'(x,y) x=locative, y=theme.
(b) be-at'(x,y) x=theme, y=locative.
(c) be-at'(x,[LS]) x=theme, LS=locative.

If the locative role of simplex *aru*, which Hasegawa shows as *be-at'(x,y)*, can be alternated with the logical structure of the embedded predicate, then there is the basis for merging her two logical structures, corrected as in (21)(b) and (c), into a single polysemous form.

In summary, although Hasegawa's development of logical structures do not seem to work out the semantic roles consistently, her essential insight is the alternation of the logical structures in the intransitivization of an Accomplishment. In the next section I develop lexical analyses for the first conjunct forms of Accomplishments, Activities and Achievements, using Pustejovsky's (1995) Event Structure

and Qualia Structure, within the CONTENT of an HPSG sign. Hasegawa's "Intransitivizing" insight dovetails with the distinction between Agentive and Formal Qualia projections in the development of the Accomplishment analysis.

5.4 First Conjunct Verb Forms

5.4.1 Process/Result Polysemy

In Pustejovsky (1995) there is an analysis of English *ion* nominalizations of Accomplishment verbs such as $create_v \longrightarrow creation_n$ showing that the nominalized form has both process and result readings. The verb 'create' has an Event Structure with process and result readings, and a Qualia Structure containing Agentive and Formal projections, but for the finite form of the verb, only the Agentive Quale is projected to syntax, with the obligatory expression of ACTOR and UNDERGOER arguments, shown in (22). The nominalization examples are repeated here with the process reading of 'creation' in (23) and the result reading in (24).

(5.22) He created a big software problem.
(5.23) The creation started on Monday morning.
(5.24) Her creation is perched on the mantelpiece.

These *ion* nominalizations project no arguments but they do seem to retrieve the process and state readings of the word.

The verb *tukuru* (bake) is a transitive Accomplishment which can appear in first conjunct position in any of the conjunctive forms listed in Chapter 2.

(5.25) Keiko ga keeki o tukuru
Keiko-NOM cake-ACC bake
"Keiko bakes a cake".
(5.26) Keiko ga keeki o tukutte iru
Keiko-NOM cake-ACC bake-TE be
"Keiko is baking a cake".

It has an Event Structure which is bieventual, partially ordered, and the events relate to Agentive and Formal Qualia. The ACTOR and UNDERGOER roles are linked to subject and complements, respectively, following Davis's (1995) linking theory. A lexical entry for finite *tukuru* is shown in Figure 5.8.

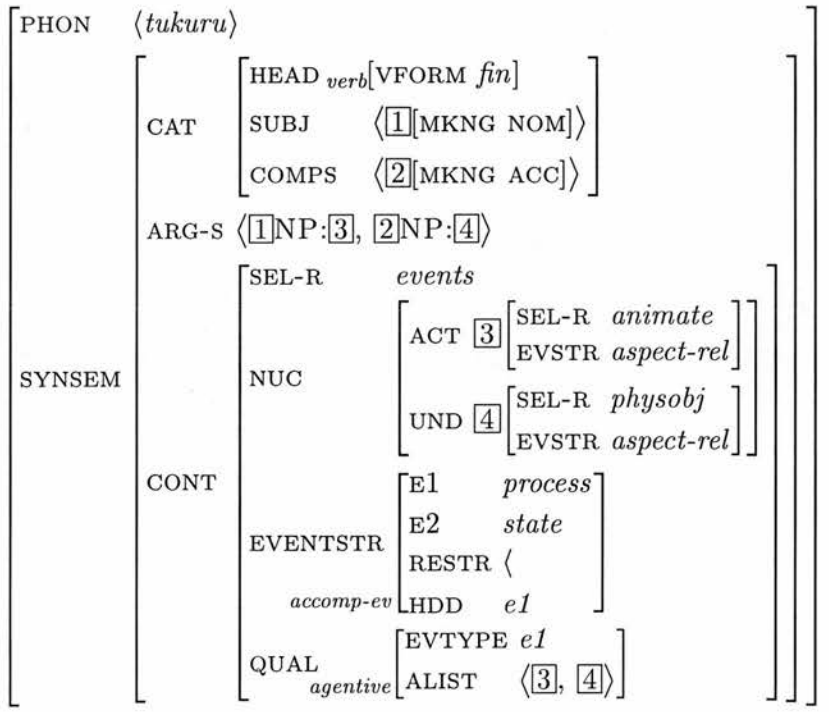


Figure 5.8: Lexical Entry for *tukuru* (with Agentive Projection)

The finite form projects both roles as the arguments. The *te* form, *tukutte* (baking) however, projects only the undergoer and suppresses the subject, after the pattern of Postal's (1974) analysis of English *ing* forms. In both cases, finite and *te* form, the predicate is E1 headed, and the Agentive, not the Formal, Quale is projected. Pustejovsky shows that in an Accomplishment, both Agentive and Formal Qualia are necessary parts of the lexical semantic structure, and yet only the Agentive Quale is projected. Hasegawa suggests that the Stative predicate resulting from the Intransitivizing process is also projected, and this is represented as the Formal Quale in Pustejovsky's system.

(5.27) keeki ga tukutte aru
cake-NOM bake-TE be
"The cake is baked".

How then, does the lexical representation of these qualia alternations play out?

The intransitivizing *te aru* form projects the undergoer of the first conjunct as subject, but as we have seen earlier, the *te* morpheme suppresses the projection of a subject, and its role is represented semantically as a (Pustejovskian) default. The subject projected in (26) above is the UNDERGOER of *iru*. Similarly, the subject projected in (27) below is the UNDERGOER of *aru*

The most prominent role of the embedded Accomplishment is not directly projected: but following the control analysis of *te miru* in the last chapter, this unexpressed subject is controlled by the subject of the higher predicate. Both *te iru* and *te aru* then, are control constructions.

5.4.1.1 Sense Enumerative Forms

In addition to providing a marking value underspecified for accusative and dative case, what the *te* morpheme seems to do is systematically suppress the projection of the leading argument, whether it is the ACTOR in a transitive predicate, or the UNDERGOER in an intransitivized form. The selection of which Quale to be projected, however, seems to be licensed by the head verb. Except for *aru*, all V2 verbs select E1; *aru*, alone, selects E2. Hasegawa's intransitivizing process is a potential of Accomplishments, but it is realized only by the selectional restrictions of the head verb. The lexical representation of the *te* form must account for both projections to syntax. Unfortunately there is a problem in Pustejovsky's representation of the arguments in his "argument structure" (HPSG's NUCLEUS), which result in the Agentive and the Formal Qualia being realized as separate lexical entries. In the Agentive Quale, the ACTOR is a default argument, represented in the NUCLEUS as D-ARG, not projected to subject, and the UNDERGOER is projected to complement with accusative case. This is the form for Accomplishment *tukutte* (baking/baked) shown as Figure 5.9(a). In the Formal Quale selected through E2 headedness of the Event Structure, the ACTOR is shadowed, represented as S-ARG in Figure 5.9(b), and the UNDERGOER is a default, represented as D-ARG, and neither of these roles is projected to syntax. The argument structure (ARG-S) list of the Agentive contains a pronominal and the UNDERGOER, but the ARG-S list of the Formal simply contains the (unprojected) pronominal.[8]

[8] A deeper investigation into the relationships between Pustejovsky's semantics, HPSG and projection to syntax might well show up some systematic relationship between Pustejovskian *defaults* and *pro*. However I will make no explicit arguments for relating these things in *control* constructions.

(a):

$$\begin{bmatrix}
\text{PHON } \langle tukutte \rangle & \\
\text{SYN} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \text{HEAD } teform[\text{VFORM } teverb] \\ \text{SUBJ} \quad \langle \boxed{1} pro \rangle \\ \text{COMPS} \quad \langle \boxed{2}[\text{MKNG ACC}] \rangle \end{bmatrix} \\
\text{ARG-S} & \langle \boxed{1}{:}pro,\ \boxed{2}\text{NP}{:}\boxed{4} \rangle \\
\text{CONT} & \begin{bmatrix}
\text{SEL-R} \quad events & \\
\text{NUC} & \begin{bmatrix} \text{D-ACT } \boxed{3} \begin{bmatrix} \text{SEL-R} \quad animate \\ \text{EVSTR} \quad aspect\text{-}rel \end{bmatrix} \\ \text{UND} \quad \boxed{4} \begin{bmatrix} \text{SEL-R} \quad physobj \\ \text{EVSTR} \quad aspect\text{-}rel \end{bmatrix} \end{bmatrix} \\
\text{EVSTR} & \begin{bmatrix} \text{E1} \quad process \\ \text{E2} \quad state \\ \text{RESTR} \ \langle \\ \text{HDD} \quad e1 \end{bmatrix}_{accomp\text{-}ev} \\
\text{QUAL} & \begin{bmatrix} \text{EVTYPE } e1 \\ \text{ALIST} \quad \langle \boxed{3}, \boxed{4} \rangle \end{bmatrix}_{agentive}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}$$

(b):

$$\begin{bmatrix}
\text{PHON } \langle tukutte \rangle & \\
\text{SYN} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \text{HEAD } teform[\text{VFORM } teverb] \\ \text{SUBJ} \quad \langle \boxed{1} pro \rangle \\ \text{COMPS} \quad \langle\ \rangle \end{bmatrix} \\
\text{ARG-S} & \langle \boxed{1}{:}pro \rangle \\
\text{CONT} & \begin{bmatrix}
\text{SEL-R} \quad events & \\
\text{NUC} & \begin{bmatrix} \text{S-ACT} \quad \begin{bmatrix} \text{SEL-R} \quad animate \\ \text{EVSTR} \quad aspect\text{-}rel \end{bmatrix} \\ \text{D-UND } \boxed{4} \begin{bmatrix} \text{SEL-R} \quad physobj \\ \text{EVSTR} \quad aspect\text{-}rel \end{bmatrix} \end{bmatrix} \\
\text{EVSTR} & \begin{bmatrix} \text{E1} \quad process \\ \text{E2} \quad state \\ \text{RESTR} \ \langle \\ \text{HDD} \quad e2 \end{bmatrix}_{accomp\text{-}ev} \\
\text{QUAL} & \begin{bmatrix} \text{EVTYPE } e1 \\ \text{ALIST} \quad \langle \boxed{4} \rangle \end{bmatrix}_{formal}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}$$

Figure 5.9: Lexical Entries for *tukutte* with Agentive and Formal Projections

5.4.1.2 *Parameterizing the Qualia*

The problem with Pustejovsky's use of D-ARG and S-ARG in argument structure is when they come to be encoded as features in the HPSG NUCLEUS, and encoding these labels directly leads to the projection of sense enumerative lexical forms, as shown in Figures 5.9(a) and (b), for the Agentive and Formal projections of *tukutte* (baking/baked). Since the status of the nuclear arguments is determined by the Quale projected, this is more accurately represented if the projection status in the NUCLEUS is underspecified, and separately specified for each quale. Thus the projection type *anyproj* has subtypes *default*, *shadow* and *project*, shown in Figure 5.10.

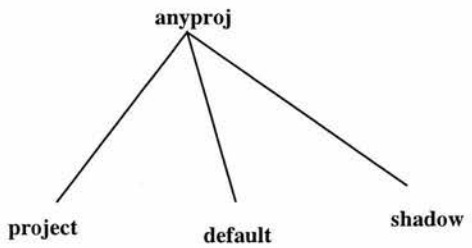


Figure 5.10: Type Hierarchy for Projection from Qualia

The value of *projs* in the NUCLEUS for any argument is *anyproj*, and the value differs in qualia, depending on the number of arguments in the quale, and whether the verbform is finite or non-finite. An Agentive projection of a finite form such as *tukuru* projects both ACTOR and UNDERGOER, so the projection value is *project*; the non-finite form *tukutte* suppresses the ACTOR, so its projection status is *default*. The Formal Quale is selected by *aru*, and its animate argument is suppressed, or quantified over, as a shadow argument, and has no direct effect on the syntax or the meaning of the predicate. With the operation of *te* the UNDERGOER argument is suppressed, represented as *default* projection status in qualia.

After rationalizing the *default*, *shadow* and *true* arguments, the Agentive and Formal projections of *tukutte* are shown in Figures 5.11(a) and (b).

(a):

$$\left[\begin{array}{l}\text{PHON } \langle tukutte \rangle \\ \text{SYN} \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD } teform[\text{VFORM } teverb] \\ \text{SUBJ} & \langle \boxed{1} pro \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG ACC}] \rangle\end{array}\right] \\ \text{ARG-S} & \langle \boxed{1}{:}pro,\ \boxed{2}\text{NP}{:}\boxed{4} \rangle \\ \text{CONT} & \left[\begin{array}{ll}\text{SEL-R} & events \\ \text{NUC} & \left[\begin{array}{ll}\text{ACT } \boxed{3} & \left[\begin{array}{ll}\text{PROJS} & anyproj \\ \text{SEL-R} & animate \\ \text{EVSTR} & aspect\text{-}rel\end{array}\right] \\ \text{UND } \boxed{4} & \left[\begin{array}{ll}\text{PROJS} & anyproj \\ \text{SEL-R} & physobj \\ \text{EVSTR} & aspect\text{-}rel\end{array}\right]\end{array}\right] \\ \text{EVSTR} & \left[\begin{array}{ll}\text{E1} & process \\ \text{E2} & state \\ \text{RESTR} & \langle \\ accomp\text{-}ev\ \text{HDD} & e1\end{array}\right] \\ \text{QUAL} & \left[\begin{array}{ll}\text{EVTYPE} & e1 \\ agentive\ \text{ALIST} & \langle [\text{PROJS } default]{:}\boxed{3},\ [\text{PROJS } project]{:}\boxed{4} \rangle\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

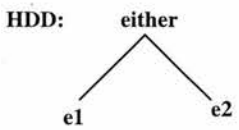
(b):

$$\left[\begin{array}{l}\text{PHON } \langle tukutte \rangle \\ \text{SYN} \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD } teform[\text{VFORM } teverb] \\ \text{COMPS} & \langle\ \rangle\end{array}\right] \\ \text{ARG-S} & \langle pro \rangle \\ \text{CONT} & \left[\begin{array}{ll}\text{SEL-R} & events \\ \text{NUC} & \left[\begin{array}{ll}\text{ACT } \boxed{3} & \left[\begin{array}{ll}\text{PROJS} & anyproj \\ \text{SEL-R} & animate \\ \text{EVSTR} & aspect\text{-}rel\end{array}\right] \\ \text{UND } \boxed{4} & \left[\begin{array}{ll}\text{PROJS} & anyproj \\ \text{SEL-R} & physobj \\ \text{EVSTR} & aspect\text{-}rel\end{array}\right]\end{array}\right] \\ \text{EVSTR} & \left[\begin{array}{ll}\text{E1} & process \\ \text{E2} & state \\ \text{RESTR} & \langle \\ accomp\text{-}ev\ \text{HDD} & e2\end{array}\right] \\ \text{QUAL} & \left[\begin{array}{ll}\text{EVTYPE} & e1 \\ formal\ \text{ALIST} & \langle [\text{PROJS } shadow]{:}\boxed{3},\ [\text{PROJS } default]{:}\boxed{4} \rangle\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

Figure 5.11: Lexical Entries for *tukutte* with Agentive and Formal Projections

As with the question of whether the separate projections of *iru* (animate be) are polysemous (in Section 5.1), the question for Agentive and Formal alternations is whether these can be represented in a single, polysemous, lexical entry. The values of CONTENT in the two alternatives shown in the figures differ only in their *headedness*: E1 for the Agentive projection and E2 for the Formal. These are unified in the single, underspecified, value *HDD: either*, as E1 and E2 are the two subtypes of the headedness value *either*.

Figure 5.12: Headedness Values:



The projections to syntax of the Qualia involve alternations in the complements value under CAT (or under *valence*, in terms of the exploratory analysis of simplex *iru* in Section 5.1) in the value of ARG-S. These alternations for Accomplishment *tukutte* (baking/baked) are illustrated in Figure 5.13, with (a) as the Agentive projection, showing the accusative marked argument as complement, and the sole element in ARG-S; and in (b), both ARG-S and COMPS are empty.

(a):
$$\left[\begin{array}{l}\text{SYN}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{SUBJ} & \langle \boxed{1} pro \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG } dat] \rangle\end{array}\right] \\ \text{ARG-S} & \langle \boxed{1}{:}pro,\ \boxed{2}\text{NP}{:}\boxed{4} \rangle \\ \text{CONT} & [\text{EVSTR}[\text{HDD } e1]]\end{array}\right]\end{array}\right]$$

(b):
$$\left[\begin{array}{l}\text{SYN}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{SUBJ} & \langle \boxed{1} pro \rangle \\ \text{COMPS} & \langle\,\rangle\end{array}\right] \\ \text{ARG-S} & \langle \boxed{1}{:}pro \rangle \\ \text{CONT} & [\text{EVSTR}[\text{HDD } e2]]\end{array}\right]\end{array}\right]$$

Figure 5.13: Projections from Qualia Alternations.

While I argued in Section 5.1 that subjects and complements might be more conveniently represented as scalar values, rather than lists, it is unlikely that the same argument can be sustained for the ARG-S list. Argument Structure is the locus of binding (in addition to linking) and there may be more list elements subject to binding constraints than there are arguments licensed by the predicate. Thus, the list notation for ARG-S cannot be explained away. It is more convenient, then, to represent the alternative *te* forms of an Accomplishment by enumerating the senses.

5.4.2 Activity Projections

The *te* form of an Activity is selected by both *iru* and *aru*, and *te iru* is interpreted as the progressive (28), in line with the same form of an Accomplishment, whereas the *te aru* form is interpreted as a sort of 'past perfect' (29), although this form is subject to mixed reactions from native speakers. It is my contention that the lack of a resultative reading (that is, the lack of a Formal Quale) leads to this 'funny' reaction. I will argue for this contention after developing a lexical structure for the *te* form of an Activity.

(5.28) kodomotachi ga asonde iru.
children-NOM play-TE be
"The children are playing".

(5.29) watashi wa nete aru wa yo
I-TOP sleep-TE be TOP EMPH
"I have slept a lot". (Hasegawa 1995)

In Pustejovsky's review of English Activities, he claims that a verb such as 'walk' has an Event Structure in which two subevents overlap: $e1 \circ_\alpha e2$, without, however, showing the Qualia associated with each subevent. Dowty's (1979) claim is that an Activity can be modelled with a predicate **do** in what amounts to a single subeventual structure. The verb 'walk' has only a single Agentive projection, in contrast with Accomplishments and Achievements which have both Agentive and Formal projections possible, so modelling the overlapping structure of an Activity is problematic. Projection from a single Agentive Quale only, in fact suggests that Dowty's view, of a monoeventual structure for Activities (cf Table 4.1(b)), is more likely to be correct. Applying this discussion to Japanese, then, if an unergative verb such as *odoru* (dance) has an Agentive Quale projecting subevent E1 with a single ACTOR argument, then the Event Structure needed to support this is a monoeventual structure, with a process event. The finite form of the unergative verb *odoru* projects the actor argument to subject, and following Matsumoto (1990) the *te* form suppresses the subject projection. So, following Pustejovsky (1995), the argument is encoded in the *nucleus* as a default. The lexical semantic structure of *odotte* (dancing) is shown as Figure 5.14.

This Activity structure is selected as an *e1* headed event by *iru* and *aru* in conjunctive forms. The intransitivizing *te aru* conjunction which selects for *e2* headedness does not select for Activities, aspectually. Selection by *iru* seems to be straightforward, as it selects for the *te* form of an *e1* headed dynamic verb such as *odotte*

$$\left[\begin{array}{ll} \text{CONT} & \left[\begin{array}{ll} \text{NUC} & \left[\begin{array}{ll} \text{RELN} & \textit{odoru} \\ \text{ACT} & \textit{animate} \end{array}\right] \\ \text{EVENTSTR} & \left[\begin{array}{ll} \text{E1} & \textit{process} \\ \text{HD} & \textit{e1} \end{array}\right] \\ \text{QUALIA} & \left[\begin{array}{ll} \text{EVTYPE} & \textit{e1} \\ \text{ALIST} & \langle[\text{PROJS } \textit{default}]\rangle \end{array}\right] \end{array}\right] \end{array}\right]$$

Figure 5.14: Activity Verb *odotte* (dancing)

in the figure. In Chapter 2 I mentioned that the acceptability of *te aru* projection of an Activity receives mixed judgements from native speakers: some speakers find this construction perfectly acceptable (see Hasegawa's (1995) examples), also those of Jacobsen (1991) and Martin (1975) (who are non-native speakers, although they no doubt received native speaker input), while other informants, who I have polled, say that the construction is a little odd, because there is no resultant state. This claim is backed up by the lexical semantic structure of an Activity, which has an Agentive Quale, but no Formal Quale, and the *e1* headed event is selected by both *iru* and *aru*. In an Accomplishment conjunction, even though the *e1* headed event is selected, the *e2* event is still present: *shadowed* in Pustejovsky's (1995) terminology. It is perhaps this shadowed event which contributes something to the meaning of *te aru* constructions, the absence of which some speakers find a little "strange".

5.4.3 Achievement Projections

Following from the analysis of Accomplishment and Activity *te* forms, the hypothesis for argument selection is that the head verb selects for a bundle of features including headedness and aspectual class and these aspectually agentive forms are selected for *e1* headedness by e.g. *iru*. Can this hypothesis be extended to account for Achievements in first conjunct position?

At first blush, the answer is no, since Pustejovsky claims that English Achievements are focused on the *result* state, and are *e2* headed in their Event Structure. Dowty (1979) claims that the Achievement is a transition with the structure *Become(State)*, differing from Accomplishments in that the transition is uncaused. It is, in fact, possible to retrieve both the left aspect and the right aspect of an Achievement such as 'die' using 'be' as an aspectualizing operator. Example (30) shows the left-headed, progressive form, (31) shows the right-headed, perfect form. The key to this analysis is that the progressive and perfect forms of 'die', which are 'dying' and

'dead', respectively, are not homophonous, and so they necessarily require separate lexical entries. The simple form is shown in (32).

(5.30) He is dying.
(5.31) he is dead.
(5.32) he dies.

If we look now at the Japanese Achievement data, the simple form *shinu* (dies), (33) entails both change-of-state and result, the *te iku* form forces a progressive reading in (34), and the *te iru* form gives a perfect reading, in (35).

(5.33) kare ga shinu
he-NOM die
"He dies".
(5.34) kare ga shinde iku
he-NOM die-TE go
"He is dying".
(5.35) kare ga shinde iru
he-NOM die-TE be
"He is dead".

Achievements do not combine with the *te aru* form, and example (36) is inadmissible. Evidently, the selectional restrictions of *aru* select only agentive verbs.

(5.36) *kare ga shinde aru
he-NOM die-TE be

Clearly there is something different about Japanese Achievement verbs which allows both consistent selection by *iru*, along with Activities and Accomplishments (which are *e1* headed), and also perfective reading, and this has to relate to the ordering of events in Event Structure.

The Event Structure possibilities in Pustejovsky's scheme are not limited to partial ordering - there is also an overlap operator, $e1 \circ_\alpha e2$, in which events *e1* and *e2* occur together. In this structure, if a head selects for *e1*, then *e2* occurs simultaneously, thus if *shinu* is an overlapped event and *iru* selects for *e1*, then a perfect reading is available in virtue of the simultaneous occurrence of *e2*. The lexical representation of this is given in Figure 5.15: the verb *shinu* (die) licenses a single, animate argument and in *te* form this argument is a default, not directly projected. The Event Structure is bieventual, with a restriction type of *overlap* ensuring that the events

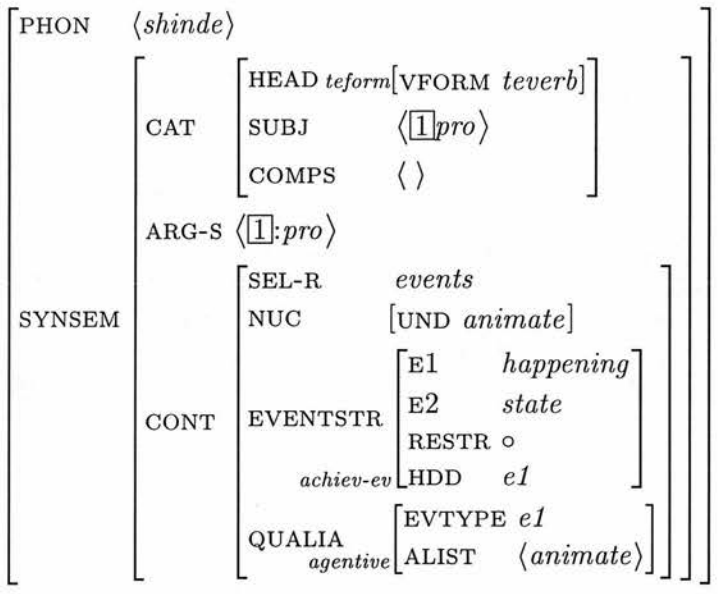


Figure 5.15: Lexical Entry for *shinde*

occur simultaneously, and the structure is *e1* headed. Both Agentive and Formal Qualia associate their events with a single, animate argument[9].

5.4.4 What is te?

In adopting Pustejovsky's system I have merged his argument structure into the HPSG NUCLEUS and retained the default - shadow - true argument designations. The problem is that when the *te* form of an Accomplishment such as *tukutte* emerges from the lexicon as a polysemous item, *te* does not "know" whether ACTOR or UNDERGOER is to be suppressed. The selection of Formal and Agentive Qualia is determined by the total composition of the sentence: the Agentive Quale is projected in the presence of the accusative argument of the first conjunct and the selecting second conjunct, which may be any of the verbs that take V2 position. When the Formal Quale is selected by *aru* (inanimate be), the ACTOR of V1 is suppressed (shadowed), the UNDERGOER is defaulted and, again, not overtly expressed. If *tukutte* is a discrete lexical item, the representation of *default* information is problematic, since in the Agentive case, the ACTOR is represented as *default*, while in the Formal case, the UNDERGOER is *default*. Since one function of *te* is to suppress the leading semantic argument, the question is whether this occurs in the lexicon as part of the same process as morphological attachment, or whether it occurs at

[9]Strictly speaking, "Agentive" is not the right term, since there is no agent involved in an Achievement. The proper concept involves a non-agentive process, not unlike the "non-agentive" Activities which Dowty (1979) identifies, and which raise a problem for the use of the **do** operator.

the same time as selection of the Formal or Agentive Quale by V2, as part of the syntactic composition of the sentence.

If *te* is a marker, then it combines with the underspecified infinitival form *tukuri* under the Head-Marker schema to form the alternate projections given in Figure 5.16, (a) Agentive and (b) Formal. However, the application of *te* to a verb entails an argument structure changing process, and both Bresnan (1995) and Manning *et al.* (1999) argue that such processes are morphological, not syntactic processes, so the simpler explanation is that *te* affixation is explained by derivational morphology. One of the functions of *te* is to give the verb a marking value, though, and in this way it resembles a case marker.

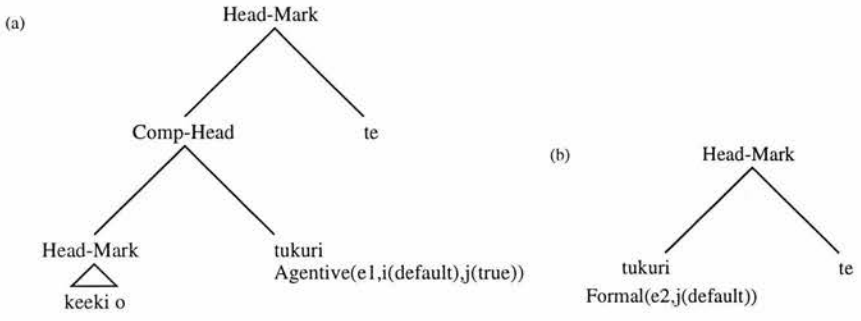


Figure 5.16: *te* Marking, Underspecified for Complement Saturation

Such morphological changes can be explained either by application of a lexical rule, or by a process of derivational morphology such as that described by Riehemann (1995). The lexical rule is described here. Projection to syntax is straightforward for Activities, as there is a single Agentive Quale which links the ACTOR argument with the first position in ARG-S, and this is linked to subject; however because *te* suppresses the leading argument, no arguments appear in the ARG-S (for a 1-place unergative verb) and there is no link to the subject.

An Activity verb such as *odoru* (dance) has an infinitival form *odori* and a *te* form *odotte*. It has a single ACTOR argument and an Agentive projection. The affixation of *te* creates a marking value underspecified for case, and suppresses the leading argument, by making it default, as has been established in the last chapter. The lexical rule for conversion from *odoru* (dance) to *odotte* (dancing) is given in Figure 5.17. The morphological change to the verbform results in the *true* argument being changed to a *default*, and this default is not projected to subject. The case of a transitive Accomplishment is a more complex phenomenon, because it is polysemous, with two possible argument projections, either of which may be selected

$$
\left[\begin{array}{ll}
\text{PHON} & \langle odoru \rangle \\
\text{CAT} & \left[\begin{array}{l}
\text{HEAD } \mathit{verb}[\text{VFORM } \mathit{fin}] \\
\text{SUBJ } \langle \boxed{1}[\text{MKNG } \mathit{nom}] \rangle \\
\text{MARKING } \mathit{unmarked}
\end{array}\right] \\
\text{ARG-S} & \langle \boxed{1}\text{NP}:\boxed{2} \rangle \\
\text{CONT} & \left[\text{QUALIA } \left[\begin{array}{ll}
\text{EVTYPE} & \mathit{e1} \\
\text{ALIST} & \langle \boxed{2}_{\mathit{agentive}}[\text{PROJS } \mathit{project}] \rangle
\end{array}\right]\right]
\end{array}\right]
$$

$$
\Longrightarrow \left[\begin{array}{ll}
\text{PHON} & \langle odotte \rangle \\
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD } \mathit{teform}[\text{VFORM } \mathit{teverb}] \\
\text{SUBJ} & \langle \boxed{1}\mathit{pro} \rangle \\
\text{MARKING} & \mathit{te}
\end{array}\right] \\
\text{ARG-S} & \langle \boxed{1}:\mathit{pro} \rangle \\
\text{CONT} & \left[\text{QUALIA } \left[\begin{array}{ll}
\text{EVTYPE} & \mathit{e1} \\
\text{ALIST} & \langle \boxed{2}_{\mathit{agentive}}[\text{PROJS } \mathit{default}] \rangle
\end{array}\right]\right]
\end{array}\right]
$$

Figure 5.17: A Lexical Rule for Converting *odoru* to *odotte*

by the head verb under syntactic combination. In a lexical operation, the leading argument of both alternative projections is systematically converted to default, and this is shown for *tukuru* $\longrightarrow$ *tukutte* in Figure 5.18.

5.5 The Lexical Forms of *aru*

Among Nuclear Conjunctions, *te aru* is unique in being associated with both the Agentive and Formal projections of a transitive Accomplishment, based on the polysemous analysis of Accomplishments given in the last section, from which example (37) is analysed as a projection of the Formal Quale of *shimete* (closed), and example (38) projects the Agentive Quale (closing).

(5.37) doa ga shimete aru.
door-NOM close-TE be
"The door is closed".

(5.38) Maiko-san ga doa o shimete aru.
Maiko-NOM door-ACC close-TE be
"Maiko has closed the door".

$$\left[\begin{array}{l}\text{PHON } \langle tukuru \rangle \\ \text{CONT } \left[\text{QUAL } \left[\begin{array}{l}\text{EVTYPE } e1 \\ \text{ALIST } \quad \langle \boxed{1}[\text{PROJS } project], \boxed{2}_{agentive}[\text{PROJS } project] \rangle\end{array}\right]\right]\end{array}\right]$$

$$\Longrightarrow \left[\begin{array}{l}\text{PHON } \langle tukutte \rangle \\ \text{CONT } \left[\text{QUAL } \left[\begin{array}{l}\text{EVTYPE } e1 \\ \text{ALIST } \quad \langle \boxed{1}[\text{PROJS } def], \boxed{2}_{agentive}[\text{PROJS } proj] \rangle\end{array}\right]\right]\end{array}\right]$$

Figure 5.18: A Lexical Rule for Converting *tukuru* to *tukutte*

No other Nuclear Conjunction participates in the E2 headed Formal projection of an Accomplishment, but *aru*, alone, alternately participates in Valence-Maintaining and Valence-Reducing conjunctions, and selects the E1 or the E2 headed event of the first conjunct verb, respectively. Simplex *aru*, it is recalled, is a Stative verb projecting inanimate UNDERGOER and LOCATIVE arguments, with the LOCATIVE projected to a dative *ni* marked argument. It would be a convenient and elegant solution if it could be shown that *aru* is three ways polysemous, with one lexical entry selecting for the nominal argument, the E1 headed Agentive projection and the E2 headed Formal projection and on the face of it may be possible to do so, if the selectional restrictions placed by *aru* on its locative argument pattern with Figure 5.20.

$$\left[\begin{array}{l}\text{NUC } \left[\begin{array}{l}\text{RELN } aru \\ \text{UND } \left[\begin{array}{l}\text{SEL-R } physobj \\ \text{EVSTR } aspect\end{array}\right] \\ \text{LOC } \left[\begin{array}{l}\text{SEL-R } phenomena \\ \text{EVSTR } agent\text{-}ev[\text{HDD } either]\end{array}\right]\end{array}\right]\end{array}\right]$$

Figure 5.19: Selectional Restrictions of the Ideal *aru*

What prevents such an analysis is the fact that the alternative projections have different adjunct scope. Thus, the Formal projection patterns with *aru* as a Stative construction, while the Agentive projection patterns with the first conjunct as an Activity or Accomplishment.

(5.39) doa ga shibaraku mae kara / *moo sugu shimete aru.
door-NOM since a while ago / soon close-TE be
"The door has been closed since a while ago/*soon".

(5.40) Maiko-san ga syuunkan ni / *shibaraku mae kara doa o shimete aru.
Maiko-NOM suddenly / since a while ago door-ACC close-TE be
"Maiko has suddenly/*since a while ago closed the door".

$$\left[\begin{array}{l}\text{PHON } \langle aru \rangle \\ \text{SYN } \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{l}\text{HEAD } _{verb}[\text{VFORM } fin] \\ \text{SUBJ } \langle \boxed{1}[\text{MKNG } nom] \rangle \\ \text{COMPS } \langle \boxed{2}[\text{MKNG } dat] \rangle_{ne-list} \\ \text{MARKING } unmarked\end{array}\right] \\ \text{ARG-S } \langle \boxed{1}\text{NP:}\boxed{3}, \boxed{2}\text{NP:}\boxed{4} \rangle & \\ \text{CONT } \left[\begin{array}{ll}\text{SEL-R} & events \\ \text{NUC} & \left[\begin{array}{ll}\text{RELN } aru & \\ \text{UND } \boxed{3} & \left[\begin{array}{l}\text{SEL-R } physobj \\ \text{EVSTR } aspect\end{array}\right] \\ \text{LOC } \boxed{4} & \left[\begin{array}{l}\text{SEL-R } phenomena \\ \text{EVSTR } _{accomp\text{-}ev}[\text{HDD } e2]\end{array}\right]\end{array}\right] \\ \text{EVSTR } _{state\text{-}ev}[\text{E1 } state] & \\ \text{QUALIA } _{formal}\left[\begin{array}{ll}\text{EVTYPE } e1 & \\ \text{ALIST} & \langle \boxed{3}[\text{PROJS } project], \boxed{4}[\text{PROJS } project] \rangle\end{array}\right] & \end{array}\right] & \end{array}\right] & \end{array}\right]$$

Figure 5.20: *aru* Selecting for Formal Quale

Thus, while the argument *doa* (door) is a projection of the conjunction *shimete aru* in (39) with the argument structures of *shimete* and *aru* merged, in (40) *aru* subcategorizes for a saturated VP, and the selectional restrictions of the two projections are incompatible: for the Formal case, *aru* selects for a LOCATIVE argument headed E2, with an unsaturated VP, as illustrated in Figure 5.21. The embedded VP *shimete* projects an UNDERGOER argument, but the operation of *te* suppresses this, and the ARG-S list contains a single pronominal element. In the merger of argument structures, the empty ARG-S of *shimete* is combined with the ARG-S of *aru*, as in (41), and the argument projected to subject is the undergoer of *aru*.

(5.41) ARG-S$\langle pro \rangle \oplus$ ARG-S$\langle NP, VN \rangle \Rightarrow$ ARG-S$\langle NP, VN \rangle$

In the agentive projection, Figure 5.22, *aru* selects for an E1 headed argument with a saturated VP. The agentive projection of *shimete* projects both ACTOR

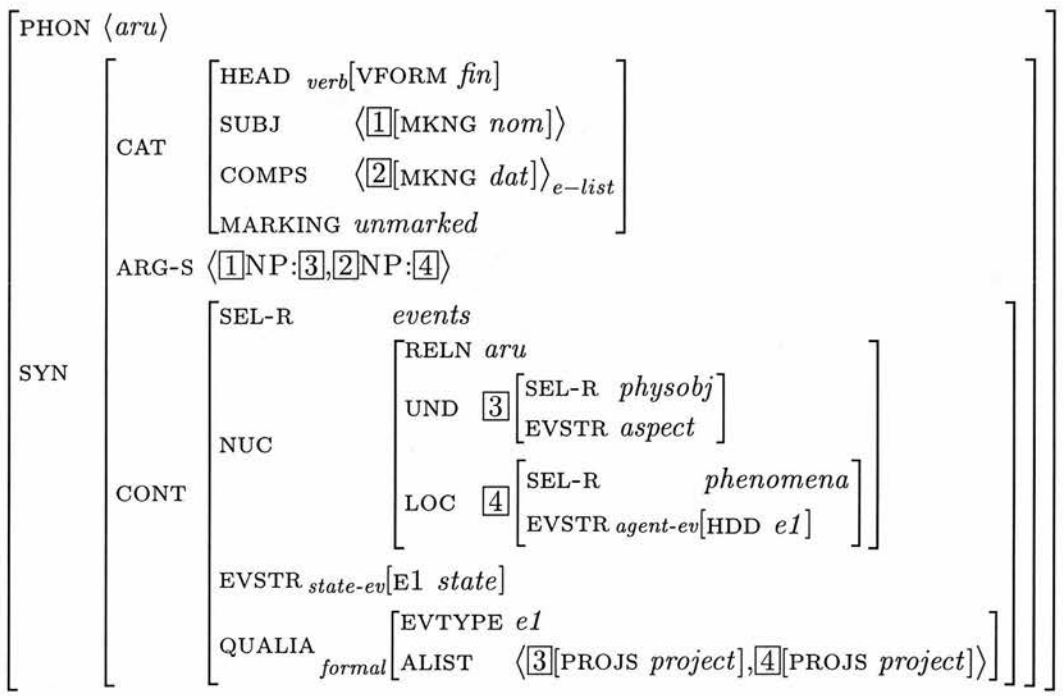


Figure 5.21: *aru* Selecting for Agentive Quale

and UNDERGOER arguments, and *te* suppresses the leading argument, the ACTOR leaving the UNDERGOER projected as accusative complement. The UNDERGOER of *aru* is projected to subject as *Maiko-san ga*, and there is a mismatch between this undergoerhood and the actorhood of the suppressed argument of *shimete*. Since it is not possible for these arguments to structure share, this is perhaps further evidence against such an analysis for *control* constructions.

There is a further distinction to be noted between the simplex and conjunct alternations of *miru* et al, on the one hand, and *aru* and *iru* on the other. This is that the UNDERGOER role of *miru* et al is underspecified for nominal and verbal content, but it is the LOCATIVE role of *aru* et al which is underspecified. The linked complement in *miru* takes accusative marking, and the linked complement in *aru* takes dative marking. This suggests that the *te* form is underspecified for both dative and accusative cases. The marking hierarchy illustrating this underspecification is given in Chapter 4, repeated here as Figure 5.23, with *tecase* as a supertype of dative (*dat*) and accusative (*acc*), respectively. In the lexicon, the marker *ni* is given dative marking value and the marker *o* is given accusative, so that nouns combining with each of these particles are maximally specified as to case; *te* marked verbs take the marking value *tecase*, allowing them to replace either *ni* markers, as with *iru*, or *o* markers as with *miru*, *morau* and *ageru*.

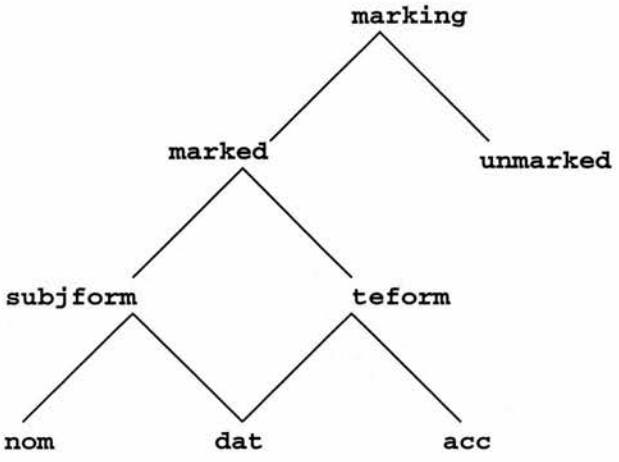


Figure 5.22: The *marking* Hierarchy

5.5.1 Summary of te aru

If the valence alternation for simplex *aru* (inanimate be) (section 5.2) follows that of simplex *iru* (animate be) (Section 5.1), then there are two simplex lexical entries, shown in Figure 5.6 (a) and (b). I have shown in this section that the two projections of conjunct *aru* arise from two further lexical entries. All four lexical forms are briefly characterized as:

(i) *ni* marked subject, nominal valued LOCATIVE.
(ii) *ga* marked subject, nominal valued LOCATIVE.
(iii) *ga* marked subject, verbal valued LOCATIVE, E2 head, unsaturated comps.
(iv) *ga* marked subject, verbal valued LOCATIVE, E1 head, saturated comps.

If the dative complement of (b) is underspecified for *te* marking, and its LOCATIVE value is underspecified for an *eventuality*, then simplex and (one of the) conjunct forms are expressions of the same lexical entry for *aru*: one value is simplex with a dative subject, one is a conjunct form taking an E1 headed agentive verb, and the third form, with a nominative subject and a LOCATIVE argument which admits a nominal object, for the simplex form, or an unsaturated, E2 headed, Accomplishment, for the conjunct form.

5.6 The Lexical Forms of *iru*

The conjunct forms of *iru* combine with Achievements, Activities and Accomplishments. In the first case, the reading is perfect, but in the other two cases it is progressive. In Section 5.4 I developed lexical analyses for these first conjunct *te*

forms which exploit the use of the event headedness value as a selectional restriction. Pustejovsky's (1995) view is that (for English) Achievements are E2 headed events, though Accomplishments and Activities are E1 headed. Since he does not develop a rigorous syntactic model, the question of consistency in selectional restrictions does not arise. If the form of *iru* (animate be) which participates in nuclear Conjunction is the same for all first conjunct complements (as seems reasonable), then the selectional restrictions must be uniform. By analysing the Event Structure of an Achievement as an E1 headed, overlapping event, as I have done in Subsection 5.4.3, *iru* is able to select E1 headedness uniformly. This is captured in the LOCATIVE role of the NUCLEUS by selecting for an Event Structure (EVSTR) of type *dynamic*, with a headednes value of E1, as shown in Figure 5.23. This conjunct form of *iru* has the LOCATIVE role taking the *te* marked verb, and the UNDERGOER role projected to *ga* marked subject.

$$\left[\begin{array}{ll}
\text{PHON} & \langle iru \rangle \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \boxed{0}\left[\begin{array}{ll}\text{HEAD} & _{verb}[\text{VFORM } fin] \\ \text{SUBJ} & \langle \boxed{1}[\text{MKNG } nom] \rangle \\ \text{COMPS} & \langle \boxed{2}[\text{MKNG } dat] \rangle \\ \text{MARKING} & unmarked \end{array}\right] \\
\text{ARG-S} & \langle \boxed{1}\text{NP:}\boxed{3}, \boxed{2}\text{NP:}\boxed{4} \rangle \\
\text{CONT} & \left[\begin{array}{ll}
\text{NUC} & \left[\begin{array}{ll}\text{RELN} & iru \\ \text{UND} & \boxed{3}\left[\begin{array}{ll}\text{SEL-R} & animate \\ \text{EVSTR} & aspect\text{-}rel \end{array}\right] \\ \text{LOC} & \boxed{4}\ _{eventuality}\left[\begin{array}{ll}\text{SEL-R} & phenomena \\ \text{EVSTR} & _{dyn}[\text{HDD } e1] \end{array}\right] \end{array}\right] \\
\text{EVENTSTR} & _{state\text{-}ev}[\text{E1 } state] \\
\text{QUALIA} & _{formal}\left[\begin{array}{ll}\text{EVTYPE} & e1 \\ \text{ALIST} & \langle \boxed{3}[\text{PROJS } project], \boxed{4}[\text{PROJS } project] \rangle \end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]$$

Figure 5.23: *iru* Underspecified for Simple and Conjunct Complements

There are two lexical forms for simplex *iru*, with alternate projections to subject, illustrated in Figure 5.1. If the complements are the LOCATIVE role of the (*ga* marked subject form) of *iru*, in 5.1(a), can be underspecified for nominal and verbal arguments, in the same way as the UNDERGOER of *miru* (see), et al, then the conjunct taking behaviour of *iru* can be explained as at least partially polysemous. The selectional restrictions on the LOCATIVE object as a nominal are not constrained as

to type, so a SEL-R value of *phenomena*[10] is possible. Similarly, the Event Structure of this nominal argument is maximally underspecified as *aspect-rel*, compatible with all nominals. Generalizing the COMPS and LOCATIVE values in the lexical entry for *iru*, the dative nominal and the *te* marked verbal are compatible with a complement of type *vn*[11]. The nominal object (*nom-obj*) and the verbal *psoa* are compatible with a CONTENT type of *eventuality*. The lexical entry for *iru* which is polysemous between simplex and conjunct forms is illustrated in Figure 5.25. This is taken together with the sense enumerative forms of simplex *iru* in Figure 5.1, and is in fact the generalization of Figure 5.1(a). Whether it is possible to describe all the behaviours of *iru* in one single lexical entry depends on whether a formalism can be developed, flexible enough to capture all polysemous behaviours. The present state of the art leads to the specification of *iru* with a minimum of two lexical entries: those given in Figures 5.1(b), and 5.25.

5.7 Remaining Verbal Alternations

The remaining verbal predicates for which data are presented in Chapter 2 include verbs *shimau* (put), *oku* (put), *iku* (go) and *kuru* (come), and adjective *hoshii* (want). The verbs which are most clearly homonymous are *shimau* and *oku* and adjective *hoshii*. The rationalization for homonymous analyses of these predicates is given in Subsection 5.7.1. The verbs *iku* and *kuru* participate in Valence-Maintaining conjunctions with Achievement, Activity and Accomplishment verbs as first conjunct, and these alternate with a LOCATIVE argument in ways which can be described with polysemous lexical entries. There is, however, an idiosyncratic construction which adds an extra LOCATIVE argument in conjunction with certain verbs such as *motsu* (hold) and *tureru* (accompany), and this prevents a completely polysemous analysis of *iku* and *kuru*. These verbs are discussed in Subsection 5.7.2.

5.7.1 shimau,oku and hoshii

The differences between simplex and conjunct forms of the predicates *shimau*, *oku* and *hoshii* all lead to homonymous lexical analyses, but in different and idiosyncratic ways. In simplex form *oku* is a ditransitive verb meaning 'to put something somewhere', as in example (42).

[10]See the *phenomena* type hierarchy in Figure 4.18, and the discussion of categorial types in Section 4.2.3.

[11]See the *head* type hierarchy in Figure 4.16.

(5.42) haha ga hondana ni kabin o oita.
mother-NOM bookshelf-DAT vase-ACC put-PAST
"My mother put the vase on the bookshelf".

In conjunct form it combines with an Activity or an Accomplishment, examples (43) and (44).

(5.43) boku ga nete oita.
I-NOM sleep-TE put-PAST
"I have slept (in preparation for something)".

(5.44) Maiko ga hoteru ni nimotsu o nokoshite oita.
Maiko-NOM hotel-DAT luggage-ACC leave-TE put
"Maiko left her luggage at the hotel".

In these cases the dative argument of *oku* is not present (the dative argument in example (44) is that licensed by *nokosu* (leave something)), so the conclusion is that the conjunct form of *oku* is a transitive verb which takes an agentive subject and a *te* marked complement. This is in contrast with the ditransitive simplex form, and so there are separate lexical entries for each.

As a simplex verb, *shimau* is already homonymous, having the synchronic ditransitive form meaning 'put something somewhere', example (45), and also a transitive form which survives in idiomatic expressions such as *mise o shimau* (to quit business, or close up the shop), in (46).

(5.45) boku ga mise o shimatta.
I-NOM shop-ACC put-PAST
"I closed up the shop". (idiom)

(5.46) Maiko ga hikidashi ni enpitsu o shimatta.
Maiko-NOM drawer-DAT pencil-ACC put-PAST
"Maiko put the pencil in the drawer".

Both these simplex forms take the purposive adjunct *shinchyou ni* (deliberately) and so they are agentive, in examples (47) and (48).

(5.47) boku ga shinchyou ni mise o shimatta.
I-NOM deliberately shop-ACC put-PAST
"I deliberately closed up the shop". (idiom)

(5.48) Maiko ga shinchyou ni hikidashi ni enpitsu o shimatta.
Maiko-NOM deliberately drawer-DAT pencil-ACC put-PAST
"Maiko put the pencil in the drawer on purpose".

The conjunct form of *shimau* combines with Achievement, Accomplishment or Activity verbs, but it yields an Achievement reading aspectually in all these cases. This is exemplified in (49) with an Accomplishment first conjunct together with the non-agentive adverbial *guuzen ni* (accidentally).

(5.49) Maiko ga guuzen ni tegami o sutete shimatta.
Maiko-NOM accidentally letter-ACC throw-away-TE put-PAST
"Maiko accidentally discarded the letter".

The above range of behaviour yields three separate lexical entries for *shimau*, which is, then, distinctly homonymous.

The projections of *hoshii* evidence a still different range of behaviour. As a simplex predicate *hoshii* is an adjective meaning 'want', with two nominative arguments, (50).

(5.50) Kukiko ga keeki ga hoshikatta.
Kukiko-NOM cake-NOM want-PAST
"Kukiko wanted a cake".

What is unique about *hoshii* is that it participates in *te* conjunction in the pattern of *te morau*, but with a meaning identical with the desiderative form, *moraitai* (want to receive). These two are compared in examples (51) and (52).

(5.51) Oguchi ga Maiko ni ie o tatete moraitai.
Oguchi-NOM Maiko-DAT house-ACC build-TE receive-DESID
"Oguchi wants Maiko to build a house for him".

(5.52) Oguchi ga Maiko ni ie o tatete hoshii.
Oguchi-NOM Maiko-DAT house-ACC build-TE want
"Oguchi wants Maiko to build a house for him".

The valence pattern of *te hoshii* is also identical with that of *te morau/moraitai*, with nominative, dative and *te* form arguments. Although one can speculate about how the correspondence between these constructions came about, the outcome for *hoshii* is that there are two separate lexical forms, the simplex with two nominative

arguments and the conjunct with (*ga, ni, te*) marked arguments, and these can be described only with separate, homonymous, lexical forms.

5.7.2 iku and kuru

The verbs of motion *iku* (go) and *kuru* (come) combine with a dative goal argument, or alternately a *te* marked dynamic verb. In the simplex form, the basic meaning is that of 'to go' or 'to come', example (53).

(5.53) Oguchi ga eki ni iku.
Oguchi-NOM station-DAT go
"Oguchi goes to the station".

In conjunct form a range of interpretations is possible, all centering on the actual or metaphorical motion, and all with the same *VP-te iku/kuru* construction. A metaphorical going away is involved in conjunction with Achievement verbs *shinu* (die) or *kieru* (be extinguished), in examples (54) and (55).

(5.54) chichi ga dandan shinde ikimashita.
father-NOM gradually die-TE go-POL-PAST
"My father gradually died".

(5.55) denki ga dandan kiete itta.
light-NOM gradually extinguish-TE went
"The light gradually went out".

Martin (1975) points out that conjunction with a non-motion verb may be interpreted as an errand, and errands are lexicalized differently in English and Japanese. For example one can say in English, as in (56) 'I will go and buy a newspaper', where in the Japanese interpretation, the motion happens after the buying event, literally 'I will buy a newspaper and come'.

(5.56) boku ga shinbun o katte kuru.
I-NOM newspaper-ACC buy-TE come
"I will go and buy a newspaper".

Conjunction with a motion verb has the interpretation 'go V-ing' or 'come V-ing', so in (57) "Saigo went climbing on Sakurajima" or (58) "Mitsuko came walking down the hall"

(5.57) Saigo ga Sakurajima o nobotte itta.
Saigo-NOM Sakurajima-ACC climb-TE went
"Saigo went climbing on Sakurajima".

(5.58) Mitsuko ga rooka o aruite kita.
Mitsuko-NOM hall-ACC walk-TE came
"Mitsuko came walking down the hall".

There is no reason to suppose that the lexical structure of these conjunctions (54) to (58) is any different though, and the basic alternation between simplex and conjunct forms is given in the stylized examples (59) and (60).

(5.59) NP ga NP ni iku/kuru.

(5.60) NP ga VP-te kuru.

Polysemous lexical entries can then be developed for *iku* and *kuru* in which the locative role in the NUCLEUS, linked to the dative complement, is underspecified for *nom-obj* and *psoa content* types, identically with the lexical entry for *iru* (animate be) given in Figure 5.24. The Event Structure differs of course, because *iku* and *kuru* are Accomplishments, not States.

What is truly anomalous about the verbs *iku* and *kuru* is the way in which they form Valence-Increasing conjunctions with certain idiosyncratic verbs such as Activities *motsu* (hold), *tureru* (accompany) and *kiru* (wear) and Accomplishment *nageru* (throw). In all these cases, the simplex form of the first conjunct is a transitive verb with nominative and accusative marked arguments, and they do not take a locative, as exemplified for *motsu* in (61) and (62).

(5.61) boku ga hon o motsu.
I-NOM book-ACC hold
"I hold a book".

(5.62) *boku ga gakkoo ni hon o motsu.
I-NOM school-DAT book-ACC hold
"*I hold a book to school".

In conjunction with *iku* or *kuru* however, the dative argument is licensed: for *motte iku* in (63) and for *turete iku* in (64).

(5.63) boku ga gakkoo ni hon o motte iku.
I-NOM school-DAT book-ACC hold-TE go
"I take a book to school".

(5.64) boku ga ensou ni Naoko o turete iku.
I-NOM concert-DAT Naoko-ACC accompany-TE go
"I take Naoko to a concert".

In the lexicalization of *iku* and *kuru*, the presumption is that the two arguments are licensed, a nominative and a dative, linked to ACTOR and LOCATIVE roles respectively, and that the LOCATIVE role licenses the alternation of a *ni* marked nominal, or a *te* marked dynamic verb. This lexicalization follows the general pattern established for polysemous analyses, and used throughout this chapter. In the cases exemplified by (63) and (64) the extra locative argument *gakkoo ni/ensou ni* is not licensed by the first conjunct, and the locative role of *iku/kuru* is already accounted for by the *te* form alternation. How, then, is the extra argument licensed? Lexicalist theory requires that all arguments be licensed by discrete lexical entries, and not from syntactic conjunctions, and the constraint which ensures this in HPSG is that the argument structure (ARG-S) list is an attribute of lexical entries not shared by any phrasal form. On this basis, then, it is not possible that *turete iku* is a ditransitive predicate which licenses ACTOR, UNDERGOER and LOCATIVE roles linked to nominative, accusative and dative arguments. There would seem to be no way to account for this construction in current lexicalist theory. So, while much of the behaviour of the verbs *iku* and *kuru* can be accounted for with polysemous lexical entries, there is one Valence-Increasing form which cannot currently be described.

5.8 Summary

While the last chapter shows the entire complement taking behaviour of some verbs can be captured with polysemous lexical entries, in this chapter I have shown that there are verbals whose behaviour can only be described with homonymous lexical entries: *shimau* (put back; complete; regret), oku (put; prepare), and *hoshii* (want) all seem to fall in this category. There are also verbs such as *iru* (animate be) and *aru* (inanimate be), whose complex range of argument taking behaviours is limited in its description by the formalism I have chosen. HPSG is quite flexible and powerful enough to capture a great deal of complex and polysemous phenomena, through the use of underspecification and multiple inheritance in the type hierarchy, but is less flexible when it comes to the description of syntactic linking alternations. Part of this inflexibility is no doubt due to the choice of a list structure representation for subjects and complements, when the cardinality of such lists has a predictable and limited range. For example in Japanese, as well as English, there only ever seems

to be one subject[12]. Moreover, the use of a list valued argument provides a gap in the constraint mechanism, because a list can take any value, so mistakes in lexicon development cannot be caught through type-checking mechanisms.

At the limits of the power and flexibility of a formalism such as HPSG, phenomena which are justifiably characterized as polysemous, must be represented by enumerating the senses in a Sense Enumerative Lexicon (*à* la Pustejovsky (1995)).

[12] Although one possible interpretation of the multiple *ga* construction in Japanese verbs is that they are both subjects.

Chapter 6

Phrasal Argument Structure

The forms of Nuclear Conjunction exhibit a further difference in the way that aspectual adjuncts attach[1]. Adverbs which attach to the *V-te miru*, *V-te morau* amd *V-te ageru* forms are compatible with the first conjunct verb, but not with the second. In contrast, adverbs attaching to *V-te iru* and *V-te aru* forms are compatible with the second, not the first conjunct.

The finite form of a transitive Accomplishment such as *tukuru* (bake), example (1), is shown in Figure 6.1(a), with the UNDERGOER (*keeki*) projected as complement, and the ACTOR (*Yoshiko*) projected as subject.

(6.1) Yoshiko ga keeki o tukuru.
Yoshiko-NOM cake-ACC make
"Yoshiko bakes a cake".

Accomplishments combine with the telic adjunct *1-jikan de* (in an hour) and the adjunct scrambles freely with the arguments, suggesting that the adjunct takes scope over the entire argument structure. The above sentence is shown with its telic adjunct attached, and examples (2) to (4) show that this adjunct can take any position before the verb. Adjuncts combine with their heads in HPSG to form a phrasal structure which has the semantic content of the adjunct, by operation of the Semantics Principle, and the syntactic content of the head, by the Head Feature Principle.

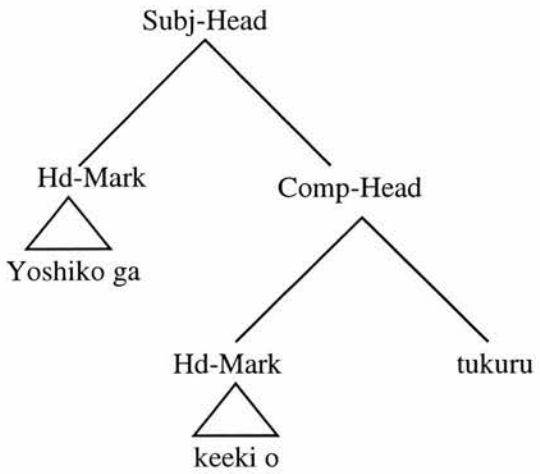
[1]Aspectual adjuncts are those which are sensitive to the features *stativity*, *telicity*, *agentivity* and *durativity*.

(6.2) 1-jikan de Yoshiko ga keeki o tukuru.
in an hour Yoshiko-NOM cake-ACC make
"Yoshiko bakes a cake in an hour".

(6.3) Yoshiko ga 1-jikan de keeki o tukuru.
Yoshiko-NOM in an hour cake-ACC make
"Yoshiko bakes a cake in an hour".

(6.4) Yoshiko ga keeki o 1-jikan de tukuru.
Yoshiko-NOM cake-ACC in an hour make
"Yoshiko bakes a cake in an hour".

(a):



(b):

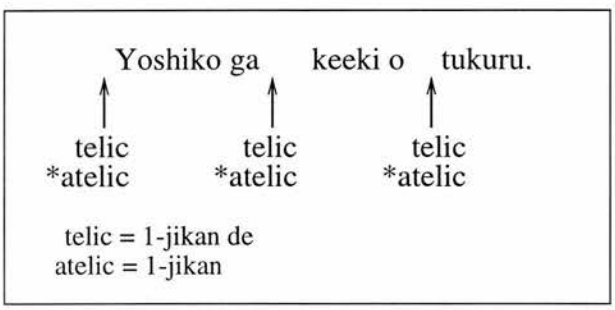


Figure 6.1: Agentive Projection of *tukuru* (a), with Adjunct Attachment (b)

Figure 6.1(b) shows possible attachment points for the Adjunct-Head Schema. The adjunct attaches through a MOD feature and selects its head, while valences, in contrast, are selected by the head. These two forms of attachment are orthogonal, so the Adjunct-Head operation has the same effect whether it attaches directly to the head, to the complement projection, or to the subject projection. This is possible because unification is an associative operation: for example adjunct, complement

and head can combine in either of the following ways, and the phrasal mothers will be identical (Note: ⊔ denotes unification.).

(6.5) (a) *adjunct(adj)* ⊔ *head(verb)* ⇒ *adjunct-head(syn:verb, sem:adj),*
(b) *comp(noun)* ⊔ *adjunct-head(verb)* ⇒ *comp-head(syn:verb,sem:adj).*

(6.6) (a) *comp(noun)* ⊔ *head(verb)* ⇒ *comp-head(syn:verb, sem:verb),*
(b) *adjunct(adj)* ⊔ *comp-head(verb)* ⇒ *adj-head(syn:verb,sem:adj).*

The aspectual data of the Nuclear Conjuncts suggests that there are distinct adjunct attachment scopes for *V-te miru* et al versus *V-te iru* et al. In its simplex form *miru* (see) is a transitive verb (7), which combines with atelic *1-jikan* (for an hour) (8), and is therefore an Activity. The accusative marked nominal alternates with the *te* form *keeki o tukutte* (baking a cake) (9). In conjunction with this Accomplishment, the atelic adverbial is inadmissible, and the adjunct form combines with telic *1-jikan de* (in an hour) (10).

(6.7) Midori ga keeki o miru.
Midori-NOM cake-ACC see
"Midori sees the cake".

(6.8) Midori ga 1-jikan ensou o miru.
Midori-NOM for an hour concert-ACC see
"Midori sees the concert for an hour".

(6.9) Keiko ga keeki o tukutte miru.
Keiko-NOM cake-ACC make-TE try
"Keiko tries baking a cake".

(6.10) Keiko ga 1 jikan de/*1 jikan keeki o tukutte miru.
Keiko-NOM in an hour/*for an hour cake-ACC make-TE try
"Keiko tries baking a cake in an hour/*for an hour".

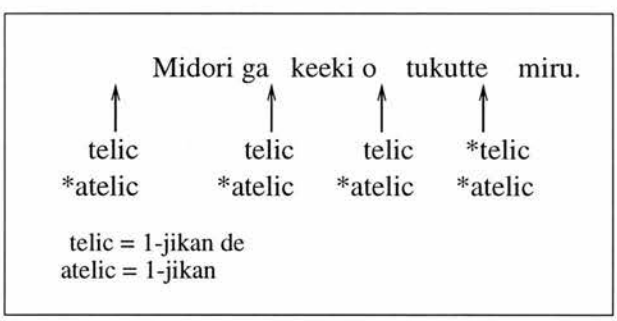


Figure 6.2: Agentive Projection of *miru* with Adjunct Attachment Points

This first conjunct adjunct scope is illustrated in Figure 6.2 with the telic adverb attaching anywhere before the first conjunct (telic) verb (*tukutte*). The atelic adverb

is inadmissible. In addition, no adjunct can intervene between the two conjunct verbs.

This is in fact in contrast with the equivalent 'try' control construction in English, example (11). The verb 'try' is an Activity compatible with atelic adverbial 'for an hour', and 'baking' is an Accomplishment, compatible with telic 'in an hour'. However examples (12) and (13) show that either one of these adverbials is acceptable, with atelic 'for an hour' combining with 'tried' in (12) and telic 'in an hour' combining with 'baking a cake' in (13).

(6.11) Delia tried baking a cake.
(6.12) Delia tried baking a cake for an hour.
(6.13) Delia tried baking a cake in an hour.

There is, then, something different about the adjunct analysis for conjunct verbs in Japanese as compared with English.

The aspectual analyses of *te iru* and *te aru* differ from those of *te miru*, *te morau* and *te ageru*, since the adjunct combines with the second conjunct verb, and not the first. The form of *iru* (animate be) shown in example (14) has a nominative UNDERGOER (*Keiko*) as subject and a dative marked LOCATIVE (*koko* (here)) as complement, and indicates a locative relation. As it combines with *shibaraku mae kara* (since a while ago) in (15), this is a State.

(6.14) Keiko ga koko ni iru.
Keiko-NOM here-LOC be
"Keiko is here".
(6.15) Keiko ga koko ni shibaraku mae kara / *moo sugu iru.
Keiko-NOM here-LOC since a while ago / *soon be
"Keiko is here since a while ago/*soon".

The adverbial *moo sugu* (soon) combines with a dynamic verb, (16) but not with a State (15).

(6.16) Keiko ga moo sugu / *shibaraku mae kara keeki o tukuru.
Keiko-NOM soon / since a while ago cake-ACC bakes
"Keiko will soon bake a cake".

The conjunction *tukutte iru* (is baking) combines with *shibaraku mae kara* but not with *moo sugu*, example (17) and is therefore Stative.

(6.17) Keiko ga shibaraku mae kara/*moo sugu keeki o tukutte iru.
Keiko-NOM since a while a go/*soon cake-ACC make-TE be
"Keiko is (has been) baking a cake since a while ago/*soon".

This pattern also applies independently of the adjunct's position in the sentence, thus in Figure 6.3, the adjunct can attach anywhere before the first conjunct, and again the conjunct internal position is blocked.

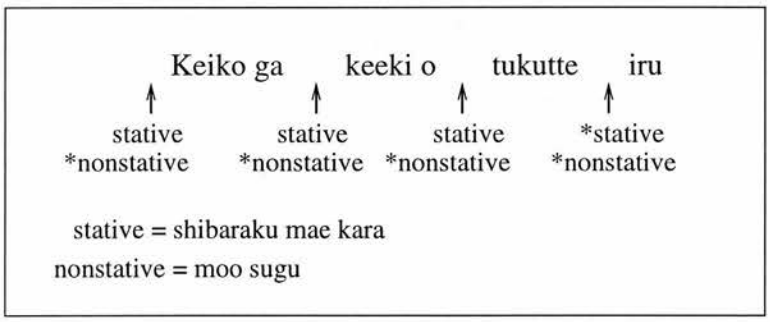


Figure 6.3: Agentive Projection of *tukutte iru* with Adjunct Attachment

The analysis of the *te miru* conjunction given in Chapter 4 shows that *miru* subcategorizes for a VP complement. The evidence from adverbials suggests that the VP complement is subcategorized together with its adjuncts. With the *te iru* conjunction, in contrast, the arguments and adjuncts project from the Stative head. The disposition of the accusative argument *keeki o* (the cake) differs between the *tukutte miru* conjunction, in (10) and *tukutte iru* in (17): the latter case can be explained if the argument is projected from the *tukutte iru* conjunction, rather than from the first conjunct (*tukutte*), as happens with *tukutte miru*. One way to address this problem is to hypothesize that the *V-te iru* (and *V-te aru*) constructions exhibit phrasal argument structure, thus, that in Figure 6.4 the conjunct *tukutte iru* combines the argument structures of the components *tukutte* and *iru*, in some way. This chapter explores a possible solution to that problem.

As a side issue to the main analysis, Section 6.1 offers a note on the analysis of adjuncts, based on the Dependency Grammar tradition (see e.g. Tesnière (1959), Melčuk (1979)), and such an analysis is incorporated into HPSG in Bouma *et al.* (1998), in which the verb has a dependent (DEPS) list in addition to argument structure (ARG-S). The DEPS list contains the append of the ARG-S and a list of adjuncts. The main phrasal argument structure analysis is based on an analogy between the *V-te iru* construction and Japanese causatives, which are lexically integral morphemes. Manning *et al.* (1999) suggest that causatives are generated by a process of derivational morphology, such as that developed by Riehemann (1995),

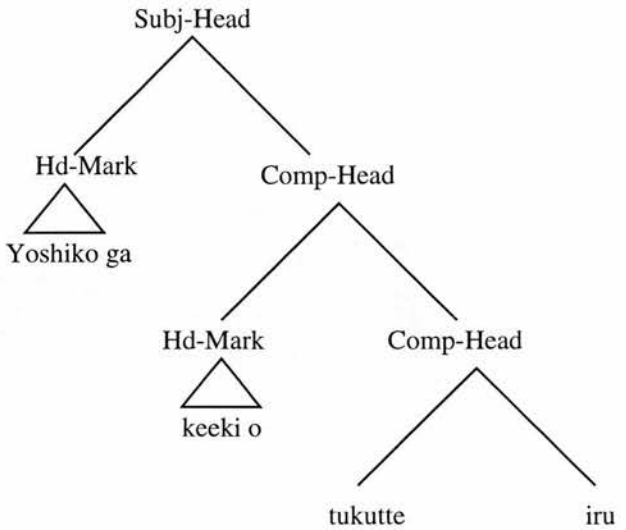


Figure 6.4: Hypothesized Analysis of *tukutte iru*

and in Section 6.3 her type based derivational morphology is reviewed, followed in 6.4 by Manning *et al.*'s (1999) analysis of Japanese causatives. Syntactic phrases are different from morphological compounds, but I propose that there is a cline of conjunction, with possible intermediate steps between these two forms. The phenomenon of phrasal argument structure falls somewhere along this cline, and I develop this mechanism in Section 6.5. The section is summarised in 6.6.

6.1 A Note on the DEPS List Analysis

The problem with analysing the adjunct attachments in example (10) (*1-jikan de ... tukutte miru*) and (17) (*shibaraku mae kara ... tukutte iru*) is that the conventional view (Pollard & Sag (1994)) is that the adjunct selects its head through a MOD feature, however there is some function of the verbal conjunction which licenses a V1 compatible adjunct or a V2 compatible adjunct. One way to explain the differences in adjunct scope is to adopt the technique of *minimal recursion semantics*, together with a syntactic analysis in which the head selects (certain adjuncts as) its dependents. This is discussed in Bouma *et al.* (1998). If all lexical items are assigned a scope value, and the adverb scopes no wider than the local semantic domain, then the relative scope values of the verbal conjuncts can determine which of the two is valid.

The general form of a lexical entry containing a DEPS list is given in Figure 6.5., and the specification of adverbials in the list is that they modify a category and content which are token identical with the containing head.

What this means for adjunct attachment is that the adjunct selects its syntactic head under the MOD feature, and provides the semantic head of the phrase, under the Semantics Principle, but also that the syntactic head licenses the category and semantic specification of the adjunct, through the DEPS list. As applied to the case of Japanese verbs and their aspectual adjuncts, such as example 6.3, the verb *tukuru* licenses a telic adverbial, on the DEPS list, and the adverbial combines under the Adjunct-Head Rule, through the MOD feature.

The DEPS list analysis might offer a possible solution to the problem of blocking adjuncts from occurring between the two verbal conjuncts in a *V-te V* construction, i.e. the fact that **tukutte 1-jikan miru* is disallowed. At present, the analysis in which the complement of *miru* or *iru* alternates between a VP and an NP does not block the intervening adjuncts. This can be solved under the DEPS list analysis by requiring a goal on the Complement-Head rule, to the effect that if the complement is a verb, then the DEPS list of the head must be unsaturated.

A more comprehensive treatment of adjunct attachment, possibly incorporating *minimal recursion semantics*, is beyond the scope of this thesis, and must await future treatment.

6.2 Lexical Argument Structure

It is a specific principle of HPSG that argument structure is projected from lexical items and not from phrases (Sag (1997)). The *te miru* conjunction is straightforwardly explained, without violating this principle. The adjunct scope of *te miru* is modelled in (18)(a) and (b). The arguments of *miru* are a nominative NP and a *te* marked VP, (18)(a). When the adjunct combines, it associates with the argument structure of the first conjunct. (18)(b). Both conjuncts retain their separate argument structures.

(6.18) (a) *miru⟨ NP(nom), VN(acc) ⟩*

$$\begin{bmatrix} \text{HEAD } \boxed{3} \\ \text{ARG-S } \boxed{1} \\ \text{DEPS } \boxed{1} \oplus \mathit{list}\left\langle \begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{HEAD } \boxed{3} \\ \text{CONT } \boxed{2} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{CONT } \boxed{2}_{mrs} \end{bmatrix}$$

Figure 6.5: The DEPS List Combining with ARG-S

(b) *VN(acc)* $\longrightarrow$ *tukutte*$\langle$ *NP(acc)* $\oplus$ *Adv(+telic)* $\rangle$

The *te iru* conjunction does not follow this pattern, as the adjunct combines with the head verb, not the subcategorized verb, and the argument structure scoping in (19) (a) and (b) reflects this difference.

(6.19) (a) *iru*$\langle$ *NP(nom), VP(te)* $\oplus$ *Adv(+state)* $\rangle$
(b) *VP(te)* $\longrightarrow$ *tukutte*$\langle$ *NP(acc)* $\rangle$

There is no phrasal structure under which the different adjunct scopes of (18)(a) and (19)(a) can be represented. Subject-Head, Complement-Head and Specifier-Head schemata 'cancel off' their valences while the argument structure is projected from the lexical head; Head-Marker structures are inapplicable because Markers, in common with other functional heads, have no argument structure; and the Adjunct-Head schema is inapplicable, again, because both conjuncts retain separate argument domains. Clearly something more than syntactic conjunction is required, so as to 'hide' the first conjunct inside the entire Stative predicate. One way to do this is to propose that argument structures of the component verbs are merged at the phrasal level. The question is though, whether such a process is available, or attested anywhere, and one possible avenue is in complex lexical structures.

There are in Japanese a variety of morphological attachments to verbs which have effects on the (combined) argument structure, and these include desiderative, potential and causative forms. The causatives in particular are relevant to this discussion, since they involve the merger of the causative morpheme's argument structure with that of the main verb. In Japanese most verbs have a causative form, and there are examples of causativized Accomplishments, Activities, Achievements and some States. All agentive verbs, whether transitive or intransitive causativize readily, thus *naosu* (repair) —⟩ *naosaseru* (let repair), *piano o hiku* (play the piano) —⟩ *piano o hikaseru* (let play the piano), *kaeru* (go home) —⟩ *kaeraseru* (let go home), and *tomaru* (stop) —⟩ *tomaraseru* (make stop). Verbs which do not causativize include some punctual Achievements *kawaku* (get dry) —⟩ **kawakaseru*, *hajimaru* (begin) —⟩ **hajimaraseru*, and some States: *aru* (be), *iru* (need) and *dekiru* (be able).

The causative morpheme is a ditransitive argument frame and the main verb is embedded, with the dative argument of the causative identified with the embedded subject argument. The verb *naosu* (correct) is a transitive Accomplishment (20)

susceptible to causativization (21), with the resultant Valence-Increase and conversion of embedded subject to dative; *hashiru* (run) in (22) is an intransitive Activity, and the embedded subject of the causative form, in (23), may be rendered as dative or accusative.

(6.20) Sensei ga ronbun o naoshimashita.
teacher-NOM report-ACC fix-POL-PAST
"The teacher corrected the report".

(6.21) Yumiko ga Sensei ni ronbun o naosasemashita.
Yumiko-NOM teacher-DAT report-ACC fix-CAUS-POL-PAST
"Yumiko let the teacher correct her report".

(6.22) Masaru ga hashitta.
Masaru-NOM run-PAST
"Masaru ran".

(6.23) Noriko ga Masaru ni/o hashiraseta.
Noriko-NOM Masaru-DAT/ACC run-CAUS-PAST
"Noriko let/made Masaru run".

A causative form with a dative argument is generally associated with a permissive reading, while the accusative is associated with a coercive reading. Various authors, including Miyagawa (1989), Ishikawa (1985), Matsumoto (1996) and Manning *et al.* (1999) argue that causatives are lexical, rather than syntactic phenomena. Manning *et al.* (1999) argue that ambiguity of adjunct scope is diagnostic of lexicality, so in example (24) the adverb *gakkoo de* (at school) is ambiguous between the 'causing' and the 'running' event.

(6.24) Noriko ga Masaru ni gakkoo de hashiraseta.
Noriko-NOM Masaru-DAT school-ADV run-CAUS-PAST
"Noriko made Masaru run at school".

In a separate test, Poser (1983) argues that the *o* accusative marker may not occur more than once in an argument domain, the ability of complex verbs to support multiple *o* markers is a consequence of multiple argument structure domains, and therefore the inability of causatives to support double *o* marking is indicative of their being a single lexical item. This is demonstrated by extending example (22) to add a traversal object (25). With this extension only the *ni* argument conversion is possible, and the *o* conversion is bad, as in (26).

(6.25) Masaru ga rooka o hashitta.
Masaru-NOM hallway-ACC run-PAST
"Masaru ran the hallway".

(6.26) Noriko ga Masaru ni/*o rooka o hashiraseta.
Noriko-NOM Masaru-DAT/*ACC hall-ACC run-CAUS-PAST
"Noriko let(/?made) Masaru run the hallway".

A further argument for lexicality is made if the conjuncts cannot be separated by intervening word forms. The *te miru* (27) and *te iru* (28) forms may take intervening focus particles such as *wa*, *mo* and *sae*, but these may not appear between the main verb and causative morpheme (29).

(6.27) Noriko ga ronbun o naoshite wa mita.
Noriko-NOM report-ACC fix-TE TOP see to-PAST
"What Noriko saw to was correcting the report".

(6.28) Noriko ga ronbun o naoshite wa ita.
Noriko-NOM report-ACC fix-TE TOP be-PAST
"What Noriko did was to correct the report".

(6.29) *Noriko ga ronbun o naoshi wa saseta.
Noriko-NOM report-ACC fix-INF TOP 'do'-CAUS-PAST

For all these reasons Manning *et al.* (1999) argue that causatives are discrete lexical forms. The *te miru* and *te morau* constructions do not exhibit the same behaviour, and I have already established in Chapter 4 that they are syntactic phenomena. There are also *te* conjunctions which do not pattern consistently with *te miru/morau*, without being pure lexical forms like the causatives. The *te hoshii* form reviewed in Chapter 2 is related to *te morau* and is clearly a syntactic construction, and yet is ambiguous between first and second conjunct adjunct scopes, a phenomenon which is otherwise diagnostic of a lexical form. The *te iru* and *te aru* forms, as we have seen above, take second conjunct adjunct scope, while otherwise patterning as syntactic phenomena. These intermediate results in fact suggest that the boundary between syntax and morphology is not so hard as is often suggested in the literature, and there may in fact be a cline of conjunction, following Hopper & Traugott (1993), as illustrated in Figure 6.6. In their analysis of Bantu noun classes, Bresnan & Mchombo (1995) argue for the complete separation of syntactic and morphological processes, however the presence of linguistic phenomena which exhibit features of both 'levels' suggests in turn that there may be some overlap between these processes.

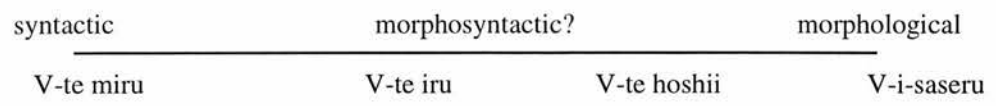


Figure 6.6: A Cline of Conjunction

In a recent HPSG analysis of Japanese causatives, Manning *et al.* (1999) adopt Riehemann's (1995) derivational morphology as a way of deriving the causative form (*naosaseru* (let repair)) from the dictionary form (*naosu* repair)), and this mechanism may provide some insight into a morphosyntactic method for analysing *te iru* and related forms.

6.3 Type Based Derivational Morphology

One theme of research into lexicalist analyses is that there are syntactic and semantic regularities in the lexicon which can be exploited in creating hierarchical structure and thereby reduce lexical redundancy, and this is demonstrated in work done by Flickinger *et al.* (1985), Flickinger (1987) and Davis (1995). There are also regularities in morphology which contribute to further economies in lexicon structure and these include rules for adjective to verb formation by *-en* suffix, thus 'red' $\longrightarrow$ 'redden', 'white' $\longrightarrow$ 'whiten'; verb to adjective formation by *-able* suffixation: 'derive' $\longrightarrow$ 'derivable', 'achieve' $\longrightarrow$ 'achievable'; verbs to nouns: 'bake' $\longrightarrow$ 'baker', 'climb' $\longrightarrow$ 'climber'; and many others, typically involving prefixation or suffixation. The mechanism for affixation in standard HPSG (Pollard & Sag (1987), Pollard & Sag (1994)) is that of the lexical rule. The passive lexical rule for *bake*, exemplified by (30) and illustrated in Figure 6.7, converts the transitive Accomplishment *bake* to its passive form *baked*, promotes the UNDERGOER argument to subject, and demotes the ACTOR to an optional 'by' headed prepositional phrase.

(6.30) Mary bakes the cake. $\longrightarrow$ The cake is baked (by Mary).

Lexical rules are unconstrained however, and in their most general form can replace anything with anything, thus yielding results which are descriptive but not explanatory, since the detailed lexical changes need not be related to any principled system, and for these reasons some researchers deprecate their use. Riehemann (1995) posits an approach using derivational morphology which "is formalized in terms of complex recursive schemata structured in a multiple inheritance hierarchy, without positing lexical rules or lexical entries for affixes". Her analysis uses data

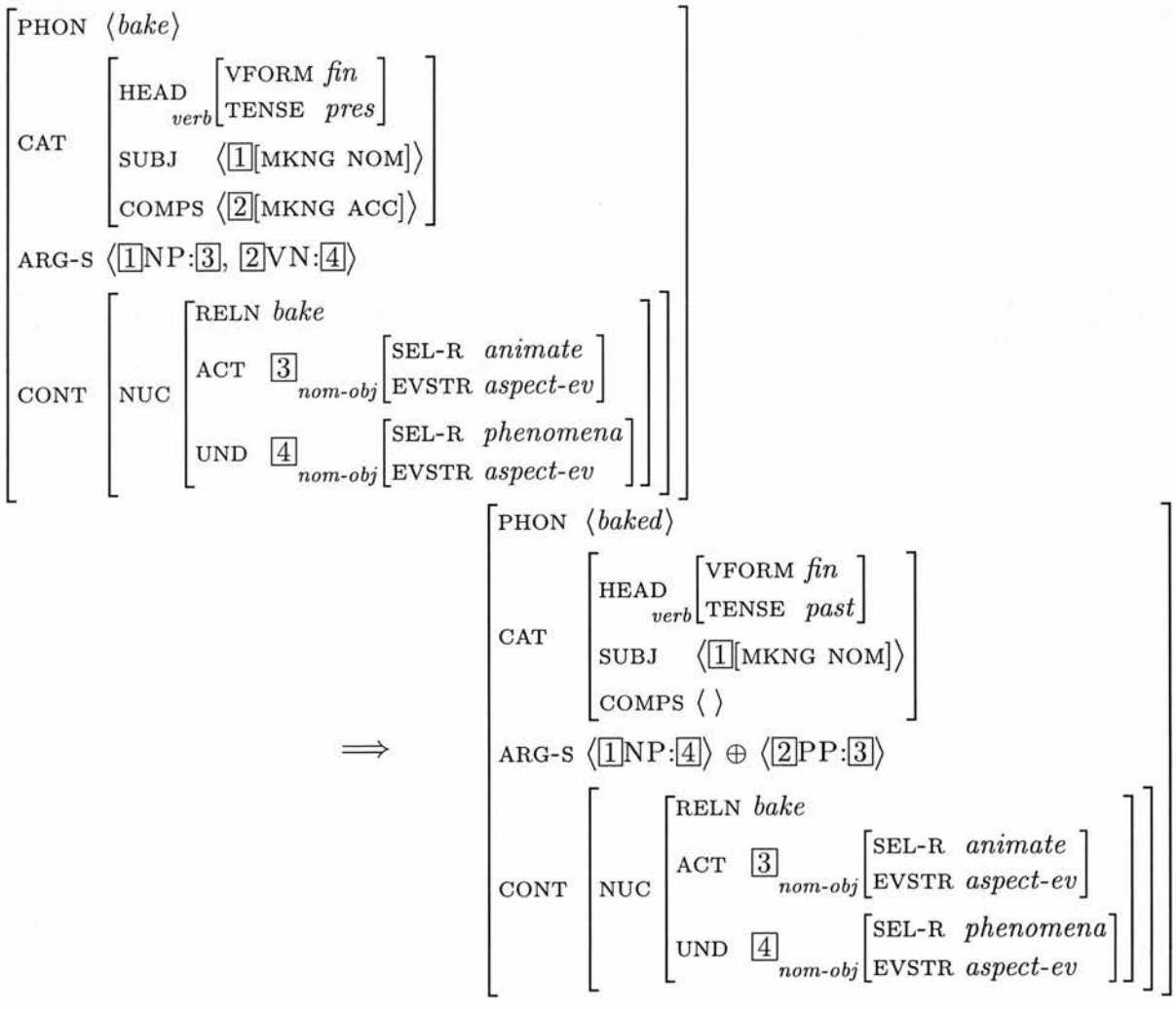


Figure 6.7: The Passive Lexical Rule

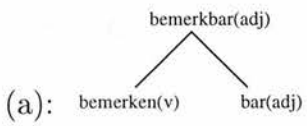
from the German *bar* adjectives, equivalent to the English '-able' adjectives, as shown in Example (31), in which the verb *bemerken* (notice) is transformed into the adjective *bemerkbar* (noticeable).

(6.31) (a) Sie bemerken die Veränderung.
They notice the change.
(b) Die Veränderung ist bemerkbar.
The change is noticeable.

The morphological operation of *bar* adjective formation works by suffixation of *bar* to the morphological root of a transitive verb, suppressing the actor and projecting the undergoer as subject[2].

[2]Following the analysis in Chapter 4, this operation can be formalized using the adaptation of Pustejovsky's (1995) system. An aspectual analysis of the German *-bar* adjectives will determine whether they are formed from agentive verbs, or more restrictively, Accomplishments. The verb

Previous approaches to derivational morphology include word syntax, as proposed by Selkirk (1982), and the use of lexical rules by e.g. Aronoff (1976). The first approach, illustrated in Figure 6.8(a), posits word internal phrase structure and Riehemann's objection to this approach is that the *bar* suffix, and morphological suffixes in general, have no independent lexical existence, yet syntactic rules require a lexicon to draw from. Evidence from grammaticalization, however, suggests that there may be a cline of lexical phenomena in which the same morpheme is both an affix and an independent word. An example from Japanese is the *mono* nominalization: while *mono* is a simplex noun meaning 'thing', it also attaches as an affix to a verbal infinitive, thus *nomu* (drink) derives *nomimono* (beverage), and *taberu* (eat) derives *tabemono* (foodstuff). Riehemann's denial of the effect of grammaticalization on independent lexical material needs more support then, but it does not invalidate the mechanism of productive word formation which she develops. In any case, a morphological phrase structural rule is clearly distinguished from a syntactic phrase structure rule, because the former yields a single lexical item with an argument structure, while the latter has no independent argument structure, according to Sag (1997). In the second approach discussed by Riehemann, shown as Figure 6.8(b), a lexical rule is used to derive the adjective *bemerkbar* from the transitive verb *bemerken*, Her objection to the lexical rule approach is not that the *bar* suffix has no independent lexical existence, but that under this operation, subregularities and exceptions cannot be cleanly explained since lexical rules allow no hierarchical variations, and multiple related lexical changes entail multiple lexical rules.



(b): $[[\text{bemerk}]_V \text{bar}]_A$

Figure 6.8: (a) Word Syntax versus (b) Lexical Rule

In Riehemann's proposal, there is a general schema[3] for the *bar* adjectives which inherits from the schema for transitive verbs. This schema involves a subject and an accusative object and is linked with Davis (1995) content relation of *act-und-rel.* The phonology is that of the verb's morphological root (i.e. *les* from *lesen*, *bemerk* from *bemerken*). Although she is at pains to point out that the *bar* schema does not warrant an independent lexical entry, it does contain its own phonology (*bar*), categorial head (adjective) and valences (subject and complements) and its content

$\longrightarrow$ adjective conversion in these cases seems to involve the relation between Agentive and Formal Qualia. This would be prefigured in the Qualia Structure of an Accomplishment.

[3]Riehemann's use of the word *schema* relates to a cluster of features in the type hierarchy, and should not be confused with Phrase Structure Schemata.

value, provided through structure sharing the content value of the verb. The lexical entry for the entire *bar* adjective, following composition with its subcategorized verb, is illustrated in the feature structure of Figure 6.9. This figure illustrates the relationship between the morphological stem identified by the feature MORPH-B, and the complex word. In the nuclear content of the verb, indexed [4], the ACTOR is suppressed, and the UNDERGOER coindexed as subject of the adjective. The comps argument, indexed [3], only appears in the case of a ditransitive verb conversion, and represents a possible LOCATIVE argument (not explained in Riehemann's original). The content of the adjective contains the subject as ARG1 and the entire content of the verb's NUCLEUS as ARG2. The phonology is that of the verb stem with *bar* appended.

$$
\left[\begin{array}{ll}
\text{PHON} & \langle \boxed{1} + bar \rangle \\
\text{MORPH-B} & \left[\begin{array}{ll}
\text{PHON} & \langle \boxed{1} \rangle \\
\text{SYN|LOC} & \left[\begin{array}{ll}
\text{CAT|VAL|COMPS} & \langle \text{NP(ACC)}\boxed{2} \rangle \oplus \boxed{3} \\
\text{CONT|NUC} & \boxed{4}\left[\begin{array}{l}\text{ACT} \ . \\ \text{UND} \ \boxed{2}\end{array}\right]
\end{array}\right]
\end{array}\right] \\
\textit{trans-verb} & \\
\text{SYN|LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \text{ADJ} \\
\text{VALENCE} & \left[\begin{array}{ll}\text{SUBJ} & \langle \text{NP}\boxed{2} \rangle \\ \text{COMPS} & \boxed{3}\end{array}\right]
\end{array}\right] \\
\text{CONT|NUC} & \left[\begin{array}{l}\text{RELN} \ \diamond \\ \text{ARG1} \ \boxed{2} \\ \text{ARG2} \ \boxed{4}\end{array}\right]
\end{array}\right] \\
\textit{reg-bar-adj} &
\end{array}\right]
$$

Figure 6.9: The Productive Schema for *bar* Adjectives

The analysis focuses on the explanation of syntactic and morphological phenomena, and she assumes Davis (1995) HPSG semantics in which the NUCLEUS of the transitive verb stem comprises unsaturated ACTOR and saturated UNDERGOER arguments. The mechanism is not extended to include an account of the argument structure operations, however in their lexical account of the Japanese causatives, Manning *et al.* (1999) employ Type Based Derivational Morphology to explain the productive formation of causatives from verbs of varying transitivity and aspect classes, and their derivation uses a richer lexical structure, including articulating the argument structures of the verb stem and the derived causative. Their analysis is particularly useful to the discussion of Nuclear Conjunctive forms.

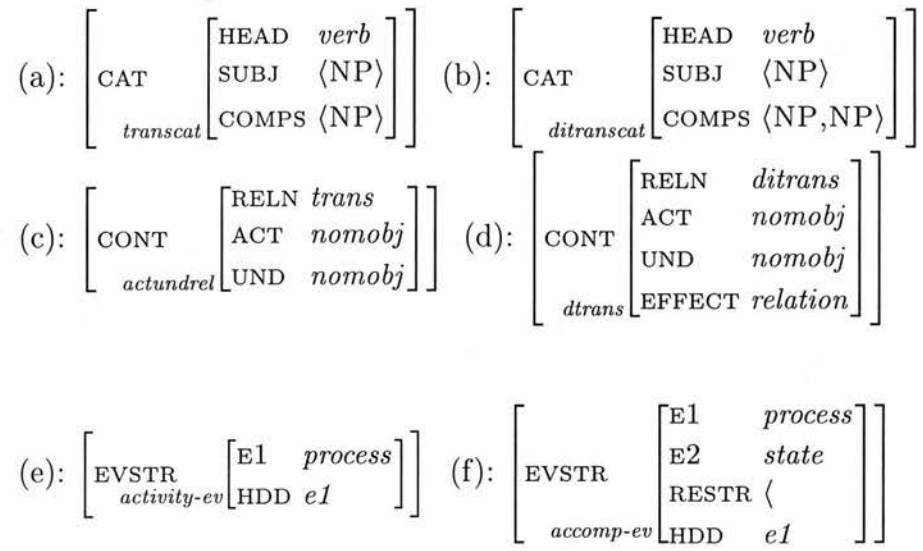


Figure 6.10: Lexical Constraints for Transitives and Ditransitives

6.4 Application of TBDM to Japanese Causatives

Both the causative morpheme *(s)aseru* and the lexical form to which it attaches are verb forms, although the causative, like the *-bar* affix in German, has no independent lexical existence[4]. The multiple inheritance approach to lexical organization adopted in HPSG identifies regularities in lexical forms such as the transitive Accomplishment verbs, and encodes them in type hierarchies containing bundles of feature descriptions. Thus, while 'give', 'receive', 'ride' and 'bake' are all verbs denoting some sort of relation between their arguments, and susceptible to common morphological processes such as *-ing* formation as in 'giving', 'receiving', 'riding' and 'baking', they differ in that 'give' and 'receive' are ditransitives: "A gives B to C", while 'ride' and 'bake' are transitive: "A bakes B". The transitives also differ in that 'ride' is an Activity with no result state, but 'bake' is an Accomplishment. The syntactic differences are captured by separating transitive and ditransitive verbs in an inheritance hierarchy, where all transitives have a complements list of length 1 (Figure 6.10(a)), while ditransitives have a complements list of length 2 (6.10(b)). The semantic differences are captured under the NUCLEUS, with ACT and UND (6.10(c)) or ACT, UND and EFFECT (6.10(d)) semantic roles, following Davis (1995), and in the event structure hierarchy, distinguishing an Activity such as 'ride', with a single process subevent (6.10(e)), from an Accomplishment such as 'bake', with process and state subevents (6.10(f)).

[4]Although the causative morpheme is clearly related to the independent lexical form of the causative of *suru* (do), which is *saseru* (let do).

Because 'bake' is defined as a transitive Accomplishment, it inherits the general constraints for those classes, and these are added to the unique characteristics including phonology, semantic relation and real world constraints supplied by CONTEXT information. The causative morpheme has no independent lexical entry, but it is articulated as a general schema in the type hierarchy of verbs. This schema, illustrated in Figure 6.11, contains phonology, which is some function of *(s)ase*, the causative morpheme; argument structure, in which the first and second arguments are causer and causee, while the third argument is the embedded argument structure of the contentive verb; the *content* is a ternary *cause-rel* relation with ACTOR, UNDERGOER and EFFECT arguments; and the stem is a template for the phonology and content of the matrix verb.

$$\begin{bmatrix} \text{PHON} & \langle Fsase(\boxed{1}) \rangle \\ \text{ARG-ST} & \langle \text{NP}:\boxed{4}, \text{NP}:\boxed{5}, \ldots \rangle \\ \text{CONT } \boxed{3} & \begin{bmatrix} \text{ACT} & \boxed{4} \\ \text{UND} & \boxed{5} \\ \textit{causerel} \; \text{EFFECT} & \boxed{6} \end{bmatrix} \\ \text{STEM} & \begin{bmatrix} \text{PHON } \boxed{1} \\ \textit{v-stem} \; \text{CONT } \boxed{3} \end{bmatrix} \end{bmatrix}$$

Figure 6.11: Underspecified Causative Morpheme

Since the causative morpheme and the verb stem are both of type verb, they both inherit features from the verb type hierarchy, roughly sketched in Figure 6.12.

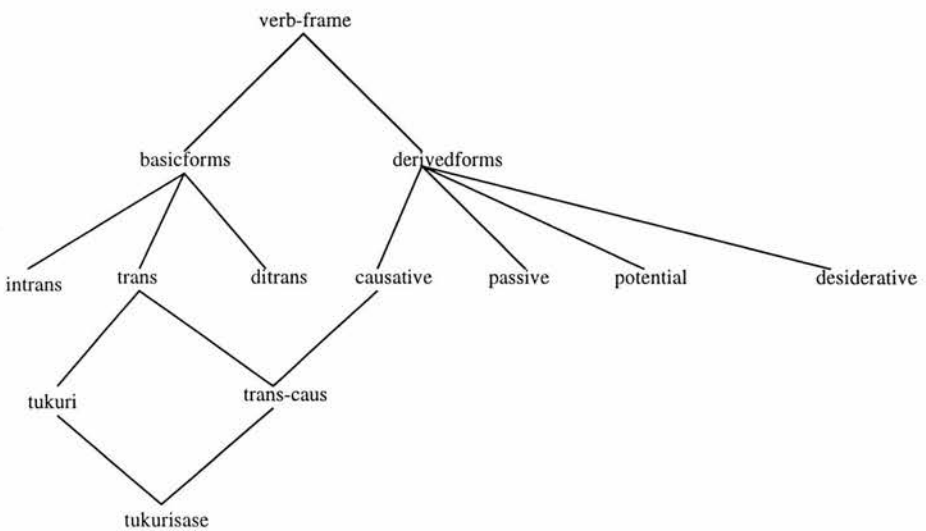


Figure 6.12: A Verb Type Hierarchy

The 'bake' stem inherits features from the transitive verb frame and gives rise to a number of lexicalized forms, including finite and non-finite inflected forms. The causative form inherits from *tukuri* and the causative feature complex to yield the lexicalized *tukuraseru* (let bake).

$$\left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \textit{verb}\\ \textsc{subj} & \langle \boxed{1}\textsc{np}\rangle \\ \textsc{comps} & \langle \boxed{2}\textsc{np},\boxed{3}\textsc{np}\rangle\end{array}\right] \\
\textsc{arg-st} & \left\langle \boxed{1}\textsc{np}{:}\boxed{4},\boxed{2}\textsc{np}{:}\boxed{5},\langle \textsc{pro}{:}\boxed{5},\boxed{3}\textsc{np}{:}\boxed{6}\rangle\right\rangle \\
\textsc{cont} & \left[\begin{array}{ll}\textsc{act} & \boxed{4}\\ \textsc{und} & \boxed{5}\\ \textsc{effect} & \left[\begin{array}{ll}\textsc{act} & \boxed{5}\\ \textsc{und} & \boxed{6}\end{array}\right]\end{array}\right]
\end{array}\right]$$

cause-stem; *causerel*; *buy-rel*

Figure 6.13: Lexical Entry for a Morphologically Derived Causative

Figure 6.13 shows the lexicalized causative, with content composed from the *causerel* content and the matrix verb. As part of this morphological operation, the argument structure is derived from the merger of the two component argument structures, as shown in (32):

(6.32) $\langle NP_4, NP_5, ...\rangle \oplus \langle PRO_5, NP_6\rangle \Rightarrow \langle NP_4, NP_5, \langle PRO_5, NP_6\rangle\rangle$

The resultant lexical entry bears a clear relationship to the object control form of type *influence* discussed in the previous section, and indeed, causative forms in English, 'let', 'make' and 'cause', which are syntactic constructions, analyse directly as control constructions (Pollard & Sag (1994)). Also in line with their analysis of English control constructions, the causee object, $\boxed{2}$, is structure-shared with the unexpressed ACTOR of the matrix verb, which is represented as a pronominal element PRO, indexed $\boxed{5}$, in the argument structure. I have shown in the previous section that the argument that *control* constructions involve an overt link to the unexpressed subject is not proven, following Chierchia (1984), so discounting PRO there is clearly a relationship between the semantically motivated control construction, and the lexically driven causativization (indeed, Manning *et al.*'s (1999) analysis explicitly identifies such a relationship). If, as I suggested earlier, (see Figure 6.6) there is a process of grammaticalization which converts syntactic constructs to complex morphological constructs, it is reasonable that there be some overlap between syntactic processes and morpholexical processes. Just such a process might

provide an explanation for similarities and differences between *te miru* and *te iru* on the one hand, and between *te iru* and morphological causatives on the other hand.

6.5 Phrasal Argument Structure

The structures examined so far include the Complement-Head schema, which is a method of syntactic conjunction, and derivational morphology, which is a method of complex word formation. The HPSG feature geometries of the resultant structures are very different: the phrasal sign retains the identity of the component daughters (HD-DTR and NON-HD-DTR), and generates a separate structure for the phrasal mother, a SYNSEM structure; the lexical structure inherits features from a type hierarchy as constraints on a single lexical entry, bearing a fully formed argument structure. There is evidently a difference between the SYNSEM structure projected as the phrasal mother and that generated morphologically, in that the former has no argument structure, and this differentiation is reflected in the *synsem* type hierarchy, illustrated in Figure 6.14.

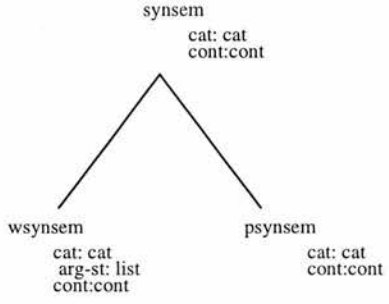


Figure 6.14: The *synsem* Hierarchy

All *synsem* structures have CAT and CONT features, bearing syntactic category and semantic content information respectively, and there are subtypes for *word* and *phrasal synsem* objects. Phrasal synsem (*psynsem*) introduces no new features, while word synsem (*wsynsem*) introduces argument structure (ARG-S), and the phrasal mother generated by application of a phrase structure rule bears the type *psynsem*, while a contentive lexical structure is of type *wsynsem*. This view of the feature geometry of HPSG follows directly from Sag's (1997) explanation.

Neither the Complement-Head analysis nor the morphological derivation are quite appropriate for the *te iru* related structures. While *te iru* patterns syntactically in that it is formed from separate lexical entries, apparently under the Complement-Head rule, it differs from pure syntactic conjunction in that the phrasal mother

combines with adjuncts in a manner which is consistent with the hypothesis that it has argument structure, and can therefore be represented with type *wsynsem*. Figures 6.15 and 6.16 illustrate the difference between the phrasal structures resulting from pure syntactic conjunction, and the kind of conjunction resulting in phrasal level argument structure. These are exemplified by *te miru* in example (33) and *te iru* in (34) respectively.

(6.33) Sugiyama-san ga 1-kagetsu ie o tatete mita.
Mr Sugiyama-NOM for 1 month house-ACC build-TE see about-PAST
"Mr Sugiyama saw about building a house in 1 month".

(6.34) Sugiyama-san ga shibaraku mae kara ie o tatete ita.
Mr Sugiyama-NOM since a while ago house-ACC build-TE be-PAST
"Mr Sugiyama has been building a house since a while ago".

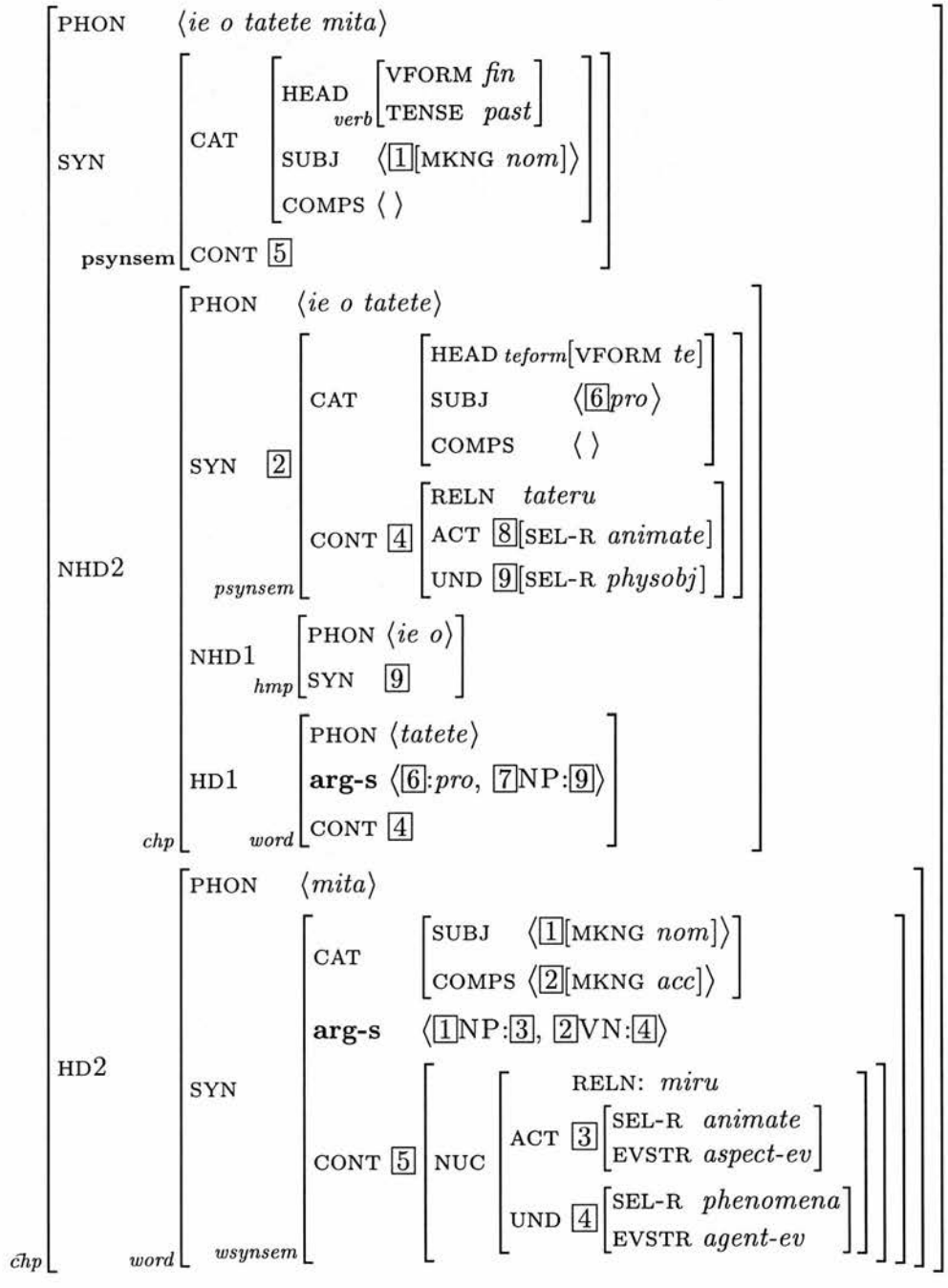


Figure 6.15: Phrasal Structure for *ie o tatete mita*

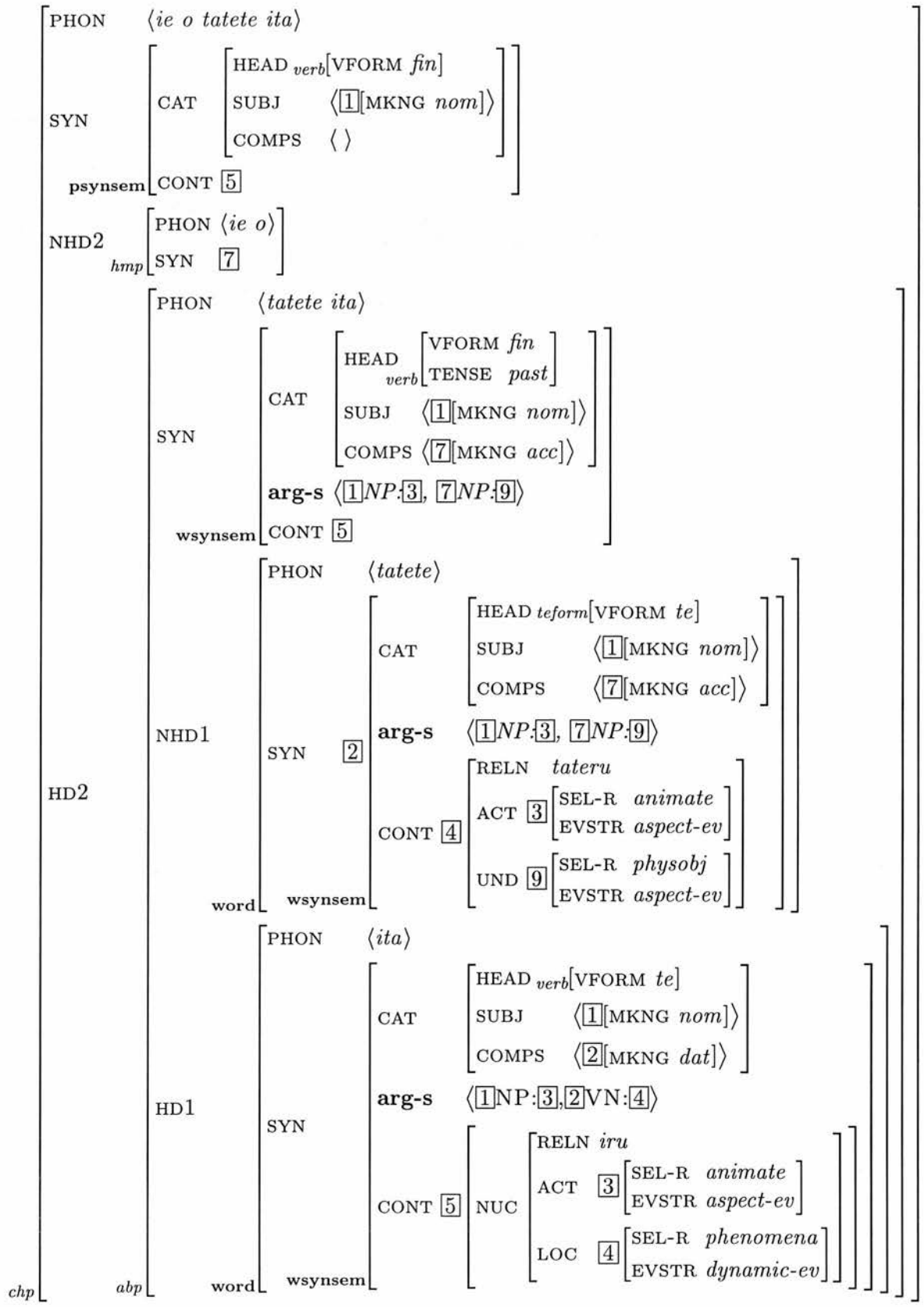


Figure 6.16: Phrasal Structure for *ie o tatete ita*

In Figure 6.15 *mita* is represented as the head daughter of a Complement-Head phrase, and it is a lexical sign (*word*), with argument structure $\langle NP, VN \rangle$. The embedded phrase *ie o tatete* is a VP representing the NON-HEAD-DAUGHTER, subcategorized as the complement of *mita*. The MARKING value of *tecase* is a supertype of *accusative*. The phrasal mother *ie o tatete mita* is a *synsem* of type *psynsem*, carrying no argument structure. Any adjunct is an adjunct of V1, *tatete*. The Complement-Head schema which generates this structure is paraphrased from Borsley (1996) as:

> **Complement Head Schema**
> A sign with the feature specification comps(saturated) can contain a head daughter with the feature specification comps(unsaturated) and any number of complement daughters.
>
> Borsley(1996):p83

The above schema does not take account of the constraint on phrasal argument structure which Sag (1997) insists on. A modified schema is therefore as follows:

1. **Phrasal Sign**
   A phrasal sign contains phon and synsem features in which synsem contains no argument structure (ARG-S) feature.
2. **Complement Head Schema**
   A phrasal sign with saturated comps can contain a head daughter and and any number of complement daughters. The head daughter is a word with unsaturated comps linked with the oblique elements of its ARG-S feature.

The *ie o tatete ita* construction shown in Figure 6.16 contrasts with the above structure in that the phrasal mother, *tatete ita* is a structure which has a phrasal ARG-S feature. The structure is characterized by a SYNSEM feature of type *wsynsem*. This phrasal mother can be interpreted as being generated by a variant of the Complement-Head rule, called the *Aspectually-Bound Rule*, which passes up some function of the ARG-S lists of the non-finite and the finite verbal daughters. A modified schema which accounts for this is:

1. **Phrasal Sign**
   A *sign* which is a strict subtype of phrasal contains PHON and SYNSEM features, with the SYNSEM articulated into CAT and CONT features. A *sign* which is a subtype of both *phrasal* and *word* types inherits the ARG-S feature from its *word* parent, in addition to the above features inherited from both parents.

2. **Complement Head Schema**
   A phrasal sign can contain a head daughter with COMPS projected from the oblique element(s) of ARG-S and any number of complement daughters.
3. **Admissibility Conditions**
   An Aspectually Binding verb such as *iru*, *aru* or *shimau* subcategorizes for a verb phrase with unsaturated COMPS, and in this case the phrasal mother projects argument structure which is the merge of the component argument structures. Any other verb which takes a verbal argument, takes a verb with saturated COMPS.
4. **Argument Structure Merge Principle**: "The ARG-ST values of NON-HD-DTR and HD-DTR merge to generate a Mother with SYNSEM value *wsynsem*, containing an ARG-ST feature which is a complex list in which the oblique element of the head daughter's ARG-ST subsumes the ARG-ST list of the non head daughter".

The *Argument Structure Merge Principle* is illustrated for *tatete iru* in (35), with the argument structure of *tatete* derived from that of *tateru* in (a), with the ACTOR argument suppressed following suffixation of *te*; and this ARG-S list taking the position of the LOCATIVE element of *iru* in (b).

(6.35) (a) $\langle NP_5, NP_6 \rangle_{tateru} \Rightarrow \langle NP_6 \rangle_{tatete}$
(b) $\langle NP_4, VN_5 \rangle_{iru} \oplus \langle NP_6 \rangle_{tatete} \Rightarrow \langle NP_4, \langle NP_6 \rangle \rangle_{tateteiru}$

There now exist two variants of the Complement-Head rule: the phrasal mother in the original rule does not carry ARG-S, while in the case of the *Aspectually Bound Phrase*, the phrasal mother does carry ARG-S. In the HPSG *sign* type hierarchy, he differences between a sign of type *word* and a sign of type *phrasal* are that the *word* sign contains ARG-S and the *phrasal* sign contains daughters. The Aspectually Bound phrase inherits from both *word* and *phrasal* signs. Sag's (1997) hierarchy of signs given as Figure 3.25, differentiates between words as one subtype, and the hierarchy of headed phrases on the other. Without further evidence, no other phrasal type has argument structure. This relationship between words and Aspectually Bound phrases is depicted in Figure 6.17, updating Sag's earlier figure.

How, though, is the Aspectually-Bound phrase realized in preference to the Complement-Head phrase? There is a particular constraint proposed in Sag (1997) which determines the saturation level of the COMPS list: this is the Empty Comps Constraint (ECC). The ECC is a precedence rule on valences which ensures that COMPS are

Figure 6.17: The Extended *sign* Hierarchy

saturated before any other valence. In a transitive verb with a subject and a complement, COMPS is saturated 'first' and then the subject[5]. In a situation where Complement-Head constructions are cascaded, however, the ECC imposes a 'precedence of saturation' which depends on the language specific Linear Precedence ordering. In English, which orders Heads before their Complements, the rightmost complement phrase is saturated first, whereas Japanese is ordered with Complements preceding their heads, and the leftmost Comp-Head phrases are saturated first. The effect of the Aspectually Bound phrase is to defeat this precedence ordering, leading to the distinction where cascaded Comp-Head phrases are left branching in the normal case of phrasal conjunction, and right branching under Aspectual Binding (as in Figure 6.4). The effect of the ECC is to ensure that the output of the Complement-Head phrase, and the input to all other phrasal types, is an empty complements list. According to Sag (1997), the ECC is a principle which

[5]In Japanese there is a phenomenon of unrestricted argument drop: the first conjunct may have complements which are dropped, but these are not then subsequently picked up by the second conjunct. A complement once dropped, stays dropped, and this is necessarily governed by a separate principle from the ECC. Complement drop does not then affect the problem of ARG-S merger.

applies to all headed phrases, but as we can see it does not apply in the cases of Aspectually-Bound phrases.

Simplex verbs select for a nominal complement, which always has an empty comps list. Conjunct verbs of the type of *miru* (see) select for a saturated Complement-Head phrase, which is a verb with complements already consumed. Verbs of the type of *iru* select for a verb with unsaturated comps, and the Aspectually-Bound phrase which governs the combination ensures that the comps of the subcategorized verb are combined onto the phrasal ARG-S. In summary: the lexical entry for *iru* has a COMPS list which specifies an unsaturated complement; the Aspectually-Bound phrase combines its daughters in a Complement-Head relationship and in addition combines the argument structures of the daughters onto a phrasal argument structure. A lexical entry for *iru* which accommodates this phrasal behaviour is given in Figure 6.18.

$$\left[\begin{array}{ll}
\text{PHON} & \langle iru \rangle \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & verb[\text{VFORM } fin] \\
\text{SUBJ} & \langle \boxed{1}[\text{MKNG } nom] \rangle \\
\text{COMPS} & \langle \boxed{2}[\text{MKNG } dat] \rangle_{ne-list} \\
\text{MARKING} & unmarked
\end{array}\right] \\
\text{ARG-S} & \langle \boxed{1}\text{NP:}\boxed{3}, \boxed{2}\text{NP:}\boxed{4} \rangle \\
\text{CONT} & \left[\begin{array}{ll}
\text{SEL-R} & events \\
\text{NUC} & \left[\begin{array}{ll}
\text{RELN} & aru \\
\text{UND} & \boxed{3}\left[\begin{array}{ll}\text{SEL-R} & animate \\ \text{EVSTR} & aspect\end{array}\right] \\
\text{LOC} & \boxed{4}\left[\begin{array}{ll}\text{SEL-R} & phenomena \\ \text{EVSTR} & dynamic\text{-}ev[\text{HDD } e1]\end{array}\right]
\end{array}\right] \\
\text{EVSTR} & state\text{-}ev[\text{E1 } state] \\
\text{QUALIA} & formal\left[\begin{array}{ll}\text{EVTYPE} & e1 \\ \text{ALIST} & \langle \boxed{3}[\text{PROJS } project], \boxed{4}[\text{PROJS } project] \rangle\end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]$$

Figure 6.18: *iru* Selecting for Unsaturated Comps

The complements linked to locative role are underspecified to accept nominal or verbal arguments, each of which has unsaturated complements. When a verb subcategorized by *iru* combines under Aspectually-Bound phrase, the phrasal ARG-S contains the subject of *iru* and any complements of the embedded verb. If an unergative verb such as *odoru* (dance) is subcategorized in the form *odotte iru*, the

ARG-S of the phrasal mother contains the subject NP only, as there is no complement to be combined. If a transitive verb such as *shimeru* (close) combines with *te iru*, the single accusative marked complement combines onto the ARG-S of the phrasal mother; if a ditransitive verb such as *oshieru* (teach) combines, both the accusative and dative arguments are passed to the mother. The complements and ARG-S lists of the phrasal mothers in these three cases are given in (36)(a) to (36)(c).

(6.36) (a) odotte iru:

odotte(comps⟨⟩) $\oplus$ *iru(comps⟨***VN(te)***⟩)*

$\Longrightarrow$ *odotte iru(comps⟨⟩)*

odotte(arg-s⟨⟩) $\oplus$ *iru(arg-s⟨***NP, VN(te)***⟩)*

$\Longrightarrow$ *odotte iru(arg-s⟨***NP***⟩)*

(b) shimete iru:

*shimete(comps⟨***NP***⟩)* $\oplus$ *iru(comps⟨***VN(te)***⟩)*

$\Longrightarrow$ *shimete iru(comps⟨***NP***⟩)*

*shimete(arg-s⟨***NP***⟩)* $\oplus$ *iru(arg-s⟨***NP, VN(te)***⟩)*

$\Longrightarrow$ *shimete iru(arg-s⟨***NP, NP***⟩)*

(c) oshiete iru:

*oshiete(comps⟨***NP, NP***⟩)* $\oplus$ *iru(comps⟨***VN(te)***⟩)*

$\Longrightarrow$ *oshiete iru(comps⟨***NP, NP***⟩)*

*oshiete(arg-s⟨***NP, NP***⟩)* $\oplus$ *iru(arg-s⟨***NP, VN(te)***⟩)*

$\Longrightarrow$ *oshiete iru(arg-s⟨***NP, NP, NP***⟩)*

In the general case a noun has no argument structure, and empty complements[6]. In the polysemous case, a locative nominal also combines with the Aspectually-Binding phrase and so its COMPS are left unsaturated, and its ARG-S merged on the phrasal mother. A noun with no complements acts like an unergative verb, cf *odotte*, 36(a), and its empty COMPS list is merged, leaving an empty phrasal COMPS, and no change to the ARG-S.

6.6 Summary

In this chapter I have taken up an issue distinct from the question of polysemy, which is nevertheless intrinsic to the lexical syntactic analysis of *te* conjunctive

[6]This follows from Pollard & Sag's (1994) assumption. They make the point that in some cases an NP may have unsaturated complements, citing predicatives (+PRD) such as *an acceptable candidate*. Whether such NP predicates exist in Japanese I don't know, and I can therefore make no predictions about how they affect this analysis.

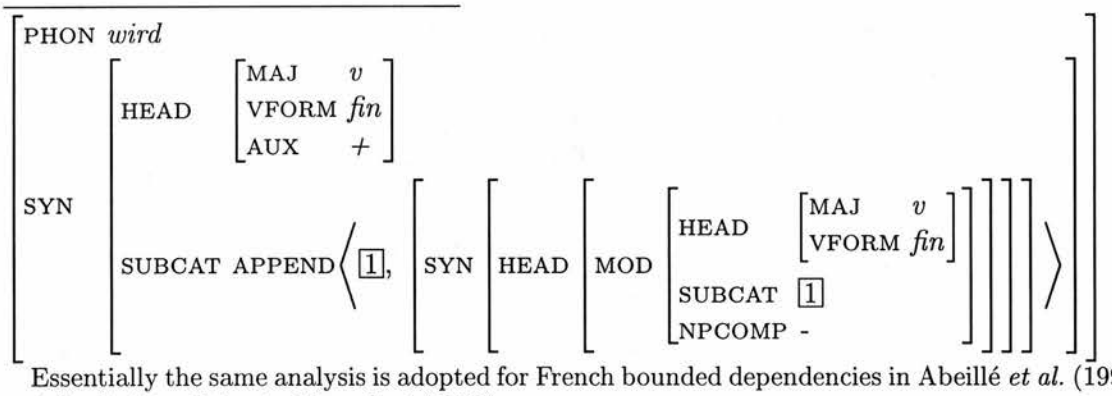
forms. It is a principle among lexicalist theories such as HPSG and LFG that syntactic operations and morphological operations belong to separate domains. Manning *et al.* (1999) are at great pains to point out that Japanese causative verbs are integral lexical morphemes, compounded from two verb forms, the contentive verb and the causative morpheme. The causative verb emerges from the lexicon with integral argument structure, and increased valence over that of the contentive verb. The nuclear conjunctive *V1-te V2* verbal conjunctions are syntactic constructions in which V1 and V2 are lexically distinct, and this has been amply demonstrated in Hasegawa (1995), Matsumoto (1996), and in this thesis. Each verbal conjunct, then, emerges from the lexicon with its own argument structure. A consequence of the Lexical Integrity Hypothesis (Bresnan (1978)) is that in the HPSG framework, argument structure is not structure-shared with (or "passed up to") the phrasal mother: in a Complement-Head phrase, the separate daughters may project argument structure, but the phrasal mother may not. This is attested by the *te miru*, *te morau* and *te ageru* (and other) structures. There is a difference between *te miru* et al, and *te iru* and related structures in that, in combination with adverbial adjuncts, an adjunct associates with the aspectual class of the first verb in *te miru*, but with the second verb in *te iru*. If adjuncts are associated with argument structure, then the first case is explained by having the adjunct associated with the argument structure of V1. This is in conformance with lexical integrity. The challenge is in how to explain the cases where the adjunct is compatible with the second conjunct, and there is no phrase structure rule whose application will reflect this phenomenon. The solution which I put forward in this chapter is to merge the argument structures of the separate conjuncts on the phrasal mother, in a distinguished form of the Complement-Head schema, and attach the adjunct to the phrasal mother[7].

[7]There is an alternative analysis available which (in its later incarnations) does not rely on phrasal argument structure. In their analysis of German auxiliary complexes, Hinrichs & Nakazawa (1989) and Hinrichs & Nakazawa (1994) develop a model of **argument composition** in which the tensed auxiliary verb takes as complements not only the base form lexical verb, but also its complements. In example (37)(a), *finden* (find) is a transitive verb, and *wird* (will) an auxiliary.

(6.37) (a) Ich glaube nicht, daß Peter das buch finden können wird.
I believe not, that Peter the book find can will
"I don't believe that Peter will be able to find the book".
(b) Ich glaube nicht, daß Peter das buch wird finden können.

Using arguments based on the fact that the auxiliary *wird* can "flip" its position, 37(a) versus 37(b), they conclude that the lexical entry for the auxiliary contains the append of the subcategorized verb (i.e. *finden*) and its arguments. Thus, the SUBCAT value of *finden* is also on the SUBCAT list of *wird*, as in the figure below.

(6.38)



Essentially the same analysis is adopted for French bounded dependencies in Abeillé *et al.* (1998) and for Italian clitics in Monachesi (1998).

The essence of the analysis is that the complex argument structure is specified on the lexical entry for the tensed auxiliary and so there is no need to posit a combination of argument structure at the phrasal level. As it stands, the argument composition approach does appear to be lexically stipulated, not arising from any deeper linguistic principles, or consequences of the type hierarchy.

In application to Japanese, the head verb *iru* or *aru* would contain COMPS and ARG-S features containing the arguments of the subcategorized *teform* VP.

Chapter 7

Concluding Remarks

In this thesis I have described the valence alternations of a set of Japanese verbs which can subcategorize for a nominal argument or a verbal argument, and I have argued that some of these verbs are polysemous in various ways. The obverse of this coin is that the remaining verbs have separate senses, and are listed homonymously. In between these two categories, I have also shown that there are verb forms with what look to be regular polysemous alternations, but because of limitations in the HPSG formalism, these have to be separately listed. These are Sense Enumerative in Pustejovsky's (1995) term. My conclusions about the forms of polysemy are developed in Section 7.1.

In developing the lexical account of the Japanese phenomena, it has been necessary to extend the HPSG formalism to take account of the aspectual phenomena which constitute part of the selectional restrictions. This has been quite successful for capturing traditional Vendler-Dowty (Dowty (1979)) categories, but still there are some problems with the underlying model for Accomplishments and Achievements, in particular. This is discussed in Section 7.2.

As a matter which is somewhat separate from the polysemy analysis, but central to the lexicalist analysis of verbs and their complements, I demonstrate in Chapter 6 that a plausible solution to the distinction between verbal conjunctions whose adjuncts are compatible with the first conjunct verb, and those having adjuncts compatible with the second conjunct verb, is that the latter forms exhibit phrasal argument structure. This analysis is based on an assumption that adjuncts are related to the argument structures of their heads in some way. There is still no really satisfactory adjunct analysis in HPSG, however. One way in which this

account might be strengthened, then, is by developing an account of (adverbial) adjuncts which clearly shows their interaction with the ARG-S list.

7.1 Polysemy, Homonymy and Sense Enumerativity

In choosing HPSG (Pollard & Sag (1994)) as a formal representation I have found that the straight valence alternations can be represented by underspecification in the type hierarchy, and the analyses of *miru*, *morau* and *ageru* best exemplify this technique. While *miru* is transitive, *morau* and *ageru* are ditransitive, and the semantic UNDERGOER role in each case takes an argument which can alternate between the syntactic forms of an accusative marked nominal and a *te* marked verb phrase. To make this work formally, the admitted complements are underspecified in the *head* and *marking* type hierarchies, and the UNDERGOER is underspecified in the *content* type hierarchy. For this to work, the *te* marked verb must bear a marking value, accusative in these cases. But there are other verbs with similar polysemous alternations in which it is not only the UNDERGOER role, but the LOCATIVE role, which projects alternately a dative marked nominal or a *te* marked verb phrase. The Stative verbs *iru* and *aru* can be encoded in this way. Much of the phenomena of the Stative verbs *iru* (animate be) and *aru* (inanimate be) can be captured in the same way, through alternate projections of the LOCATIVE role to a dative marked nominal, or a *te* marked verbal. Similarly, the verbs of motion *iku* (go) and *kuru* (come) can be represented by underspecification in the LOCATIVE role. If *te* is underspecified in the *marking* hierarchy for accusative and dative cases, then the entire set of Nuclear Conjunctive forms is represented uniformly in the lexicon.

Of the phenomena which turn out to be completely homonymous on the basis of valence alternations, there is one adjective *hoshii* (want) and two verbs *oku* (put; do in preparation) and *shimau* (put away; complete or regret)[1]. I have discussed these verbs in Chapter 5, and there seems to be no common pattern by which they can be related: while *te oku* involves the dropping of a locative argument vis à vis the simplex form of ditransitive *oku*, the distinction between forms of *te shimau* and simplex *shimau* involves valence change as well as a change of aspectual class, with conjunctive *te shimau* consistent with Achievementhood. Finally, the *te hoshii* (want) form would seem to be more strongly related to the desiderative form of *te morau*, which is *te moraitai* (want to receive) than the simplex adjective *hoshii*

[1]There are other verbs which take *te* conjunctions which I have not investigated in this thesis, including *mieru* (be visible; seem) and miseru (show). Preliminary inspection indicates these are probably homonymous.

(want), with two nominative marked arguments. The reasons for the distinction between polysemous and homonymous lexical forms may well be rooted in processes of grammaticalization (Hopper & Traugott (1993)), but this study has had little to say about homonymy as such, or about the aspects of grammaticalization which might lead to homonymy, as I have focused on the mechanisms necessary for explaining polysemous behaviour.

There is a middle ground, in between the forms which are, and can be represented as, polysemous, and the forms which are homonymous. This middle ground includes verb forms which appear to involve polysemous alternations, but the formal framework is unable to accommodate them as such. These are the alternate projections to subject of *iru* and *aru*, analysed in Chapter 5, and the alternate Agentive and Formal (= resultative) projections of Accomplishment *te* forms, discussed in the same chapter. In the cases of *iru* and *aru*, both these verbs have two roles, one which is a LOCATIVE, and one which can be described as an UNDERGOER. Depending on the context, either one of these roles can be projected to subject, and in principle it should be possible to underspecify the *category* hierarchy in HPSG, along the lines of what Davis (1995) has done with the *content* hierarchy, so as to capture related behaviour in a single lexical form. There are difficulties in describing the contents of a list within a type hierarchy though, discussed in Chapter 5, which effectively prevent polysemous phenomena being represented as such. A solution which allows greater flexibility of representation may well involve reformulating HPSG categories to do away with list valued valences. A second area in which verbal alternations are apparently polysemous, but can only be represented in current HPSG by Sense Enumeration, is the alternation between Agentive and Resultative forms of a transitive Accomplishment, which is realized in the distinction between *te iru* and *te aru* projections. These are indicated by examples (1) and (2), where *shimete iru* indicates the progressive 'closing' form, and *shimete aru* indicates the perfect 'closed' form.

(7.1) Maiko ga doa o shimete iru
Maiko-NOM door-ACC close-TE be
"Maiko is closing the door".

(7.2) doa ga shimete aru
door-NOM close-TE be
"The door is closed".

This Agentive $\longrightarrow$ Resultative intransitivization process was identified by Hasegawa (1995) but the formalism she used, Role and Reference Grammar (Foley & Van Valin

(1984)), incorporating Montague Semantics (Montague (1974)), is unable to countenance a polysemous solution, due to the lack of complex hierarchical types. By adapting Pustejovsky's Event Structure and Qualia Structure as part of the *content* type hierarchy in HPSG, it becomes possible to describe the intransitivization process as a semantic alternation between subtypes. Again though, the problem comes with the attempt to project these alternations to syntax, though this time it is the linkage between argument structure (ARG-S) and complements (COMPS) which is hard to capture in a single lexical entry, in HPSG.

The data underlying the above analyses is suggestive of a process of grammaticalization in the manner of Hopper & Traugott (1993). To place this inference on a stronger basis, a diachronic analysis of the verbal data would be useful. The processes of grammaticalization can perhaps be more readily deduced from such an analysis.

7.2 About Aspect

Although Hasegawa's (1995) work prefigured the need to take account of aspect in formal linguistic frameworks, HPSG has not hitherto incorporated such machinery. However, the phenomena of aspect are an important semantic classifier of verbs, with a set of regular relationships between valences, semantic roles and aspectual classes. Of further interest for this thesis is the discovery that selectional restrictions on verbal conjunctions are sensitive to aspectual features. Accordingly, I have classified the data in this thesis with respect to the Vendler-Dowty aspectual classes, and their underlying features. Thus, *miru*, *oku* and *ageru* subcategorize for an *agentive* verb, *iru* subcategorizes for a *dynamic* verb and *aru* subcategorizes for either an *agentive*[2] verb or an Accomplishment. The nominal valence alternations are accommodated in this scheme if they are maximally underspecified for aspect, and this goes towards permitting the polysemous alternations to be formally realized. The model of aspect I have adapted for HPSG is based on a hybrid of the aspectual calculus of Dowty (1979), and the Event Structure of Pustejovsky (1995). There are, however, limitations in both of these models which lead to a less than complete characterization of the broad range of aspectual phenomena.

In his development of the lexical semantics of verbs, Pustejovsky notes the alternation between causative and unaccusative verbs such as 'break', in examples (3) and

[2]The aspect type hierarchy containing these types: *agentive*, *telic*, *dynamic* and so on is depicted in Chapter 4, Figure 4.9.

(4), by claiming an alternation between the Event Structures of Accomplishments and Achievements.

(7.3) Jiko broke the glass.
(7.4) The glass broke.

His claim is that many verbs are underlyingly bieventual and that some of the event pairs are alternates. In particular, he claims that Accomplishments and Achievements are alternations of the same Event Structure: a *process* followed by a *state*, in strict partial order. Their projections include a single UNDERGOER in the case of Achievements, and ACTOR and UNDERGOER in the case of Accomplishments. These projections are represented by Agentive and Formal Qualia, and are selected by a headedness value, E1 or E2 in this scheme. A verb which is underspecified for this headedness value is polysemous between the two readings. This would be a convenient solution if it were true, but an Achievement is no more, or less, oriented towards the result than is an Accomplishment, and this is perhaps better reflected in Dowty's (1979) calculus, where an Achievement is predicated around a *Become* change of state, and an Accomplishment is encoded as a *Cause* change of state. I claim that it is the distinction in agentivity that marks the difference between these telic verbs, and not any resultative orientation: thus, Accomplishments are bieventual with a *caused process* as first event, while Achievements are bieventual, with an *uncaused process* as first event. Pustejovsky's (1995) causative/unaccusative distinction in the English verb form does not, in any case, carry over to Japanese, because there are no phonological forms with both readings, and projections. The unaccusative form of 'break' is translated as *kowareru* and the causative is *kowasu*; the unaccusative form of 'stop' is *tomaru*, the causative is *tomeru*. There are endless numbers of these verbs in Japanese in which the Achievement and Accomplishment forms differ by their suffices, but which are represented in (one of their) English translation(s) by a single phonological form. These include: *agaru/ageru* (go up); *ataru / ateru* (hit); *chikazuku / chikazukeru* (approach); *fueru / fuyasu* (increase); *hajimaru / hajimeru* (begin); *heru / herasu* (decrease); *kakaru / kakeru* (hang); *kakureru / kakusu* (hide); *kawaku / kawakasu* (dry); *kaeru / kawaru* (change); *magaru / mageru* (bend); *nobiru / nobasu* (extend); *shimaru / shimeru* (close); *shizumu / shizumeru* (sink); *susumu / susumeru* (advance); *tatsu / tateru* (stand); *tokeru / tokasu* (melt); and *tsuzuku / tsuzukeru* (continue). The Event Structures of these verb pairs are not duals differing only by a headedness value, but quite different structures. In fact I have shown that it is the process and result alternations of Accomplishments which are duals of the same Event Structure. The transitive

Accomplishment in each of these pairs, *hajimeru* (begin) and *shimeru* (close), for instance, has process and result alternations of the *te* form, but only a process reading, with the Agentive projection, in finite forms. The Agentive Quale projects both ACTOR and UNDERGOER arguments, but the Formal (or resultative) Quale projects only the UNDERGOER. The Event Structure solution for Achievements is more problematic, because there is a need to reconcile the perfective reading of the *te iru* conjunction with consistency of selectional restrictions across all dynamic aspectual types, by *iru*. If complement selection is by headedness, then for *iru* to be consistent in its conjunction with verbal complements, it must select for E1 headedness uniformly. Pustejovsky's (1995) solution, selection of the Stative projection by E2 headedness, would yield an inconsistency with respect to the E1 selection of Activities and Accomplishments. I have suggested for Achievements that E1 and E2 are overlapping events, and their qualia project the same arguments. In virtue of E1 selection, E2 happens at the same time, and a reading which is ambiguous between process and result ensues. This is not entirely satisfactory, because it cannot separately capture progressive behaviour in some Achievements and perfective behaviour in others, i.e. it cannot separate the *punctual* from the *durative*.

In summary, while my model of Event Structure captures the uniformity of selectional restrictions over all aspect classes and also allows for selection of resultative Accomplishments by *aru*, it does not really provide a satisfactory account of the behaviour of Achievements. Taken together with Dowty's (1979) misgivings about the model of Activities, and about the encoding of Incremental Themes in Accomplishments, there is clearly work still needed on the underlying model of aspect before a better formalization can be produced.

Bibliography

Abeillé A, Godard D, Miller P & Sag I (1998). French Bounded Dependencies, *in Romance in HPSG*, CSLI Publications, Stanford, CA, pp. 1–54. (eds) Balari, S. and Dini, L.

Aronoff M (1976). *Word Formation in Generative Grammar*, Linguistic Inquiry Monograph 1, MIT Press, Cambridge, MA.

Bach E & Partee B (1980). Anaphora and Semantic Structure, *Chicago Linguistic Society* **16**.

Borsley R (1987). Subjects and Complements in HPSG, *CSLI Reports* **99**.

Borsley R (1996). *Modern Phrase Structure Grammar*, Blackwell, Oxford.

Bouma G, Malouf R & Sag I A (1998). A Unified Theory of Complement, Adjunct, and Subject Extraction, *in* G Bouma, G J M Kruijff & R T Oehrle, eds, *Proceedings of the FHCG98. Symposium on Unbounded Dependencies*, Universität des Saarlandes und DFKI, Saarbrücken, pp. 83–97.
*http://www-csli.stanford.edu/ sag/inst-sem/bms.ps

Bresnan J (1978). A Realistic Transformational Grammar, *in* M Halle, J Bresnan & G Miller, eds, *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, MA, pp. 1–59.

Bresnan J (1995). Lexicality and Argument Structure, *in Paris Syntax and Semantics Conference.*

Bresnan J, ed. (1982). *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, MA.

Bresnan J & Kanerva J (1989). Locative Inversion in Chichewa: A Case Study of Factorization in Grammar., *Linguistic Inquiry* **20**, 1–50.

Bresnan J & Mchombo S (1995). The Lexical Integrity Principle: Evidence from Bantu, *Natural Language and Linguistic Theory* pp. 181–254.

Cann R (1999). Specifiers as Secondary Heads, *in* D Adger, S Pintzuk, B.Plunkett & G Tsoulas, eds, *Specifiers*, Oxford University Press, Oxford. To Be Published.

Carpenter B (1992). *The Logic of Typed Feature Structures*, number 32 *in* 'Cambridge Tracts in Theoretical Computer Science', Cambridge University Press, New York.

Carpenter B & Penn G (1998), The Attribute Logic Engine User's Guide, Version 3.1 Beta. Published on the World Wide Web: http://www.sfs.nphil.uni-tuebingen.de/ gpenn/ale.html.

Chierchia G (1984). Anaphoric Properties of Infinitives and Gerunds., *Proceedings of the Third West Coast Conference on Formal Linguistics* pp. 28–39.

Chomsky N (1986). *Knowledge of Language*, Praeger, New York, NY.

Crimmins M & Perry J (1989). The Prince and the Phone Booth: Reporting Puzzling Beliefs, *The Journal of Philosophy* **86**, 685–711.

Cruse D (1986). *Lexical Semantics*, Cambridge University Press, Cambridge.

Dalrymple M (1993). *The Syntax of Anaphoric Binding*, CSLI, Stanford, CA.

Davis A (1995), Lexical Semantics and Linking in the Hierarchical Lexicon, PhD thesis, Stanford University.

Dixon R (1979). Ergativity, *Language* **55**, 59–138.

Dowty D (1979). *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*, D. Reidel, Dordrecht.

Dowty D (1985). On Some Recent Analyses of Control, *Linguistics and Philosophy* **8**, 1–41.

Dowty D (1991). Thematic Proto-Roles and Argument Selection, *Language* **67**, 547–619.

Dowty D & Jacobson P (1989). Agreement as a Semantic Phenomenon, *Proceeding of the Fifth Eastern States Conference on Linguistics (ESCOL '88), Powers, J and de Jong, K (eds)* .

Flickinger D (1987), Lexical Rules in the Hierarchical Lexicon., PhD thesis, Stanford University.

Flickinger D, Pollard C & Wasow T (1985). Structure Sharing in Lexical Representation, *ACL* **23**, 262–267.

Foley W & Van Valin R (1984). *Functional Syntax and Universal Grammar*, Cambridge University Press, Cambridge.

Gazdar G, Klein E, Pullum G & Sag I (1985). *Generalized Phrase Structure Grammar*, Basil Blackwell, Oxford.

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

Gruber J S (1976). *Lexical Structures in Syntax and Semantics*, North Holland, Amsterdam.

Gunji T (1987). *Japanese Phrase Structure Grammar*, Studies in Natural Language and Linguistic Theory, D. Reidel Publishing Company, Dordrecht.

Hasegawa Y (1995). *A Study of Japanese Clause Linkage: the Connective TE in Japanese*, CSLI, Stanford, CA.

Hellan L (1988). *Anaphora in Norwegian and the Theory of Grammar*, Foris, Dordrecht.

Hinrichs E & Nakazawa T (1989). Flipped Out: AUX in German, *Chicago Linguistic Society* **25**.

Hinrichs E & Nakazawa T (1994). Linearizing AUX'es in German Verbal Complexes, *in German Grammar in HPSG*, CSLI Publications, Stanford, CA, pp. 11–37. (eds) Nerbonne, J. and Netter, K. and Pollard, C.

Hiroo (1993). *The Complete Japanese Verb Guide*, Tuttle Language Library, Tokyo.

Hopper P & Traugott E (1993). *Grammaticalization*, Cambridge University Press, Cambridge.

Iida M (1995). *Context and Binding Theory in Japanese*, CSLI, Stanford, CA.

Ishikawa A (1985), Complex Predicates and Lexical Operations in Japanese., PhD thesis, Stanford University.

Jackendoff R (1972). *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, MA.

Jackendoff R (1974). A Deep Structure Projection Rule, *Linguistic Inquiry* **5**, 481–506.

Jackendoff R (1976). Toward an explanatory semantic representation, *Linguistic Inquiry* **7**, 89–150.

Jacobsen W (1991). *The Transitive Structure of Events*, CSLI, Stanford, CA.

Kay M (1980). Algorithm Schemata and Data Structures in Syntactic Processing, *in Readings in Natural Language Processing*, Morgan Kaufman, Los Altos, CA.

Kempson R (1977). *Semantic Theory*, Cambridge University Press, Cambridge.

Kindaichi H (1950). *Nihongo Doshi no Tensu to Asupekuto. In: Nihongo Doshi no Asupekuto*, Mugi Shobo, Tokyo.

Klein E & Sag I (1982). Semantic Type and Control, *Linguistics and Philosophy* **5**.

Kuno S (1973). *The Structure of the Japanese Language*, Vol. 3 of *Current Studies in Linguistics Series*, MIT Press, Cambridge, MA.

Lakoff G (1965), On the Nature of Syntactic Irregularity, PhD thesis, University of Indiana.

Lascarides A & Copestake A (1999). Default Representation in Constraint based Frameworks, *Computational Linguistics* **25**, 55–105.

Levin B (1993). *Towards a Lexical Organization of English Verbs*, University of Chicago Press, Chicago, IL.

Manning C (1994). *Ergativity: Argument Structure and Grammatical Relations*, CSLI, Stanford, CA.

Manning C & Sag I (1999). Dissociations Between Argument Structure and Grammatical Relations, *in* G Webelhuth, J Koenig & A Kathol, eds, *Lexical and*

Constructional Aspects of Linguistic Explanation, CSLI, Stanford, CA, pp. 63–78.

Manning C, Sag I & Iida M (1999). The Lexical Integrity of Japanese Causatives, *in Studies in Contemporary Phrase Structure Grammar*, Cambridge University Press, pp. 39–79. (eds) Robert Levine and Georgia Green.

Martin S (1975). *A Reference Grammar of Japanese*, Yale University Press, New Haven, CT.

Matsumoto Y (1990). On the Syntax of Japanese 'Intransitivizing' *te aru* Construction: Nonlexical Function Changing, *in Proceedings of the Chicago Linguistics Society*, Vol. 26, University of Chicago, Chicago, IL, pp. 277–291.

Matsumoto Y (1996). *Complex Predicates in Japanese: A Syntactic and Semantic Study of the Notion 'Word'*, CSLI, Stanford, CA.

McClure W (1994), Syntactic Projections of the Semantics of Aspect., PhD thesis, Cornell University.

Melčuk I (1979). *Studies in Dependency Syntax*, Karoma Publishers, Ann Arbor, MI.

Miyagawa S (1989). *Structure and Case Marking in Japanese*, Vol. 22 of *Syntax and Semantics Series*, MIT Press, Cambridge, MA.

Moens M & Steedman M (1988). Temporal Ontology and Temporal Reference, *Computational Linguistics* **14**, 15–28.

Molla Aliod D (1997), Aspectual Composition and Sentence Interpretation: a Formal Approach., PhD thesis, University of Edinburgh.

Monachesi P (1998). Decomposing Italian Clitics, *in Romance in HPSG*, CSLI Publications, Stanford, CA, pp. 305–358. (eds) Balari, S. and Dini, L.

Montague R (1974). *Formal Philosophy*, Yale University Press, New Haven, CT.

Moriyama T (1988). *Nihongo Dooshi Zyutugoban no Kenkyuu (Investigations of Japanese Verbal Predicates)*, Meiji Syoin, Tokyo.

Nightingale S (1996), An HPSG Account of the Japanese Copula and other Phenomena, Master's thesis, University of Edinburgh.

Perlmutter D (1978). Impersonal Passives and the Unaccusative Hypothesis, *in Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, University of California, Berkeley, CA, pp. 159–189.

Pollard C & Sag I (1987). *Information Based Syntax and Semantics*, CSLI, Stanford, CA.

Pollard C & Sag I (1994). *Head Driven Phrase Structure Grammar*, The University of Chicago Press, Chicago, IL.

Poser W (1983), What is the 'Double *o*' Constraint a Constraint on? MS, Stanford University, Stanford, CA.

Postal P (1974). *On Raising*, MIT Press, Cambridge, MA.

Pustejovsky J (1995). *The Generative Lexicon*, MIT Press, Cambridge, MA.

Riehemann S (1995), Type Based Derivational Morphology. MS, Stanford University, Stanford, CA.

Rogers A (1971). Three Kinds of Physical Perception Verbs, *Chicago Linguistics Society* **7**, 206–222.

Rosenbaum P (1967). *The Grammar of English Predicate Complement Constructions*, MIT Press, Cambridge, MA.

Ross J (1972). Act, *Semantics of Natural Language, Donald Davies and Gilbert Harman (eds)* pp. 70–126.

Ryle G (1949). *The Concept of Mind*, Barnes and Noble, London.

Sag I (1997). English Relative Clause Constructions, *Journal of Linguistics* **33**, 431–484.

Sag I & Wasow T (1999). *Syntactic Theory: A Formal Introduction*, CSLI, Stanford, CA.

Selkirk E (1982). *The Syntax of Words*, MIT Press, Cambridge, MA.

Shieber S (1986). *An Introduction to Unification Based Approaches to Grammar*, CSLI, Stanford, CA.

Siegel M (1996). Die Japanische Syntax im Verbmobil Forschungsprototypen, *in Verbmobil*, number 133 *in* 'Verbmobil', BMB+F, Saarbrucken.

Talmy L (1978). Figure and Ground in Complex Sentences, *in* J Greenberg, ed., *Universals of Human Language*, Vol. 4, Stanford University Press, Stanford, CA, pp. 625–649.

Tateishi K (1994). *The Syntax of Subjects*, CSLI, Stanford, CA.

Tesnière L (1959). *Elements de Syntaxe Structurale*, C. Klincksiek, Paris.

Tomabechi H (1989), Uniform Treatment of Japanese Postpositions using the Head Marker Structure in HPSG. MS, Carnegie Mellon University, Pittsburgh, PA.

Touretzky D (1986). *The Mathematics of Inheritance Systems*, Research Notes in Artificial Intelligence, Pitman, London.

Vendler Z (1967). *Verbs and Times*, Cornell University Press, Ithaca, NY.

Verkuyl H (1993). *A Theory of Aspectuality: The Interaction Between Temporal and Atemporal Structure*, Cambridge University Press, Cambridge.

Wechsler S (1995). *The Semantic Basis of Argument Structure*, CSLI, Stanford, CA.

Xue P, Pollard C & Sag I (1994). A New Perspective on Chinese *ziji*, *in Proceedings of the West Coast Conference on Formal Linguistics*, CSLI Publications, pp. 432–447.

Yoshimoto K (1996), Tense and Aspect in Japanese., PhD thesis, University of Stuttgart.