

A DEPENDENCY ANALYSIS OF COORDINATION IN ENGLISH AND JAPANESE

Hisami Konishi Springer and Stanley Starosta

University of Hawai'i

Abstract: Coordination can be analyzed within a pure lexicalist dependency framework, without the need for additional constituent structure machinery or orthogonal dimensions of representation. Gapping constructions are fragments relatable to full grammatical counterparts by statable processing strategies. When lexcase coordination analyses are applied to English and Japanese data, Japanese is found to have phrase-level coordination but not clausal coordination.

Keywords: constituent structure, coordination, dependency, English, gapping, Japanese, lexcase

I. INTRODUCTION

In our paper, we make the following assumptions:

- Linguistics is a science.
- The goal of a science is to construct a theory, a falsifiable model of reality.
- To be falsifiable, a theory must be formal and explicit (i.e. generative) and constrained.

We have chosen dependency grammar as our framework of analysis because we believe that it is the best currently available vehicle for constructing a scientific theory of human language. Dependency grammar in its original modern form, as presented by Lucien Tesnière in 1959, is a much more explicit and constrained framework than the increasingly vague and powerful versions of Chomskyan grammar, for example, and research in the lexcase version

of DG since 1971 has resulted in a conceptually simple and formal notation embodying a set of constraints which go well beyond those assumed in Lucien Tesnière's original formulation.

The content of a theory is its constraints. A stronger theory makes a weaker claim, and the contribution that dependency grammar makes to our knowledge of human language is directly proportional to the number of things it rules out as impossible. Our purpose in organizing this panel was to explore ways in which it might be possible to constrain dependency analysis even further, and also ways in which it might be necessary to increase its descriptive power to accommodate facts and generalizations about natural language. In our own contribution, we would like to explore one facet of this question, the analysis of coordinate constructions.

A. *Coordination in dependency grammar*

It is a thesis of dependency grammar in general that syntactic structure can and should be represented as a network of pairwise word-to-word regent-to-dependent links, rather than in terms of constituency, a successive partitioning of sentences into smaller and smaller chunks. While this has worked smoothly in most cases, a number of dependency approaches have maintained that coordinate constructions cannot be analyzed in this way, and instead require that the expressive power of the theory be increased by for example allowing for constituent structure analysis be incorporated in the dependency system. Frameworks which require this kind of exceptional treatment of coordination include Word grammar (Hudson 1989:57-94, esp. 62-65, 86), Mel'čuk's dependency grammar (Mel'čuk 1988), Pragian Functional Generative Dependency Grammar, and Sebastian Shaumyan's applicative Universal Grammar:

'Some linguists contend that the dependency concept applies to coordinate construction whereas others deny that coordinate constructions can be defined in terms of dependency (Hudson 1984; Melcuk 1988).' (Sebastian Shaumyan, dg-net, 4 October 1996)

Word grammar (Hudson, Rosta)

'From a Word Grammar perspective, coordination seems the chief, and perhaps only, area of syntax where dependency analyses won't work.' (And Rosta, dg-net, 14 Nov 1994)

'It is possible to dispense with constituent structure in this theory because everything that needs to be said about the syntactic structure can be said in terms of binary relations between single words (except for the special case of coordinate structures).' (Hudson 1990:12)

'According to WG, there are just two parts of the grammar where there are larger units made up of smaller ones. We have just discussed one of them -- morphology, where the larger units are words. The other is coordination, where the larger units are word-strings and the parts are conjunctions and other word-strings.' (Hudson 1990:97)

'The notion 'word-string' corresponds closely to 'constituent', but is used in WG only in the analysis of coordination so as to avoid the conflicts raised by examples like (1b), where the constituents demanded by coordination are not phrases as defined by dependency.' (Hudson 1990:84)

'Special sub-types are recognized for both 'word' and 'word-string' -- 'noun', 'verb', etc. for 'word', and 'coordination' for word-string.' (Hudson 1990:84-85)

Applicative Universal Grammar (Shaumyan)

'...My linguistic theory Applicative Universal Grammar abandons the treatment of the concept of dependency as an intuitive undefined primitive notion. Instead it defines the concept of dependency concept in terms of more fundamental concept--the typed functional relation OPERATOR-OPERAND....In terms of this relation, I distinguish two concepts of dependency: 1) applicative dependency and 2) constitutive dependency....The above definitions have several advantages....Third, they allow to widen the scope of dependency grammar to cover coordination.' (Sebastian Shaumyan, dg-net, 4 October 1996)

Praguian Functional Generative Dependency Grammar. An alternative modification of the theory is to add an additional dimension of dependency analysis to the representation of coordinate constructions, as in the Praguian approach:

'If certain basic linguistic intuitions, made convincing by a broad agreement among linguists of different schools, are reflected, then a fundamental difference is to be made between (a) those relations which can be rendered by dependency or (admittedly) by constituency, and (b) coordination (and perhaps apposition).' (Petr Sgall, dg-net, 28 September 1993)

'Coordination (as well as apposition, the formal properties of which are similar) has to be handled by a specific means. However, it seems that a treatment of coordination and apposition as relations of a third dimension...is well possible within a dependency framework, see Pit'ha and Sgall (1972); a formalism rendering this dimension by a specific kind of brackets in the linearized notation was presented by Platek et al. (1984), and recently elaborated by Petkevic (1988).' (Sgall and Panevová 1989:81)

'...working with a network of three dimensions; this may seem too complex a device to be understood as an adequate description of a relatively easily learnable system (cf. the position of Chomsky's UG), but, fortunately, the relationships between the three dimensions are strongly restricted (cf. the condition of projectivity, or Dick Hudson's 'adjacency'), so that the whole network can be linearized if two kinds of parentheses are used (for dependency and for coordination); in other words, the device is rather close to a string.' (Petr Sgall, dg-net, 28 September 1993)

B. Pure dependency

Other dependency frameworks however maintain that coordinate constructions can be accommodated without abandoning a pure monostratal dependency approach, but that this requires the invocation of precessing strategies to account for conjunction reductin and gapping. These include David Milward's 'Dynamic Dependency Grammar', lexicase (Starosta 1988: 210-216, 246-255), and computational implementations by Hening Lobin (1993) and Peter Hellwig.

David Milward, Dynamic dependency grammar

'I know of no analyses of coordination which use just dependency relations. However this shouldn't be taken as a criticism of dependency grammars, since I think it is fair to

argue that there are no really successful treatments of 'coordination in terms of constituency either. The best attempts are in the Lambek Calculus tradition, but these suffer from problems of over generation. In my own work I've been led reluctantly to the conclusion that coordination has to be treated at a second level of analysis, possibly as part of processing.' (David Milward, dg-net, 15 Nov 1994)

Lexicase. Lexicase dependency grammar handles coordination in terms of dependency rather than constituency. To make this work, it is necessary to regard coordinate constructions as exocentric constructions with special properties, and to analyze reduced and gapped coordinate constructions as non-grammatical fragments rather than as full well-formed configurations. The remainder of our paper will address these two points, and then look at the special case of coordination in Japanese from the point of view of the system we have established at that point.

II. COORDINATE CONSTRUCTIONS AS TRANSPARENT EXOCENTRIC CONSTRUCTIONS

A. Exocentric constructions

Definition and formalization. An EXOCENTRIC CONSTRUCTION is defined in lexicase dependency grammar as a word with one or more structurally necessary dependents. That is, heads of exocentric constructions never occur without certain dependents, even in 'movement' or other unbounded dependency situations. If X is the head of an exocentric construction and Y is one of its obligatory dependents, then X will bear a lexical contextual feature [?+[+Y]]. Exocentric constructions are represented in stemmas by a horizontal line:

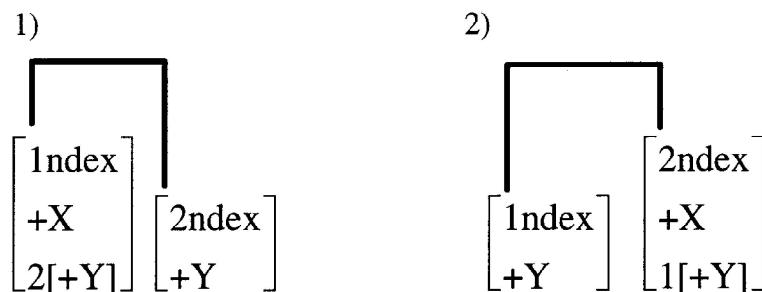


Fig. I. Exocentric stemmas

where 1ndex and 2ndex are addresses, numbers reflecting a word's relative position in a linear string. In both stemmas, X is the REGENT (the dominant member of a pair of grammatically related words) and Y is the DEPENDENT. This is represented in two ways: 1) by placing Y lower in the stemma than X, and 2) by marking X as the member of the relation that licenses Y as its dependent.

Linear precedence may also be represented lexically in a lexicase dependency grammar by contextual features on the regent word indicating the the relative position of the dependent:

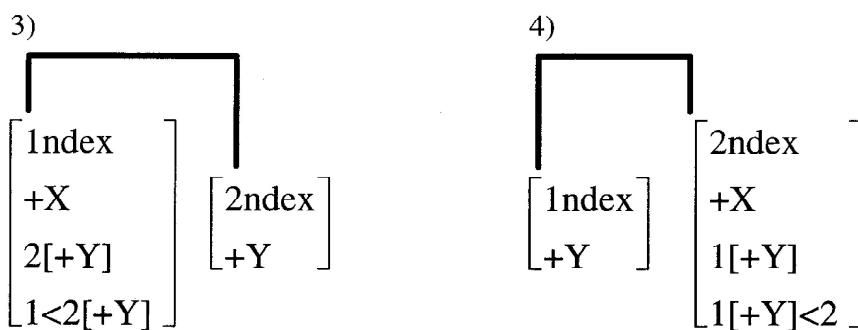


Fig. II. Linear precedence

Here the feature $1<2[+Y]$ indicates that the regent X [1ndex] precedes the dependent Y [2ndex] in 3), and follows it in 4). Since all the information about dependency links and linear precedence is represented in the lexical matrices, the stemma itself is redundant, and the same information can be given equivalently in a randomly ordered string of words:

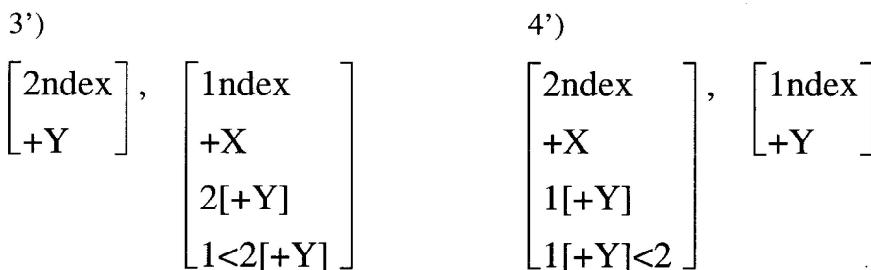
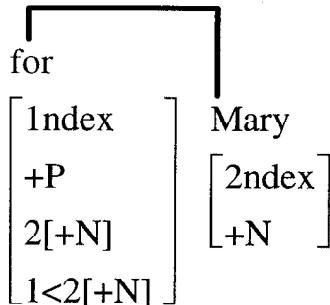


Fig. III. String representation

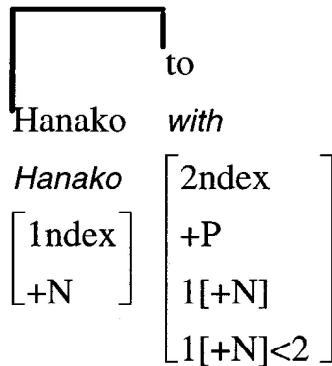
This aspect of the notation has important implications for computational applications, but the stemma notation will be retained in our paper because it is more familiar-looking and thus hopefully more legible.

PP and CnjcP. Of the eight syntactic word classes allowed in the lexcase framework (*V*, *Adv*, *P*, *N*, *Adj*, *Det*, *Cnjc*, *Sprt*), two are found in exocentric constructions: *P* and *Cnjc*. English and Japanese examples of these two patterns are given in Figure IV.

5) English PP



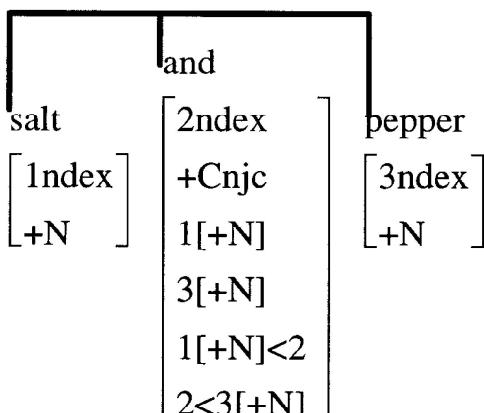
6) Japanese PP



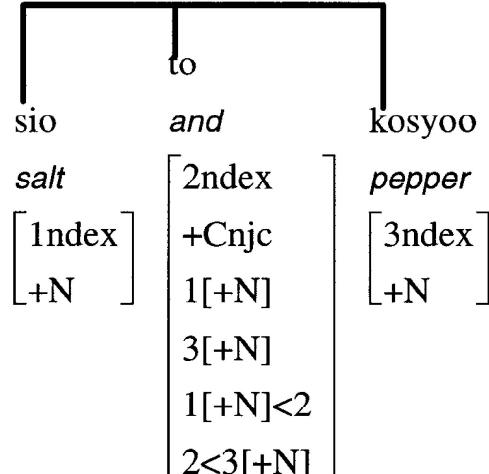
‘with Hanako’

Fig. IV. Prepositional and postpositional phrases

7) English NP



8) Japanese NP



‘salt and pepper’

Fig. V. Coordinate noun phrases

As indicated in these diagrams, prepositions and postpositions differ from conjunctions in that the former have only one structurally required dependent while the latter have two.

Optional dependents. Heads of exocentric constructions may also have structurally optional dependents. As in classical dependency grammar stemmas, structurally optional dependents are indicated by slanting edges. The lexical representation for a structurally optional dependent has the form $[?([+Z])]$. Thus English prepositions and Japanese postpositions may take adverb dependents, and conjunctions may take one or more additional left conjuncts:¹

¹This analysis of left conjuncts was first proposed by Lawrence Reid.

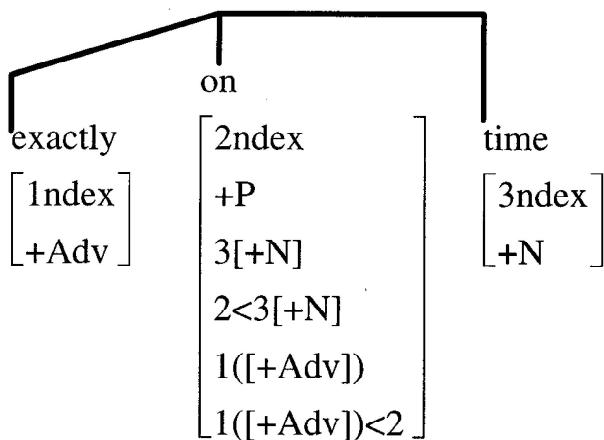
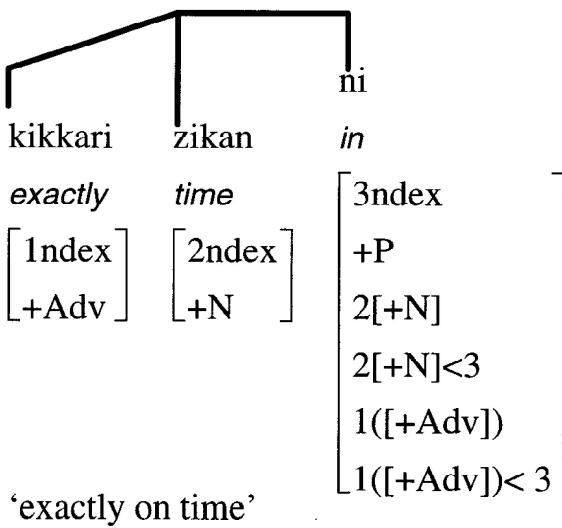


Fig. VI. Optional dependents of P: English

Note that since there is no restriction on the relative order of the adverb and the N dependent of P, both orders are allowed in the Japanese examples in Figure VII.

9)



10)

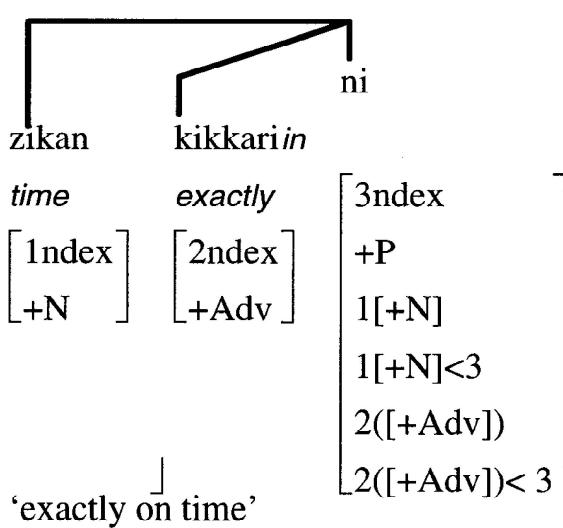


Fig. VII. Optional dependents of P: Japanese

In such stemmas, the horizontal line should be regarded as belonging to the lexical head of the construction, so that *vinegar*, *salt*, and *pepper* are all dependents of *and*. The linear precedence feature $[1([+N])<2([+N])]$ in Figure VIII indicates that the optional dependent *vinegar*[1index] must precede the obligatory dependent *salt*[2ndex]. The Japanese counterpart of this pattern, shown in Figure IX, is identical in structure.

Multiple optional dependents. A given exocentric head word may allow any number of optional dependents. This is accounted for in this notation by feature FEATURE CLONING: in parsing, each time the question marks in the feature $[?([+X])]$ and any associated linear precedence features such as $[@<?([+X])]$ are instantiated by indices of actually cooccurring dependent words, the original contextual features are copied back into the matrix to allow for an additional dependent (see Figures X and XI). There is no structural limit on how often this may be repeated.

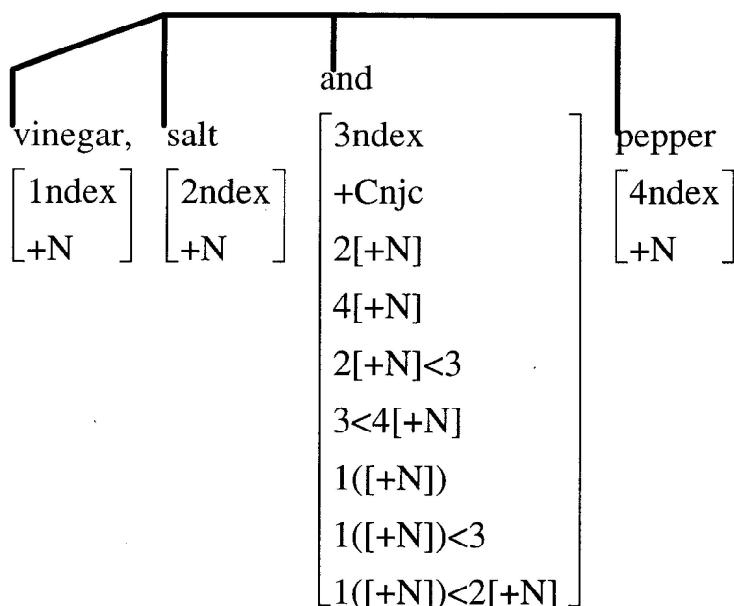


Fig. VIII. Optional dependents of Cnjc: English

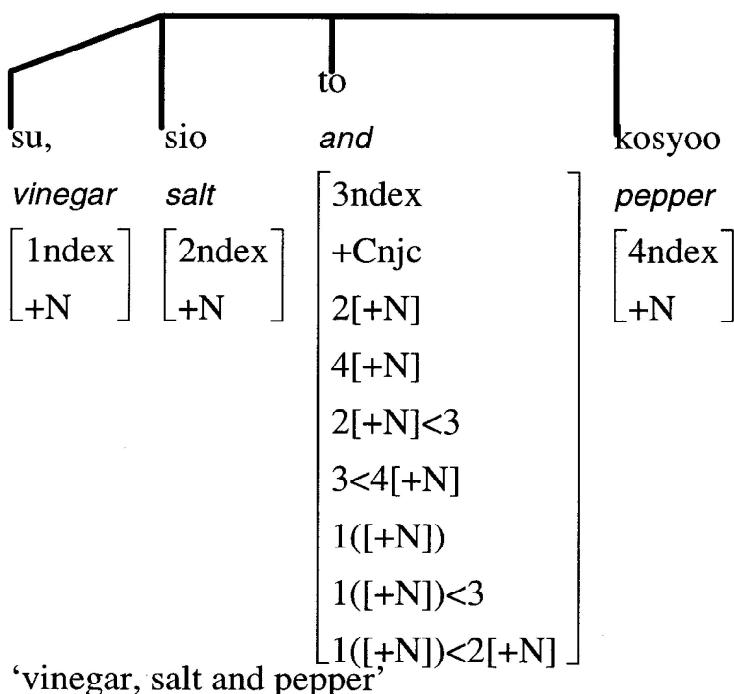


Fig. IX. Optional dependents of Cnjc: Japanese

It should be noted at this point that so far, no special notational provisions have been made for coordinate constructions as such. Coordinate constructions are just a special case of exocentric construction, and everything needed to account for coordination in a lexically based dependency system is also independently required for prepositional phrase constructions. This is equally true of the final property which needs to be incorporated into the analysis: TRANSPARENCY.

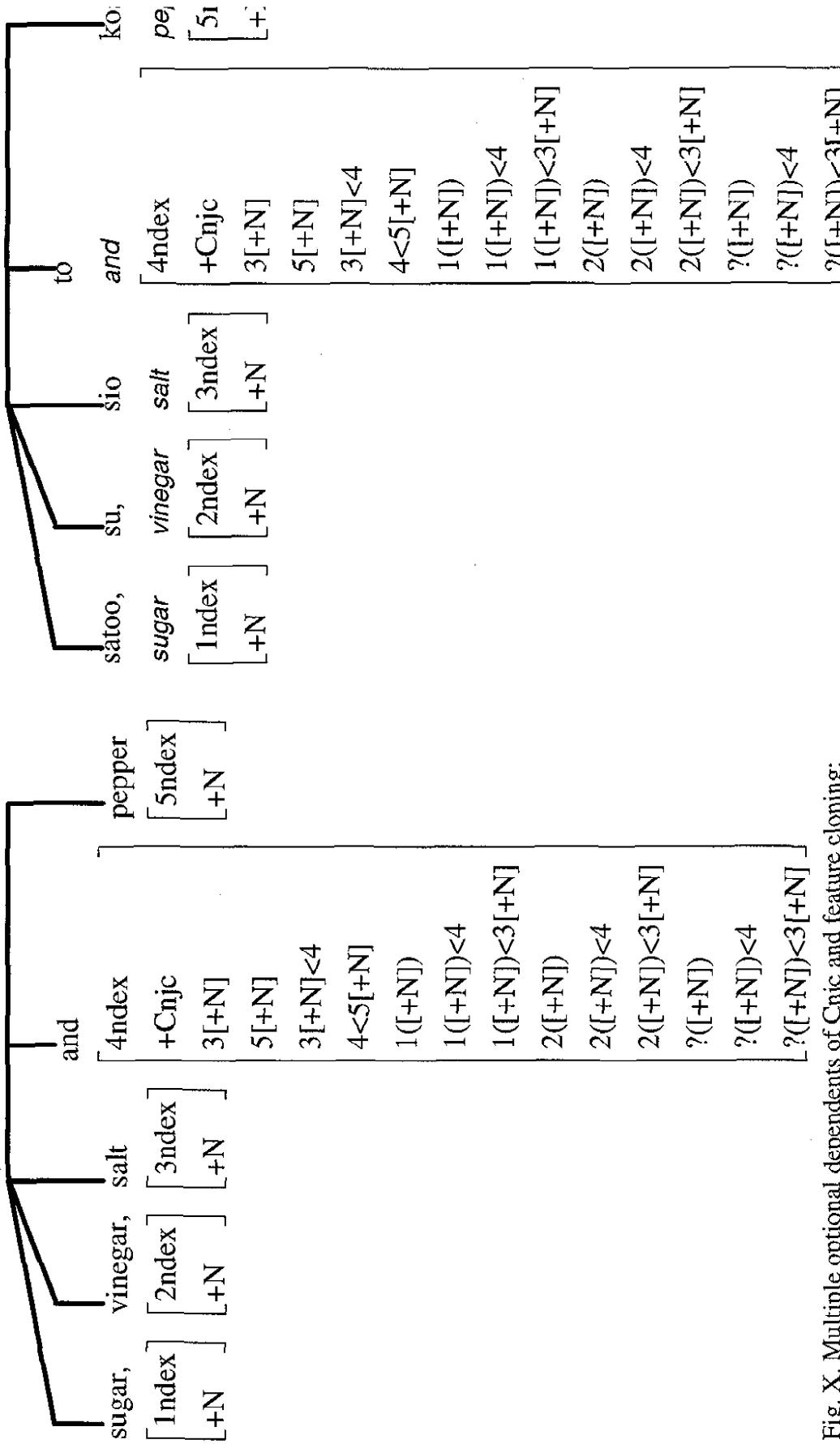


Fig. X. Multiple optional dependents of Cnjc and feature cloning:
English

'sugar, vinegar, salt, and pepper'

Fig. XI. Multiple optional dependents of Cnjc and feature cloning:
Japanese

B. Transparency

PPs and conjoined NPs. Lexical heads of exocentric constructions are 'transparent' to requirements imposed by their regents on their dependents; in another metaphor, they carry the requirements imposed by their regents, and may pass them on to their dependents. Thus Ps bear case relations for their obligatory N dependents, and conjunctions bear the number, case form, and case relation for their dependent NPs. Thus in Figure XII), *stayed* assigns the case relation PAT directly to the 'subject' *Murphy*, but assigns the case relation LOC 'Locus' to *in*, which carries it for the whole construction *in the closet*. In Figure XIII, *todomatta* assigns the case relation LOC 'Locus' to the postposition *ni*, which carries it for the whole construction *osiire ni* 'in the closet', and also assigns PAT 'Patient' to the postposition *ga*, which carries it for the 'subject' PP, *Murphy ga*.

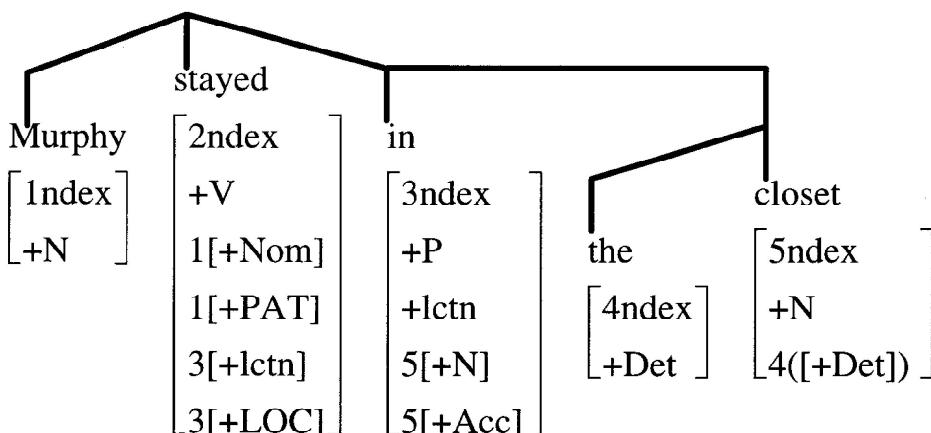


Fig. XII. Transparency in PPs: English

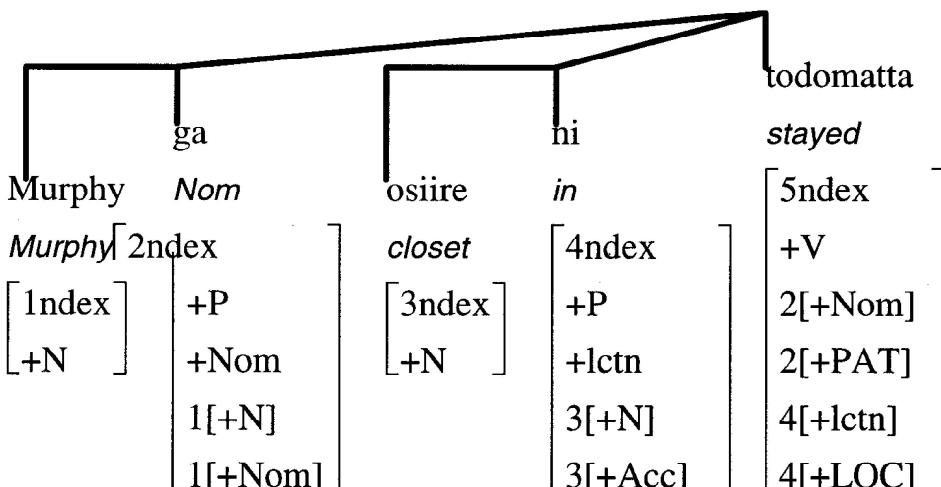


Fig. XIII. Transparency in PPs: Japanese

'Subject' agreement features in English are selectional features in the sense of Chomsky 1965, rather than grammatical features. They mark interpretations which a regent word imposes on a dependent, and a construction is acceptable to the degree that a hearer accepts the interpretation. In the English example, the verb *disagree* is looking for a plural nominative noun dependent, and it interprets *and* as a plural nominative noun. *and*

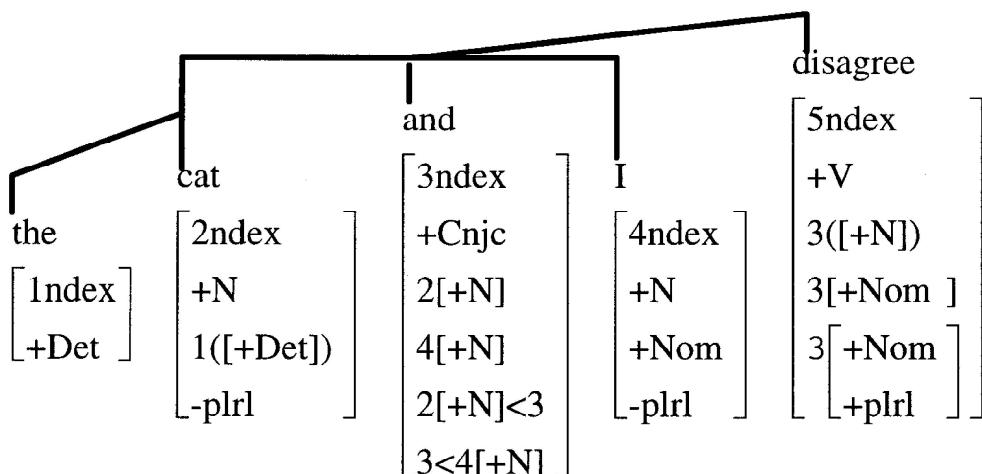
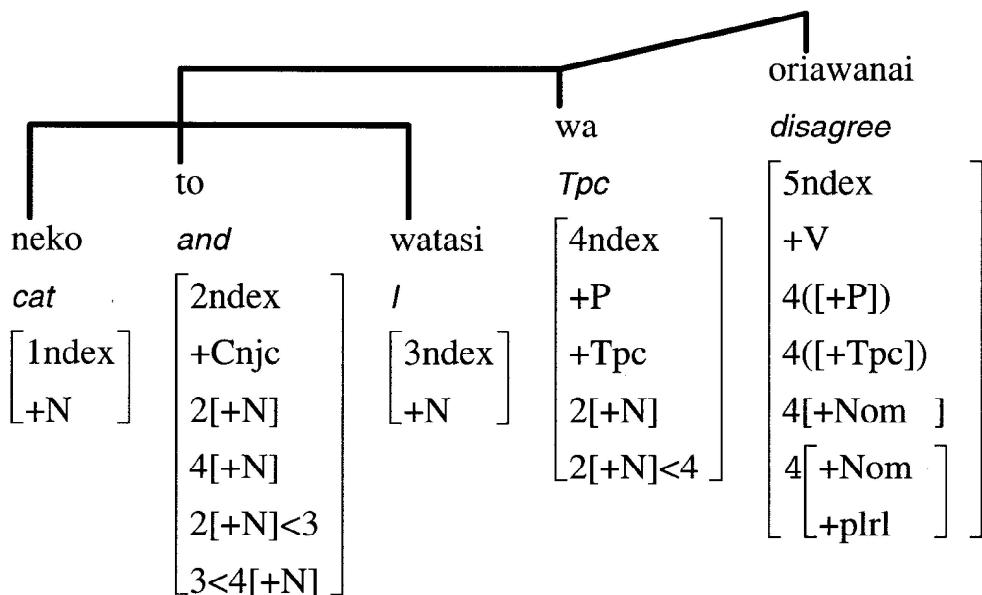


Fig. XIV. Transparency in conjoined NPs: English



‘The cat and I disagree.’

Fig. XV. Transparency in conjoined NPs: Japanese

passes on this interpretation to its noun dependents, *cat* and *I*.² *I* is lexically nominative and *cat* is lexically unmarked for case, so they do not contradict this interpretation. *cat* and *I* are both lexically singular, but the plural interpretation is assigned to the head of the construction, *and*, and a plural interpretation of the whole NP is not pragmatically inconsistent with a conjunction of two singular elements, so the sentence is acceptable. The Japanese example

² A similar example is mentioned in Sgall and Panevova 1989:81 (7):

7) Jane and my brother are a happy couple.

‘in such cases as (7) the coordination can be described as taking place between the two nouns as the heads of the corresponding syntagms (subtrees).’

works in the same way, with the additional machinery required to accommodate the *wa* topic postposition.³

Sentences containing NPs with more than two conjuncts then will be treated in exactly the same way. Thus in an example like 11),

11)⁴ The teacher, the politician, the postman, and the sailor all signed up for the quiz show.

signed assigns the Nominative case to *and*, and *and* passes this interpretation on to the NP as a whole.

This notation also takes care of hierarchical coordination, such as found in the example represented in Figures XVIII and XIX.

The fully specified lexical entries for conjunctions such as *and* and *to* are shown as Figure XVI.

12)

and

@ndex
+Cnjc
?[+N]
?[+N]
?([+N])
@<?[+N]
?[+N]<@
?([+N])<@
?([+N])<?[+N]

13)

to 'and'

@ndex
+Cnjc
?[+N]
?[+N]
?([+N])
@<?[+N]
?[+N]<@
?([+N])<@
?([+N])<?[+N]

Fig. XVI. Lexical entries for English and Japanese coordinating conjunctions

where @ ('self') is replaced by a sequential index in context, and where additional optional dependent features are cloned as necessary. Most of the information shown here is predictable by general rule, so that only the information shown in Figure XVII actually needs to be recorded in the lexicon.

³Cf. Springer 1993b for a discussion of Japanese topicalization in a dependency framework.

⁴Brian MacWhinney, dg-net, 16 November 1994

$$\begin{aligned}
 14) \quad & \text{Bob} \\
 & \text{and} \\
 & \left[\begin{array}{c} +\text{Cnjc} \\ ?[+N] \end{array} \right]
 \end{aligned}
 \quad
 \begin{aligned}
 15) \quad & \text{and} \\
 & \text{to 'and'} \\
 & \left[\begin{array}{c} +\text{Cnjc} \\ ?[+N] \end{array} \right]
 \end{aligned}$$

With the conventions defined so far, this entry is sufficient to generate all the coordinate NP constructions defined so far.

Clausal coordination. English coordinate clauses are analyzed in the same way as coordinate NPs, except that the conjunctions are subcategorized for verbs instead of nouns, as shown in Figure XX. Multiple causal coordination is analogous to multiple NP coordination; following Laurie Reid's analysis, two conjuncts are complements and any additional ones are adjuncts. This is illustrated in Figures XXI and XXII.

Fig. XVII. Minimal lexical entries for English and Japanese coordinating conjunctions

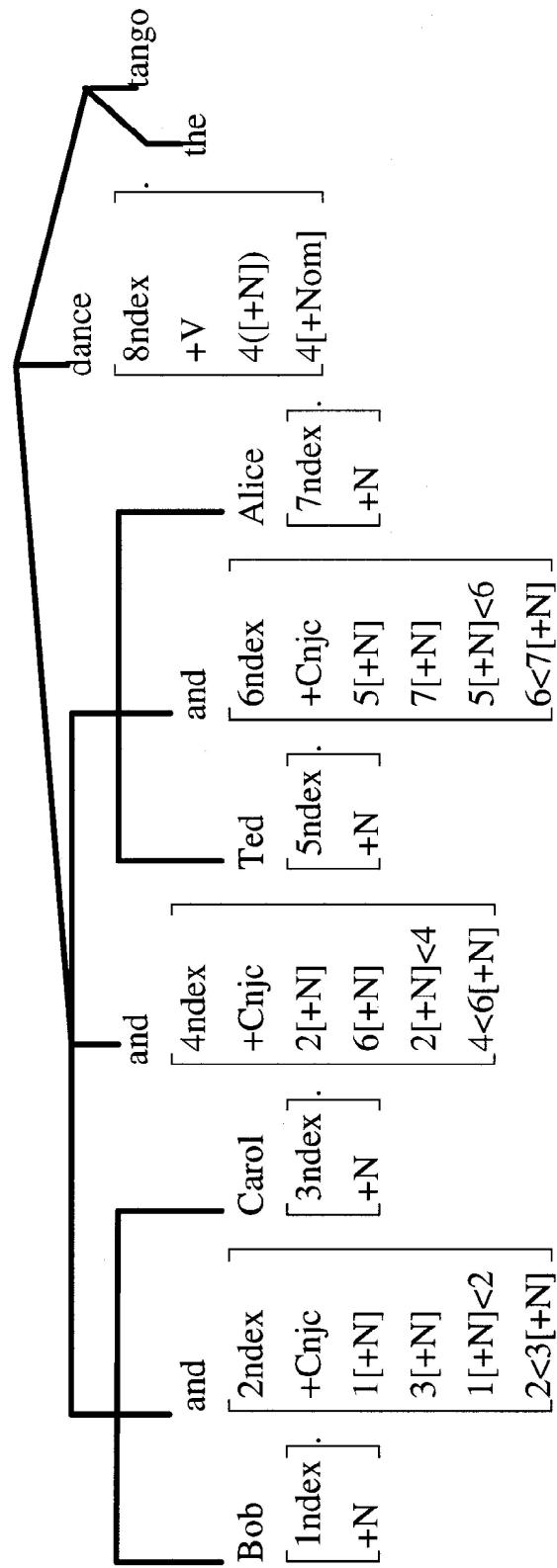


Fig. XVIII. Hierarchical coordination: English

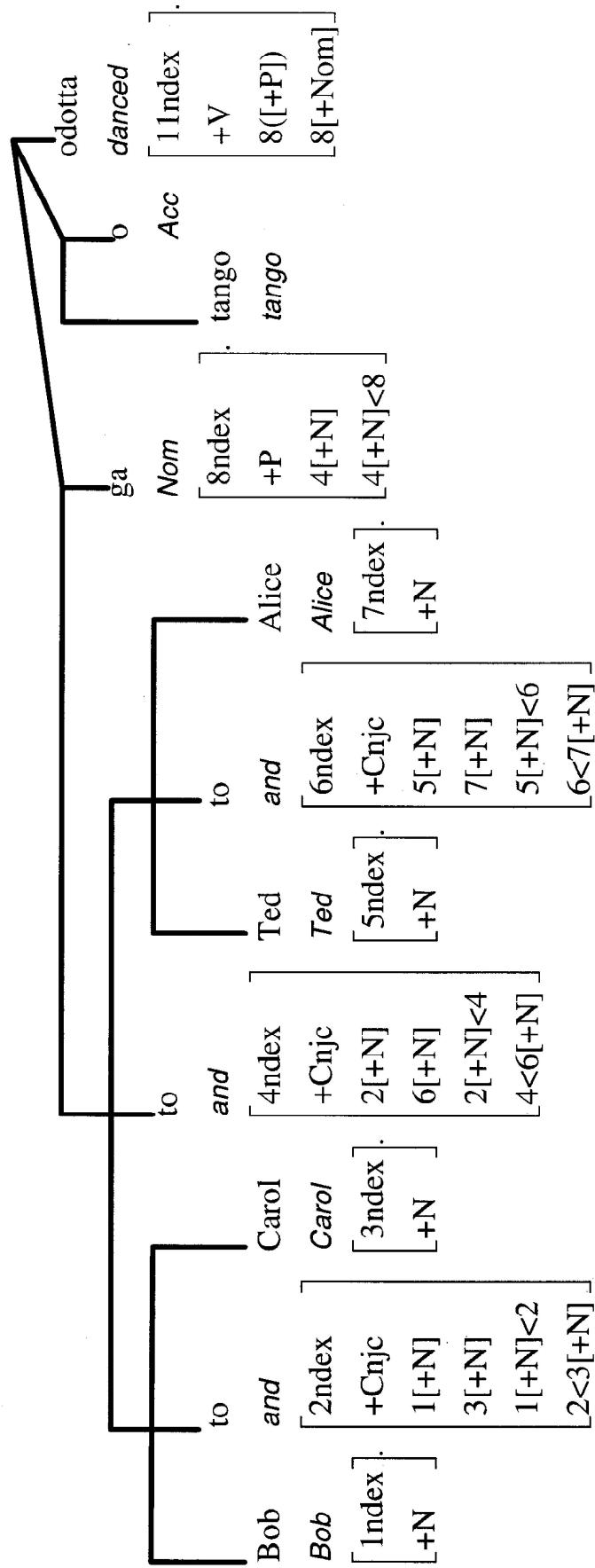


Fig. XIX. Hierarchical coordination: Japanese

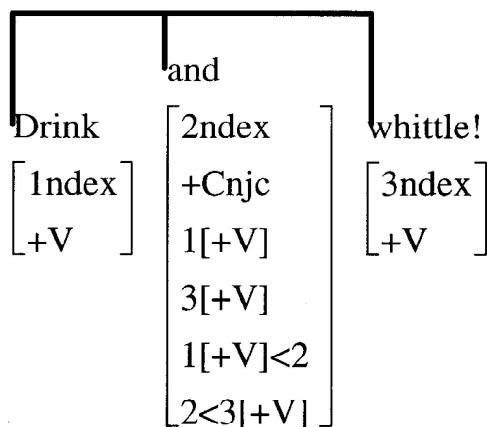
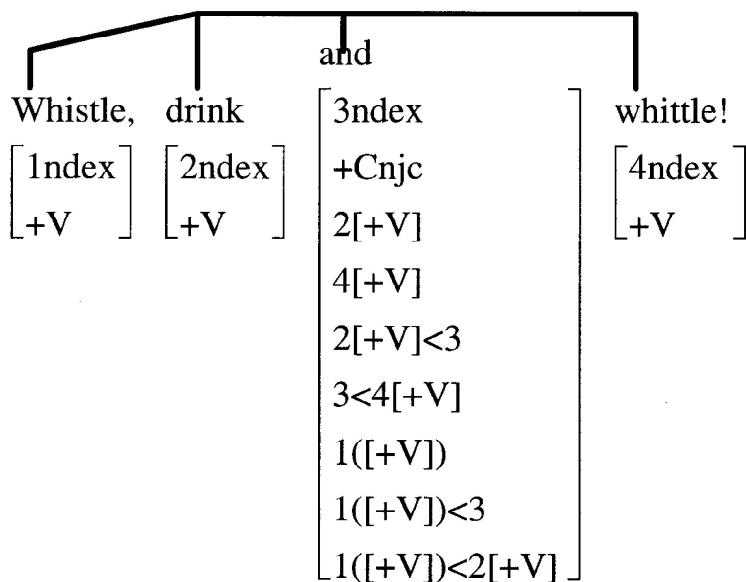
Fig. XX. Simple clausal coordination⁵

Fig. XXI. Multiple clausal coordination

As far as we can determine, direct Japanese structural counterparts to these clausal coordination constructions do not exist. Instead, the Japanese translations are subordinating constructions, with no conjunction and with the non-final clauses in non-root or infinitival or imperative inflected forms. Examples are given as 16) - 20).

16) <u>Nonde</u> utae!	'Drink and sing!'
17) <u>Nomi</u> , utae!	'Drink and sing!'
18) <u>Nome</u> , <u>sosite</u> utae!	'Drink and then sing!'
19) <u>Kutibue</u> o <u>huki</u> , <u>nomi</u> , <u>kirikizame</u> !	'Whistle, drink, and whittle.'

⁵The examples in this section might be uttered as answers to the question, 'What do I need to do in order to qualify as a 'good ol' boy'?

20) Geppu o si, kutibue o huki, nomi, kirikizame! ‘Belch, whistle, drink, and whittle.’

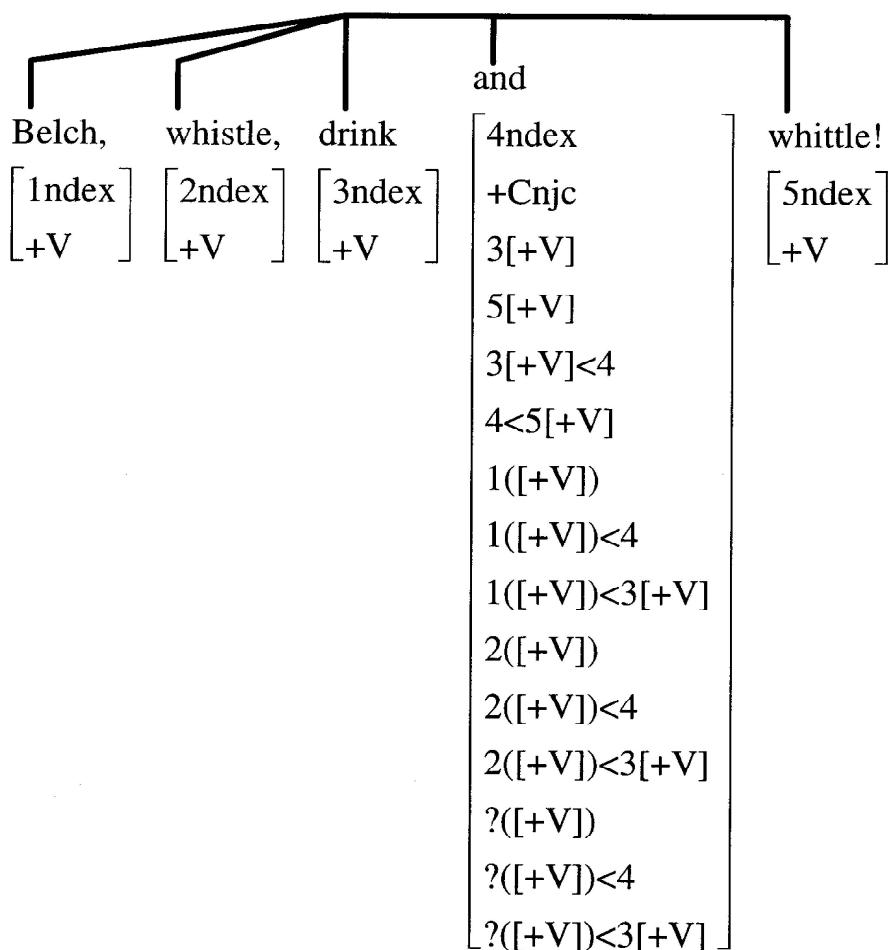


Fig. XXII. Multiple adjuncts, feature cloning

The example in Figure XXIII illustrates the property of transparency in the embedding of coordinate clauses. The Japanese translation of this example however is structurally not coordinate but subordinate, as indicated by the bracketing and non-root verb forms:

21) Watasi wa [[[inu ga hoe] neko ga kamu] to sitte] imasu
I Tpc dog Nom barking cat Nom bite that knowing am
 ‘I know that dogs bark and cats bite.’, literally something like
 ‘I am knowing that with dogs barking, cats bite.’

The example in Figure XXIV, English hierarchical clausal coordination works the same as phrasal coordination. Once again, the Japanese counterpart is a subordinating rather than a coordinating construction:

22) Nonde untensuru ka nomanaide ikinokoru ka karera wa sasizusita
drinking drive or not drinking remain alive or they Tpc ordered
 “They ordered (someone to choose between) driving (while) drinking or staying alive by not drinking.”

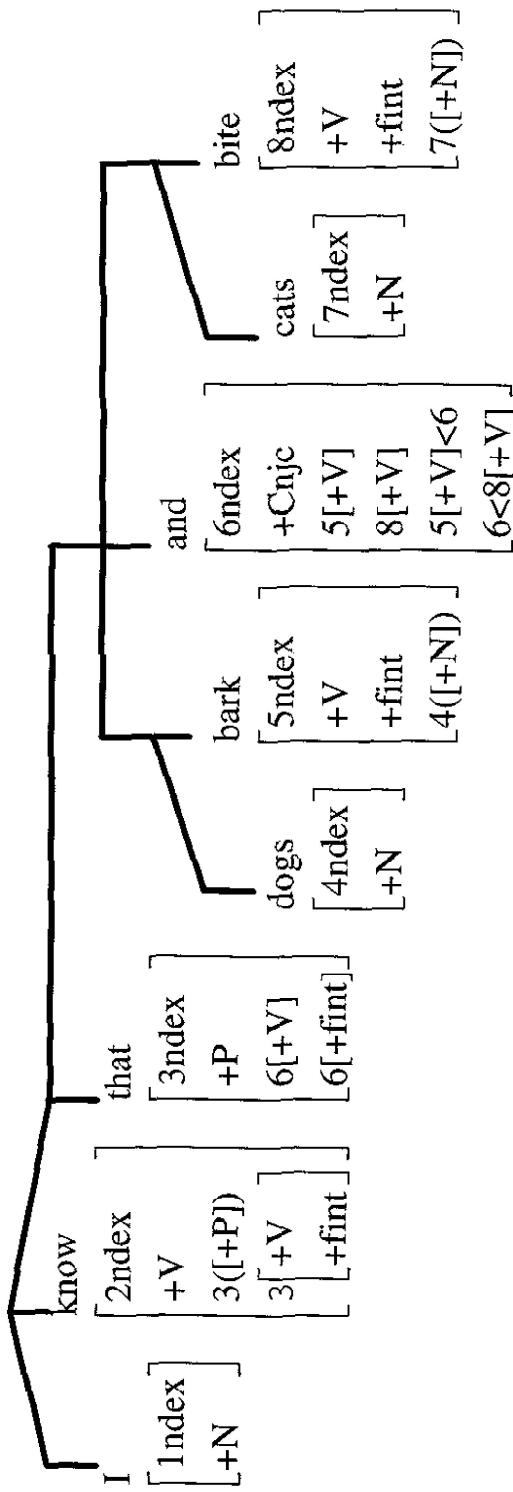


Fig. XXIII. Transparency in clausal complements: English

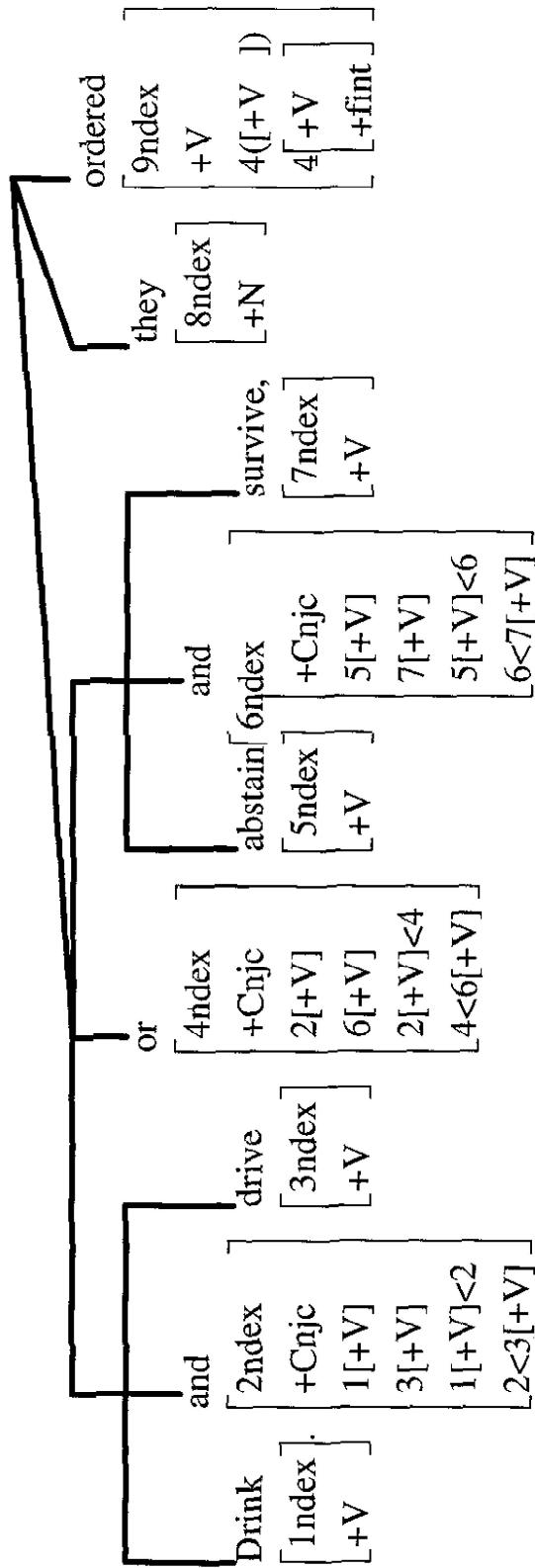


Fig. XXIV. Hierarchical clausal coordination: English

The correct analysis of the *ka* in this example is not immediately obvious. It is almost certainly related to the interrogative Sentence Particle *ka*, and does not have the syntactic distribution of coordinating conjunctions.

III. COORDINATION REDUCTION

The more interesting and troublesome aspect of coordinate constructions of course is not how to account for what is there, but how to account for what is not there, as illustrated in the following sections.

A. *Dependent omission*

Data. Coordination reduction examples

23)

That young man and woman are engaged.

Sono wakai dansei to zyosei wa konyaku site iru.
that young single man and single woman Tpc engaged doing are
 'That young man and woman are engaged.'

24)

The dog and the cat in the basement are fighting again.

Tikasitu ni sumu inu to neko ga mata arasotte iru.
basement in reside dog and cat Nom again fighting are
 'The dog and the cat in the basement are fighting again.'

25)

A king and queen who traveled by dogsled became famous.

inuzori de tabi o sita oo to zyooo ga yuumei ni natta.
dogsled at trip Acc did king and queen Nom famous in became
 'A king and queen who traveled by dogsled became famous.'

26)

Maybelle got up and washed her hair.

Maybelle ga okite kanozyo no kami o aratta.
Maybelle Nom rising she of hair Acc washed
 'Maybelle got up and washed her hair.' [subordinate construction]

27)

Harvey rearranged and Bob vacuumed the living room.

Harvey ga moyoogae o site Bob ga ima ni sooziki o kaketa.
Harvey Nom arranging Acc doing Bob Nom living room in vacuuming Acc did
 'Harvey rearranged and Bob vacuumed the living room.' [subordinate construction]

28)

Kelly hugged and kissed the wombat.

Kelly ga dakishimete wombat ni seppun sita.
Kelley Nom hugging wombat in kiss did
 'Kelly hugged and kissed the wombat.' [subordinate construction]

Missing elements. Intuitively and/or grammatically, all of these examples have one or more missing elements, indicated by the Δ s in the following representations.

23')

That young man and Δ woman are engaged.

Sono wakai dansei to Δ zyosei wa konyaku site iru.

In 23), the common singular noun *woman* requires a determiner. It has none, but the sentence is acceptable nonetheless, and the noun is interpreted as if it had the determiner *a*. The two persons described in the Japanese and English sentences are both interpreted as being young.

24')

The dog Δ and the cat in the basement are fighting again.

Tikasitu ni sumu inu to Δ neko ga mata arasotte iru.

In 24), the dog is normally interpreted as being in the basement along with the cat.

25')

A king Δ and Δ queen who traveled by dogsled became famous.

Inuzori de tabi o sita oo to Δ zyooo ga yuumei ni natta.

In 25), the king is interpreted as being in the dogsled with the queen, and both are interpreted as indefinite singular in the English version.

26')

Maybelle got up and Δ washed her hair.

Maybelle ga okite Δ kanozyo no kami o aratta.

In example 26), *washed* is a finite verb and thus requires a subject. Even though it lacks a subject here, the sentence is quite acceptable, and is interpreted to mean that Maybelle is the subject of both finite verbs. In the Japanese example, both *okite* and *aratta* are finite verbs, though *okite* is non-root, and both thus allow a clausemate subject. When the second is missing, it is interpreted as identical with the first.

27')

Harvey rearranged Δ and Bob vacuumed the living room.

Harvey ga Δ moyoogae o site Bob ga ima ni sooziki o kaketa.

In 27), *rearranged* is a transitive verb, and requires a cooccurring object. Here it has none, but the sentence is acceptable and *rearranged* is interpreted as having the object *the living room*.

28')

Kelly hugged Δ and Δ kissed the wombat.

Kelly ga Δ dakishimete Δ wombat ni seppun sita.

Finally, *hugged* in 28) is a transitive verb missing its object, and *kissed* is a finite verb requiring but lacking a subject. Both missing elements are consistently identifiable, and the sentence is acceptable.

Each of the examples above is a less complete version of the following sentences:

23")

That young man and that young woman are engaged.

Sono wakai dansei to sono wakai zyosei wa konyaku site iru.

24")

The dog in the basement and the cat in the basement are fighting again.

Tikasitu ni sumu inu to tikasitu ni sumu neko ga mata arasotte iru.

25")

A king who traveled by dogsled and a queen who traveled by dogsled became famous.

Inuzori de tabi o sita oo to inuzori de tabi o sita zyooo ga yuumei ni natta.

26")

Maybelle got up and Maybelle washed her hair.

Maybelle ga okite, Maybelle ga kanozyo no kami o aratta.

27")

Harvey rearranged the living room and Bob vacuumed the living room.

Harvey ga ima o moyoogae site, Bob ga ima ni sooziki o kaketa.

28")

Kelly hugged the wombat and Kelly kissed the wombat.

Kelly ga wombat o dakishimete, Kelly ga wombat ni seppun sita.

Fragment analysis. Over the years, a number of attempts were made within the lexicase framework to generate such reduced coordinate examples directly and to identify the missing elements by grammatically conditioned rules of the same type used in control and unbounded dependency ('movement') constructions, but the efforts were not successful. Even if the expressive power of the framework was dramatically increased, it was still not possible to account explicitly for examples such as 23), in which the missing sequence *that young* is not a constituent in either the constituency or dependency sense.⁶

The conclusion we have reached as a result of these efforts is that the theory is dutifully performing its assigned heuristic function. Part of the content of a theory is a claim about which phenomena belong together. A good theory tells us how to 'carve nature at its joints', and the decision about which phenomena to group together is an empirical one, not one that is given a priori. If an account of missing elements in coordinate constructions cannot be accommodated as a purely grammatical matter, then a different and not purely grammatical analysis may be sought for such examples. The approach adopted in lexicase (Starosta 1988:253-254) is to regard examples such as those in 23) - 28) as fragments which are related to their grammatically complete counterparts 23") - 28") by means of processing strategies (cf. Lobin 1993). Such strategies enable a speaker to compact a sentences in such a way that the hearer can unambiguously reconstruct the missing elements.⁷ This is the correct analysis if it is superior to alternative analyses which do not make this assumption, applying the usual criteria of generativity, minimal expressive power, and language-specific and cross-linguistic generality in comparing alternative accounts.

LXC representation. Our dependency representations of sentences such as 23) - 28) will show the missing elements in angle brackets, as shown in Figure XXV. Stemmas with the bracketed elements present will be directly generated by the grammar, and the stemmas with the bracketed elements omitted will be an acceptable but ungrammatical fragments derived from the corresponding grammatical sentences by processing rules.

We maintain that this is not just a facile way of shoving the problem under the rug. If we can provide an explicit account of the principles of compacting and reconstructing such fragments, then we will have made a testable claim about the nature of coordinate constructions, and will have shown the adequacy of dependency notation for accounting for all grammatically well-formed constructions, including coordinations. Following are the procedures and constraints we have found so far.

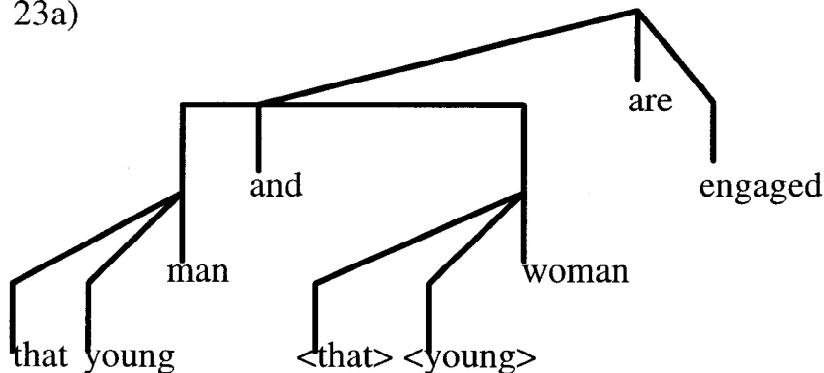
⁶The term CONSTITUENT when used in a dependency context refers to any word (the HEAD of the constituent) plus all its direct and independent dependents. Similarly, a 'noun phrase' (NP) is a noun plus its dependents, and a clause or sentence (S) is a predicate plus all its dependents. In lexicase dependency grammar, these terms are used only for expository purposes. They are derivatively defined, and play no role in the formulation of grammatical generalizations. It may be noted in this connection that Hudson's Word Grammar does not distinguish the concepts of HEAD and REGENT, and uses the term 'head' in both senses.

⁷A computational analogy is compacting files to save storage space and time in transfer.

- A following dependent in the first conjunct or a preceding dependent in the second conjunct may be omitted under identity.

More precisely, if a coordinate construction has the form shown in Figure XXVI, where X and Y are contiguous sequences of words which depend directly or indirectly on their regents R_1 and R_2 but are not necessarily constituents, where X and Y are identical in form to W and Z respectively, and where X and Y bear the same grammatical relations to their respective regents R_1 and R_2 as Z and W do,⁸ then X and/or Y may be omitted without affecting the acceptability of the resulting sequence.

23a)



23a')

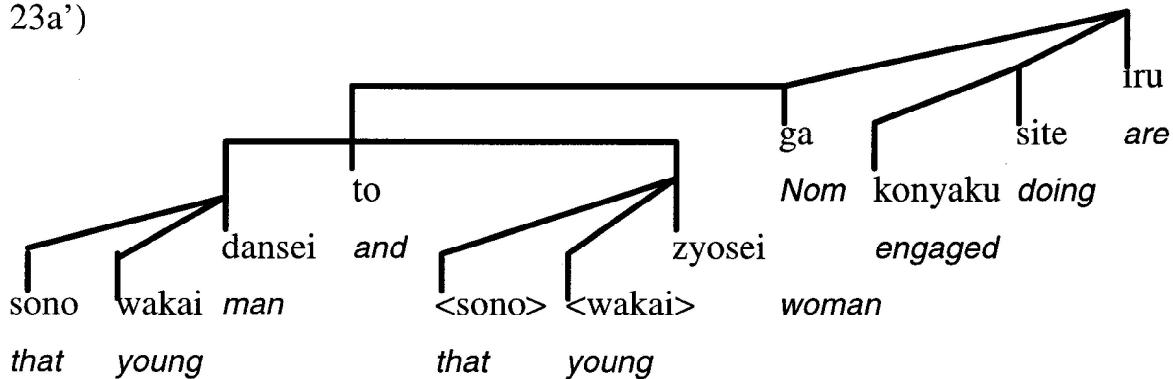


Fig. XXV. Hierarchical clausal coordination: English

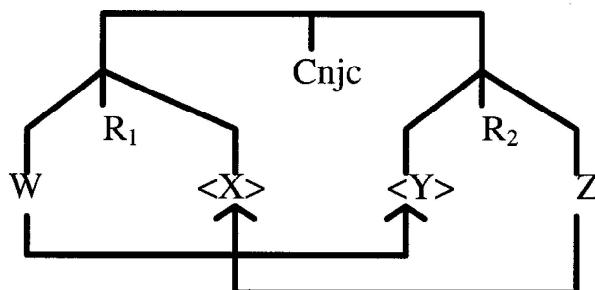


Fig. XXVI. Gappable sequences in coordination

⁸Since regent-dependent relations in this framework are encoded as contextual features in the lexical matrix of the regent item, the requirement on identity of grammatical relations can be formalized by requiring the regents to belong to the same syntactic class and subclass.

Only two types of elements may be compacted in this way, coordinated nouns (Figure XXVII) or coordinated predicates (Figure XXVIII).⁹

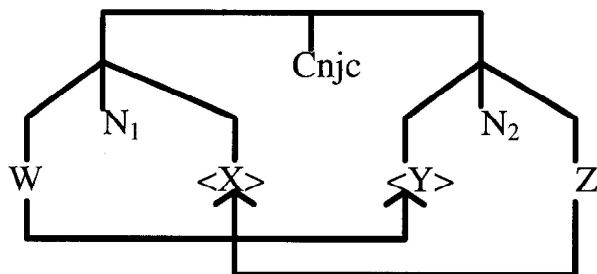


Fig. XXVII. Coordinated Ns, preceding and/or following dependents of N omissible

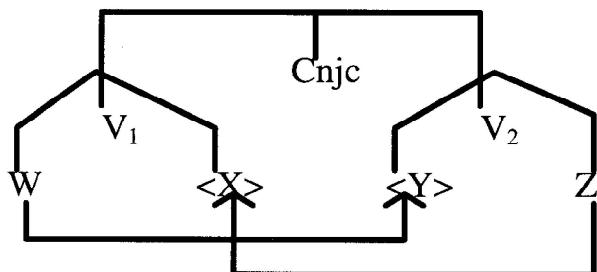
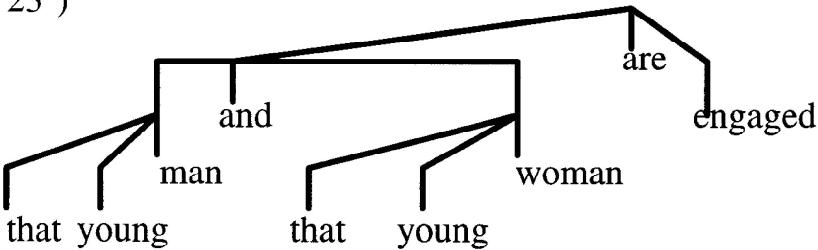


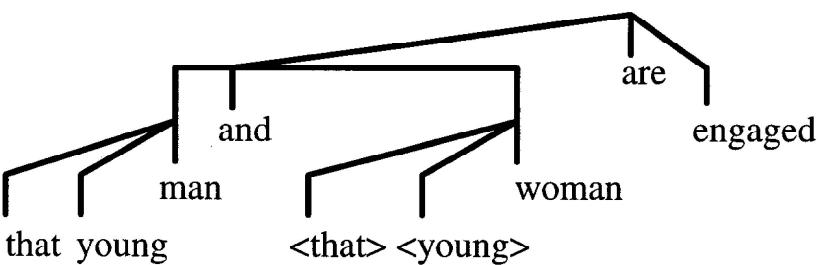
Fig. XXVIII. Coordinated Vs, preceding and/or following dependents of V omissible

Examples 23) - 28) are thus incomplete sentence fragments related to the grammatically complete sentences 23") - 28") by processing rules: If 23") - 28") are grammatical and satisfy these requirements, then 23) - 28) will be acceptable, though not, strictly speaking, grammatical:

23")

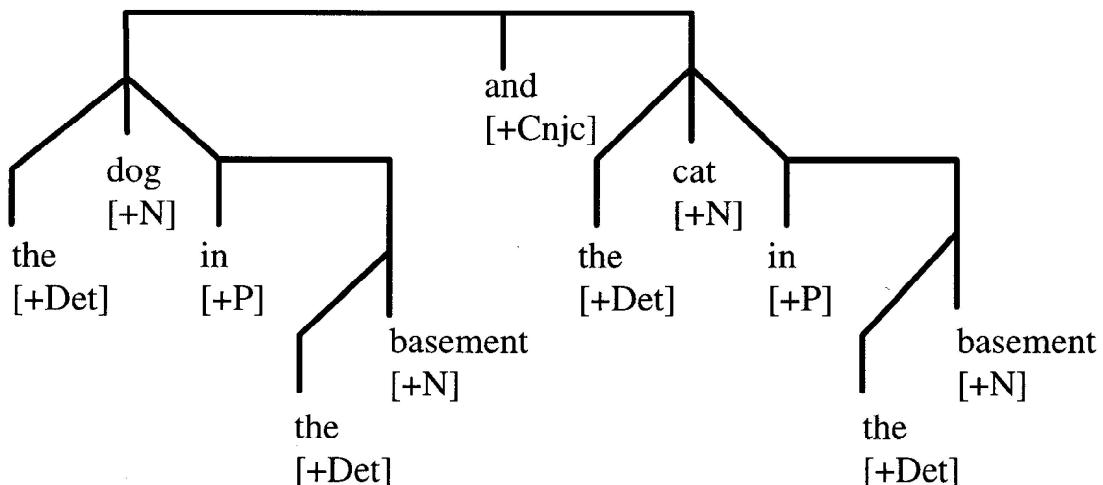
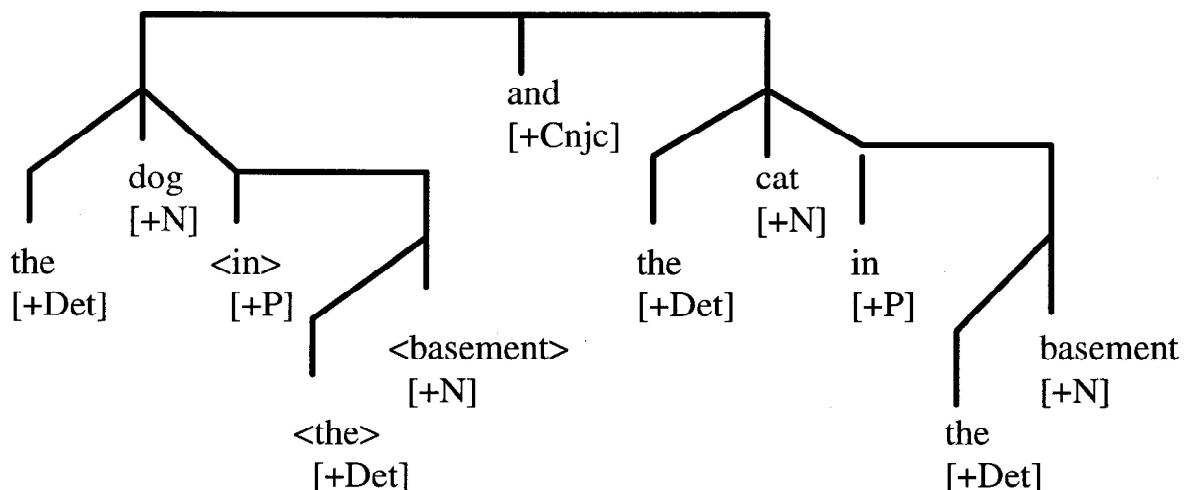


23"") Omitted Det and Adj; forward reduction (W → Y)

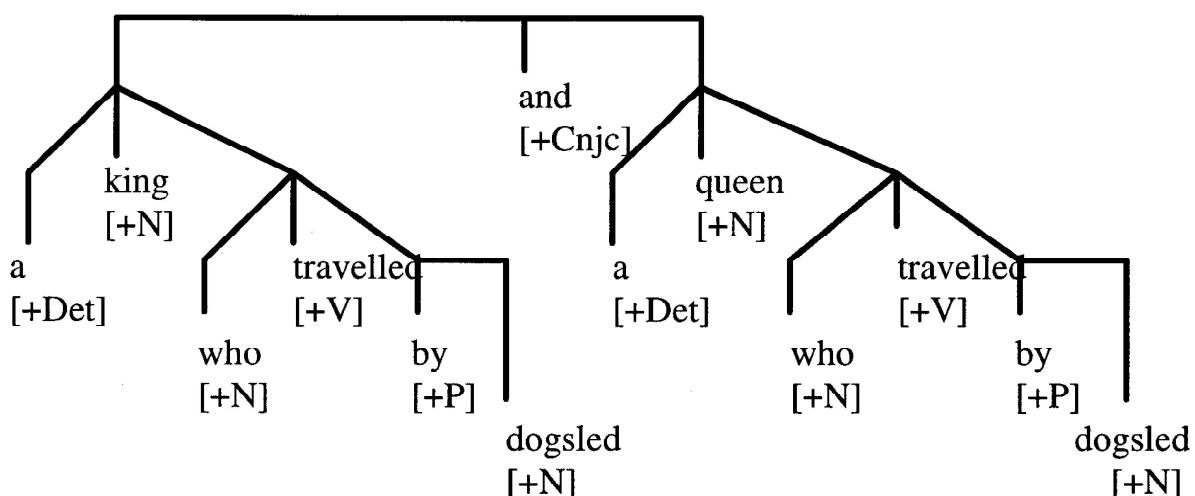


⁹ Predicates may be either verbs, nouns, or prepositions, depending on the language. For the remainder of this paper, we will use verbs to represent the more general class of predicates.

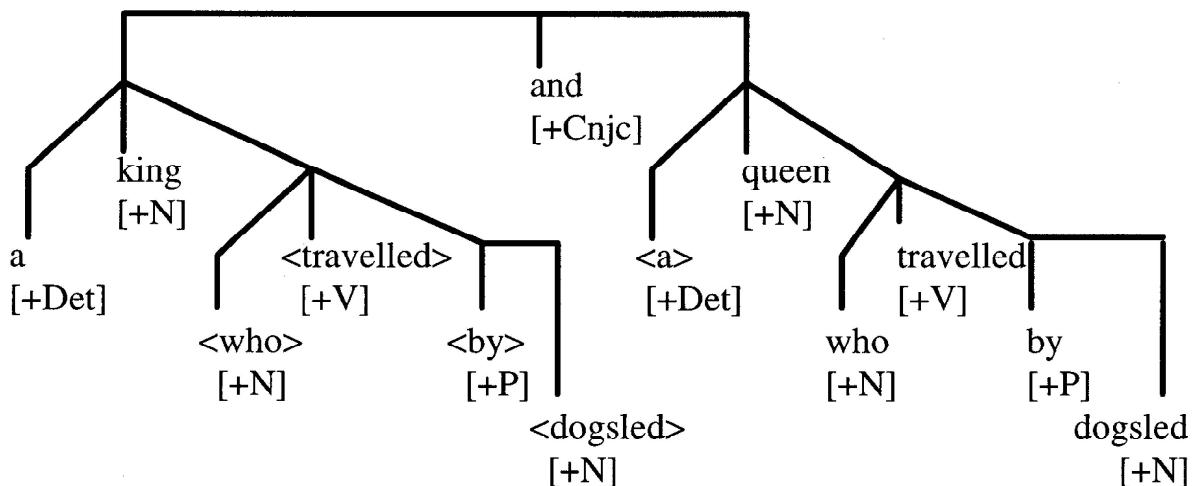
24")

24'') Omitted PP: backward reduction ($Z \rightarrow X$)

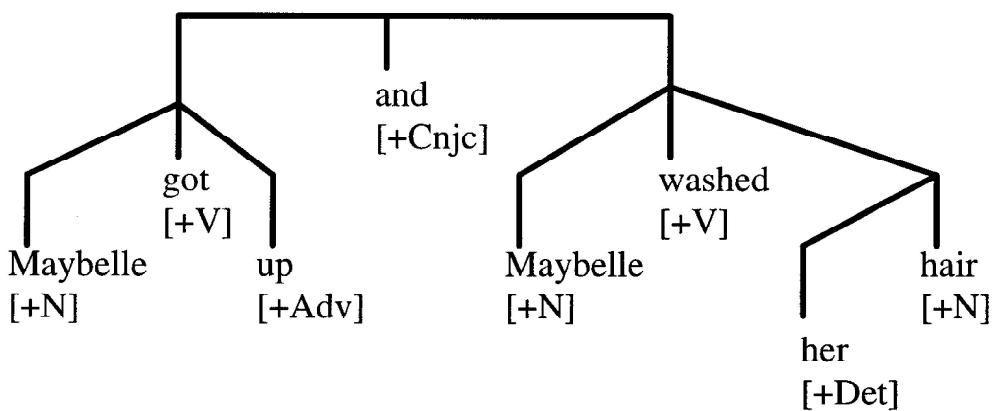
25")



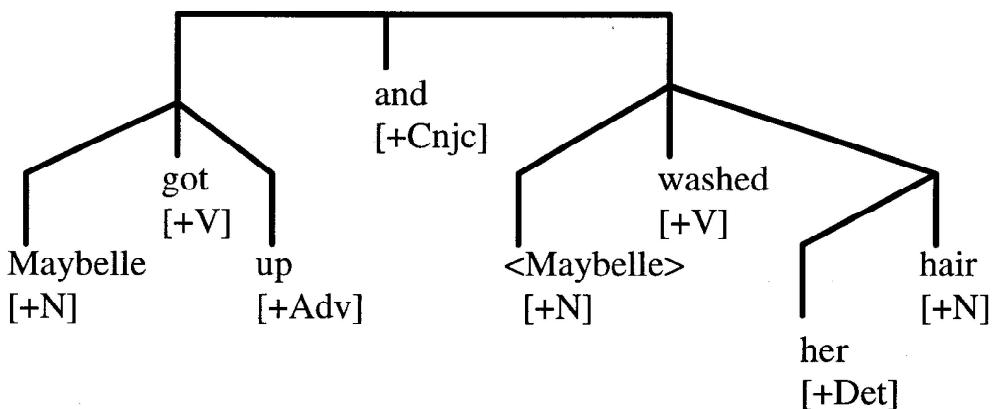
25'') Omitted Det and relative clause: forward ($W \rightarrow Y$) and backward ($Z \rightarrow X$) reduction



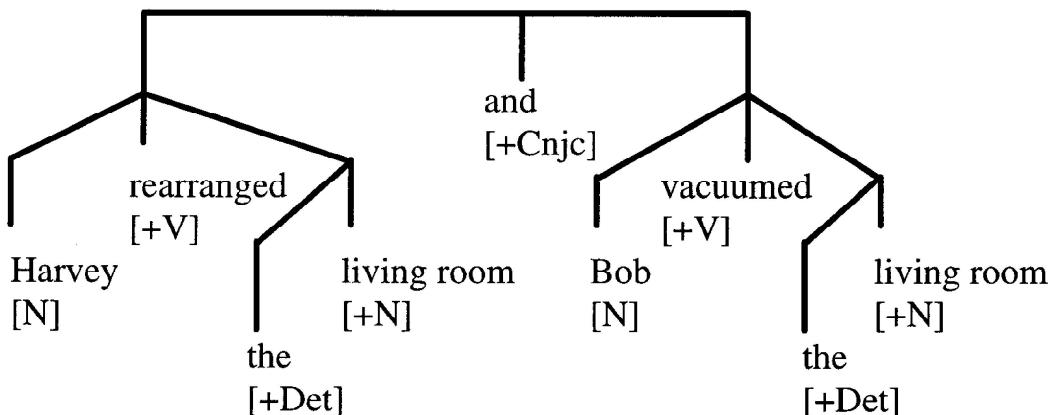
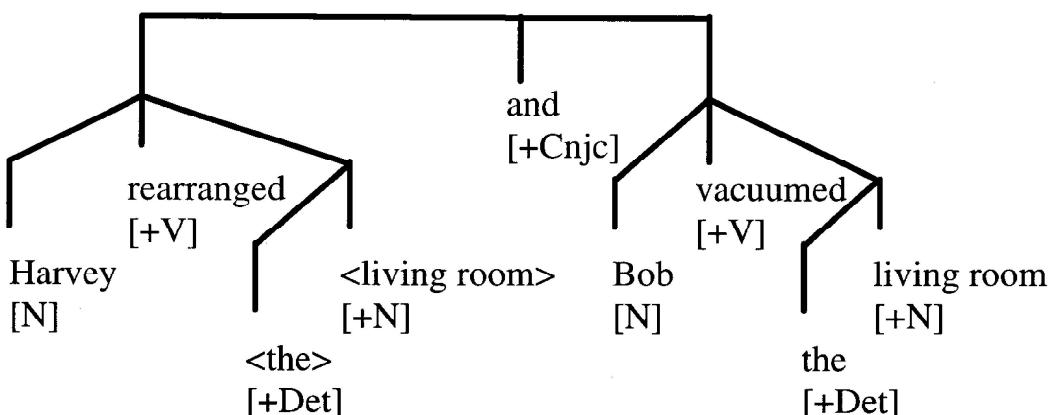
26'')



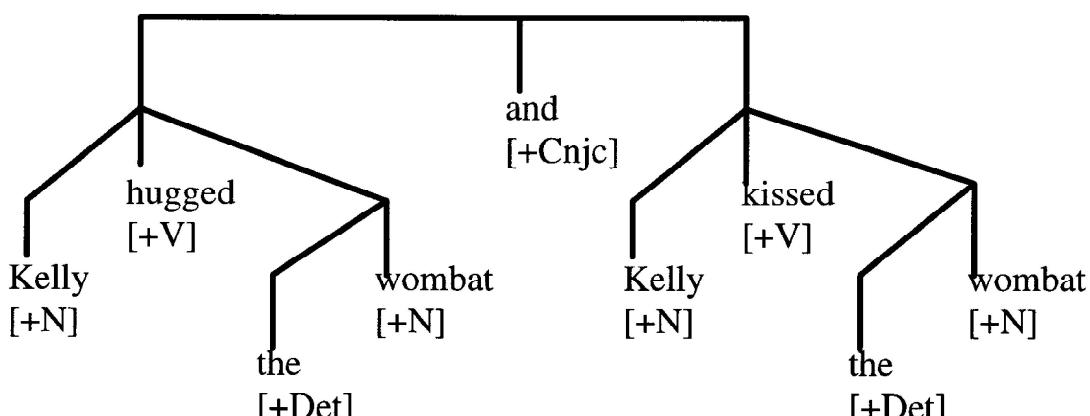
26'') Subject omitted; forward reduction ($W \rightarrow Z$)



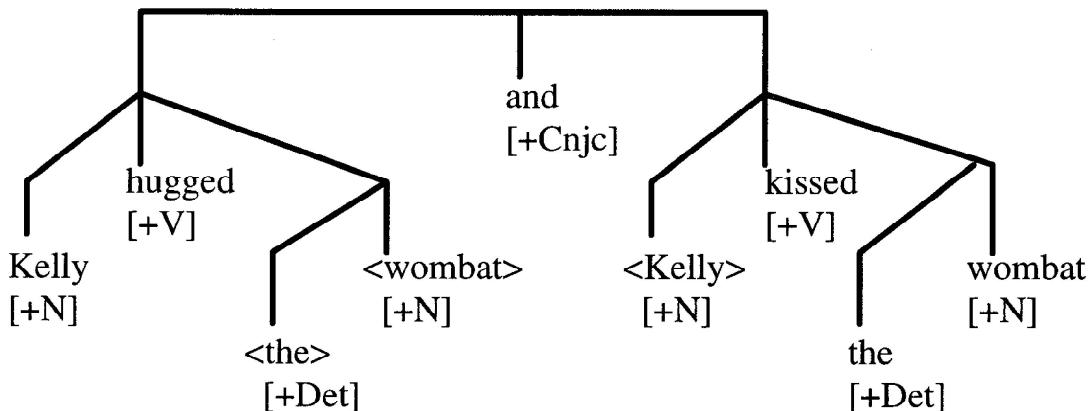
27")

27") 'Direct object' omitted; backward reduction ($Z \rightarrow X$)

28")



28'') 'Subject' and 'object' reduced; forward ($W \rightarrow Y$) and backward ($Z \rightarrow X$) reduction



One of Hudson's examples is of a somewhat different type from the examples given above:

29) (1a) (Hudson 1990:84)

Fred {[is writing an article] and [has nearly finished]}

According to Hudson's analysis, *Fred* and *is writing an article and has nearly finished* are apparently the two immediate constituents of the sentence, and neither of the finite verbs *is* and *has* have a dependent nominative NP, thus violating a strong rule of English syntax. In the corresponding lexicase analysis, the example is not and cannot be so analyzed without violating metatheoretical constraints. There is thus a difference in expressive power between the two dependency sub-theories, and a difference in the analyses they allow which can in principle be investigated empirically. However, the lexicase analysis (29') and 29'') is also not completely straightforward. That is, we seem to have two cases of forward reduction, one of the $W \rightarrow Y$ type and a new type, $X \rightarrow Z$. Before proposing a new strategy for this one, we need to determine how general it is.

B. Gapping

The phenomenon. In the examples of coordinate structure reduction given so far, dependents of coordinated regents are omitted under identity of form and grammatical relation to the regent. However, it is sometimes possible to omit one of the conjoined regents as well. When the regents in question are predicates, this has been referred to in the transformational literature as GAPPING.

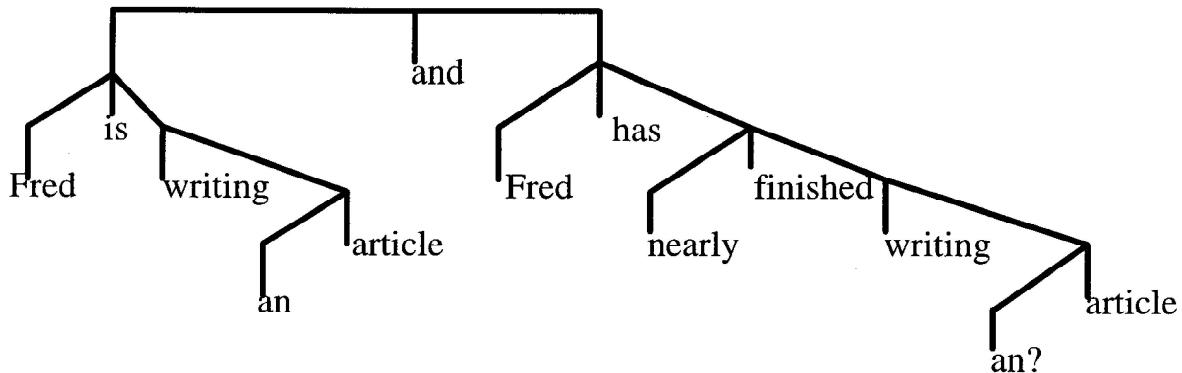
Gapping has been notoriously difficult to handle transformationally because the elements omitted are often not constituents, and because the requirements on identity of form may supercede requirements on lexical and grammatical identity. We believe that a processing analysis stated within a dependency approach promises solutions to both these difficulties. In the present paper, space limitations will only allow us to outline the proposed solution.¹⁰

As in the case of the analysis of missing dependents in the preceding section, a gapping analysis requires that missing elements be identical in shape, and that the regent-dependent

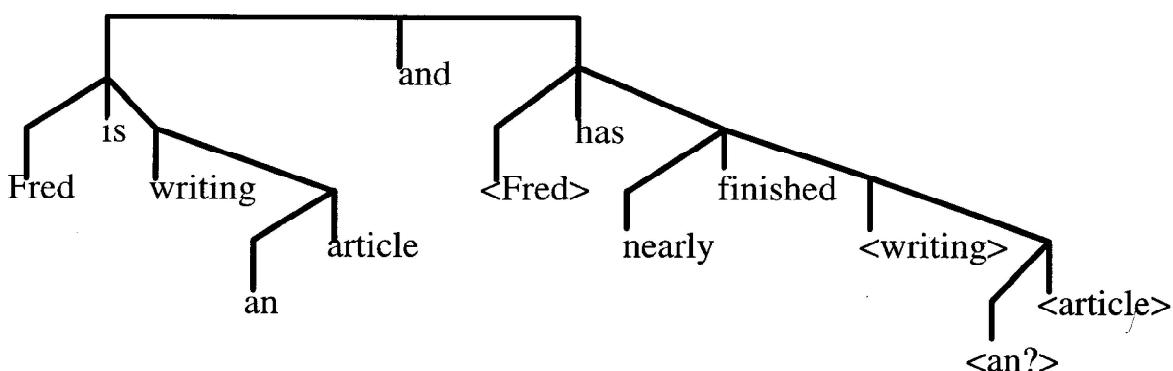
¹⁰The constraints listed below evolved from a set of lexicase reformulations of Word Grammar constraints proposed in Hudson 1989.

relationships in the conjuncts be identical. In addition, the six constraints below must be met in order for word sequences to be gapped.

29')



29'')



- The highest gap in a gapping construction must be a predicate:
- Every gap must depend directly on another gap, or be a predicate depending on a conjunction. That is, no gap can depend on a REMNANT in the sense of Hudson 1989:57, where a remnant is any direct or indirect overt dependent of the gapped predicate.
- The first conjunct can't contain a gap [English and other SVO or V-initial languages]
- The object of a preposition can't be gapped. Apparent counterexamples are cases of reanalysis of prepositions as adverbs ('particles'), and such reanalysis can apply only to complements, not to adjuncts:
- There cannot be two gaps, both of which are flanked by remnants. Or maybe, 'there cannot be more than two remnants'.
- A suppressed word (a word bracketed by <>s in a lexicase stemma) cannot serve as an antecedent for a gap.

Illustrations of these constraints were given in a discussion on the dg-net in 1994, and are omitted here because of length considerations.

IV. JAPANESE

One of the things that inspired us to write this paper was a suggestion by Richard Hudson that Japanese might not have coordination at all.

'Different languages certainly produce different phenomena; e.g. some lgs are full of clitics, others have very few; some lgs have very fully developed coordination systems (e.g. English, I'd say), others hardly seem to have coordination at all (e.g. Japanese?).'
(Richard Hudson, dg-net, 17 October 1996)

Our comparison of English coordinate constructions with their Japanese translations in the first part of the paper suggests that this is at least partly right. That is, while we found clear cases of NP coordination, we did not find clause-level coordination. Examples 30)- 33) further illustrate phrase-level coordinate constructions.

30) Nihon to Eikoku wa simaguni da.
Japan and England Tpc island country be
 'Japan and Great Britain are island nations.'

31) Taroo wa Hanako to Mitiko ni seppun sita.
Taro Tpc Hanako and Michiko in kiss do
 'Taroo kissed Hanako and Mitiko.'

32) Umi to yama ni asobu.
sea and mountain in amuse self
 '(One) takes vacation by the sea and the mountain.'

33) 3 to 5 de 8 ni naru.
3 and 5 at 8 in become
 '3 and 5 makes 8.'

In these four examples the two words, X and Y, are linked by the word *to*: 'X to Y'. As expected, the sentences obtained by exchanging X and Y, 30a) - 33a) are also grammatical.

30a) Eikoku to Nihon wa simaguni da.
 'Great Britain and Japan are island nations.'

31a) Taroo wa Mitiko to Hanako ni seppun sita.
 'Taroo kissed Michiko and Hanako.'

32a) Yama to umi ni asobu.
 '(One) takes vacation in the mountain and at the sea.'

33a) 5 to 3 de 8 ni naru.
 '5 and 3 makes 8.'

In this respect, the exocentric constructions 30) - 33) share the same structural properties with the exocentric construction examined in 34) and 35). Moreover, extraction and placement of the N-*to* sequences in 30) - 33 immediately before the predicate or the regent verb yield the ungrammatical constructions 30b) - 33b).

30b) *Eikoku wa Nihon to simaguni da.

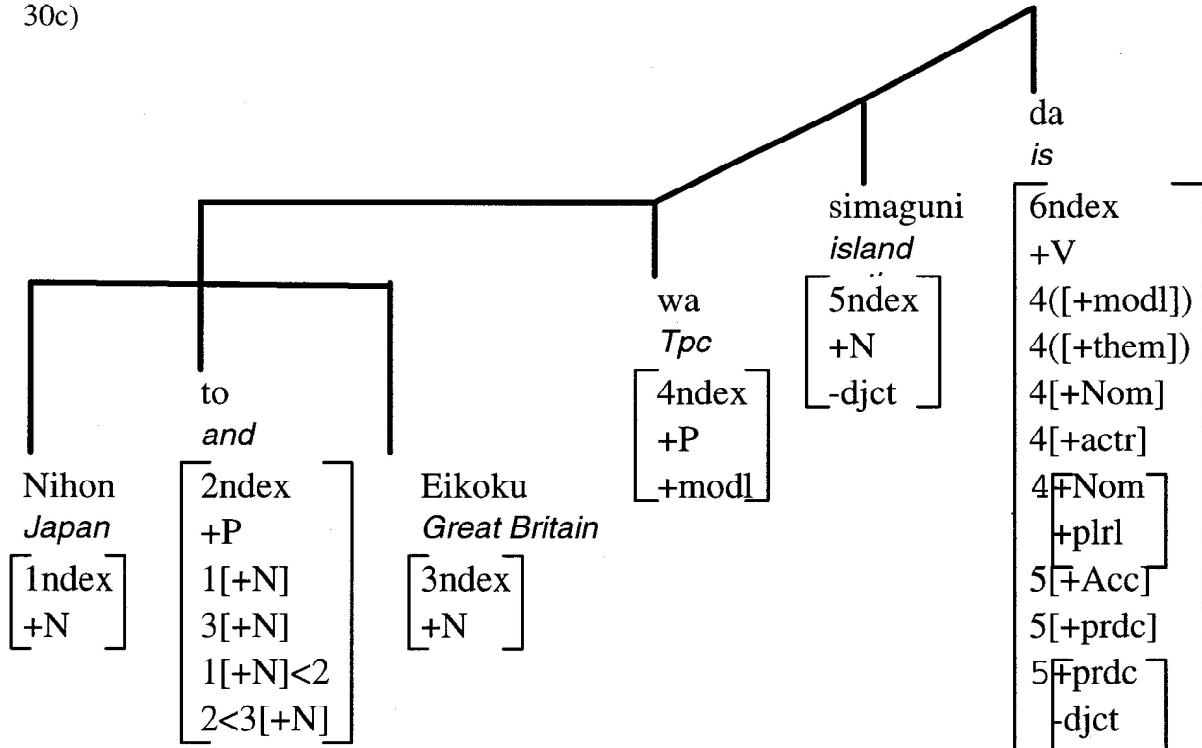
30b) *Taroo wa Mitiko ni Hanako to seppunsita.

32b) *Yama ni umi to asobu.

33b) *5 de 3 to 8 ni naru.

This indicates that the *N-to* sequences in 30) - 33) are not separate constituents bearing their own case relations to the regent verb. We conclude that the examples in 30) - 33) are cases of true coordination, and constitute a partial refutation of Hudson's suggestion that Japanese lacks coordination. In the following section, we present lexicase analyses of examples 30) - 33), in each case treating the *N to N* sequence as a single exocentric construction.

30c)



'Japan and Great Britain are island nations.'

In 30c), the postposition *to* bears the case form and case relation for its obligatory N dependents, *Nihon* and *Eikoku*. Since lexical heads of exocentric constructions are 'transparent' to requirement imposed by their regent, the PAT of this construction is interpreted as borne by both conjunct nouns.

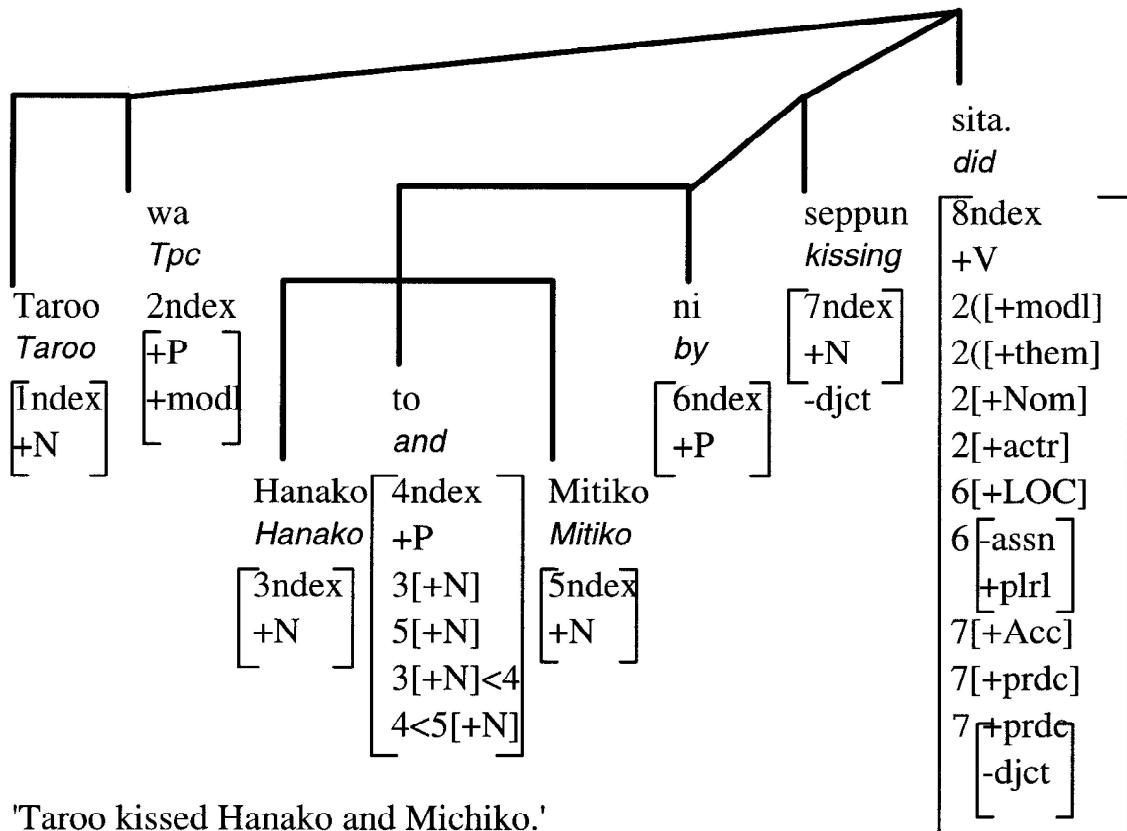
Similarly, in 31c), the postposition *to* bears the case form and case relation for its obligatory N dependents, *Hanako* and *Mitiko*. Since lexical heads of exocentric constructions are 'transparent' to interpretations imposed by their regent, the LOC of this construction is interpreted as applying to both *Hanako* and *Mitiko*. In example 32c), the PAT of the construction must be supplied depending on the larger context in which this sentence occurs. The index for an actant derived from the discourse context is the index *x* marked on contextual features in the matrix of the regent verb, *asobu*. In 33c), the construction lacks PAT case relations. The regent verb is an impersonal verb, marked in the matrix of the regent verb as [+mprs].

One factor that clouds the picture is the fact that the coordination conjunction *to* is homophonous with a 'comitative' postposition *to* and a finite 'complementizer' postposition *to*, and it is not always a simple matter to distinguish between the first two. As examples, consider 34) and 35).

34) Hanako to Taroo ga arasoimasita.
'Hanako and Taroo competed.'

35) Hanako to Taroo ga kimasita.
'Hanako and Taroo came.'

31a)



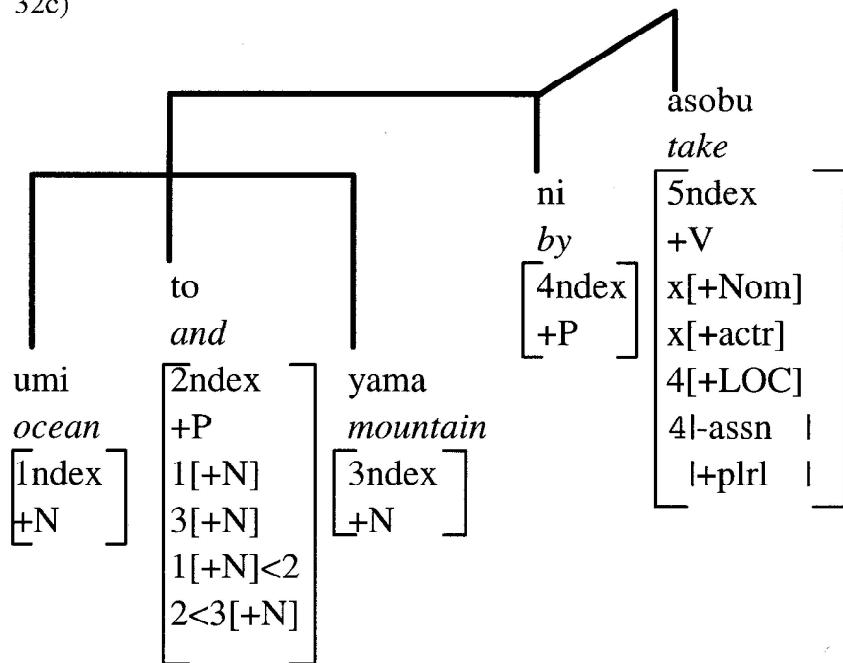
'Taroo kissed Hanako and Michiko.'

Although these constructions have coordinate English translations, they are not to be analyzed as coordinates in Japanese. One test for coordination is that the two conjuncts can be permuted without affecting grammaticality, and this appears to be true of the two examples above. Thus we find that transposing the two NP's 'Hanako' and 'Taroo' in 34) and 35) yields grammatical sentences 34a) and 35a).

34a) Taroo to Hanako ga arasoimasita.
35a) Taroo to Hanako ga kimasita.

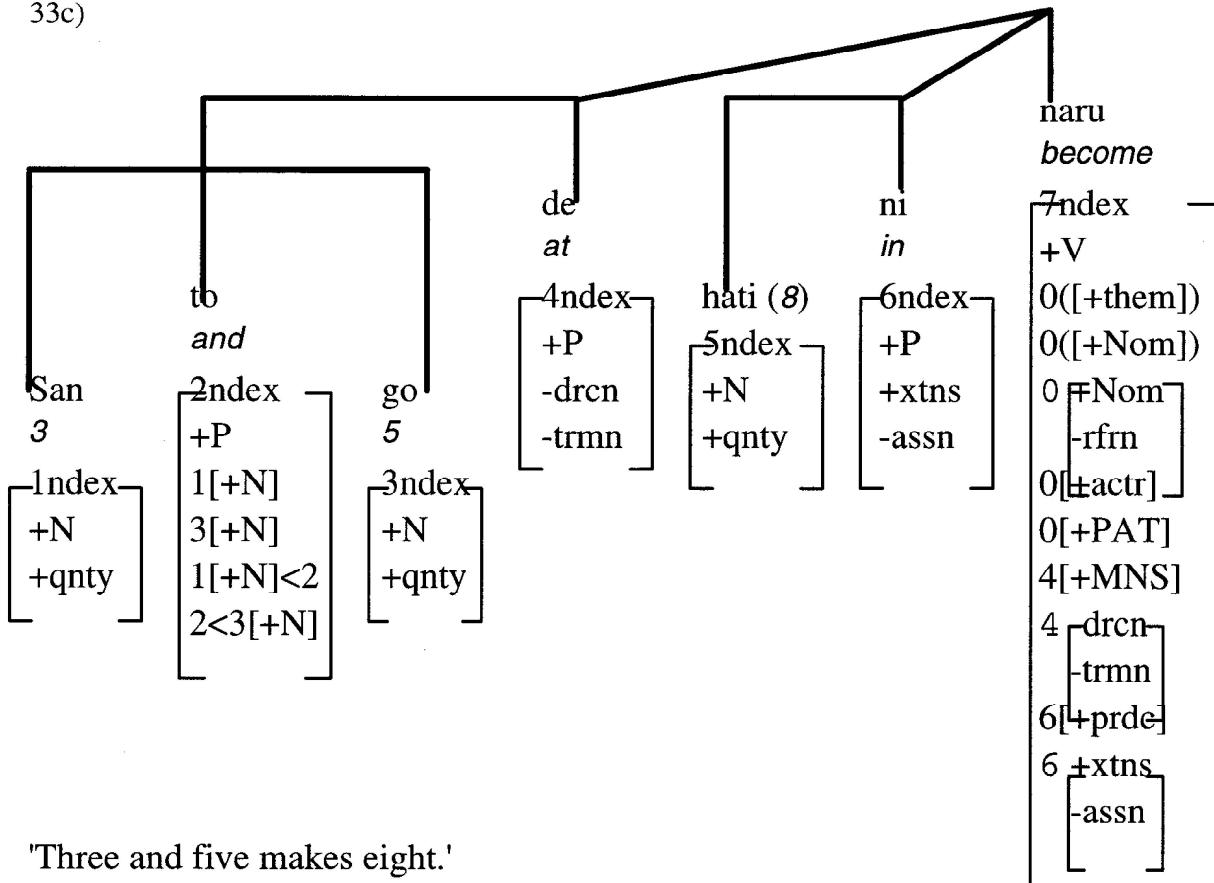
This is consistent with a claim that *Hanako to Taroo* in 34) and 35) is a coordinate construction headed by Cnjc *to* (cf. Watanabe 1971, p.230-231). However, this analysis is not tenable as the sentences 34b) and 35b) below are both grammatical. They illustrate that a

32c)



'One takes vacation by the ocean and the mountain.'

33c)



'Three and five makes eight.'

part of the conjoined exocentric construction, *Hanako to* in 34) and 35) may be extracted and placed immediately before the regent verbs *arasoimasita* and *kimasita* respectively.

34b) Taroo ga Hanako to arasoimasita.

35b) Taroo ga Hanako to kimasita.

It should not be possible to extract a conjunct plus its following conjunction from a coordinate construction. Rather, the standard inference that can be drawn from this kind of invertibility is that *Taroo ga* and *Hanako to* in these two examples are separate and mutually independent dependents of their regent verbs.

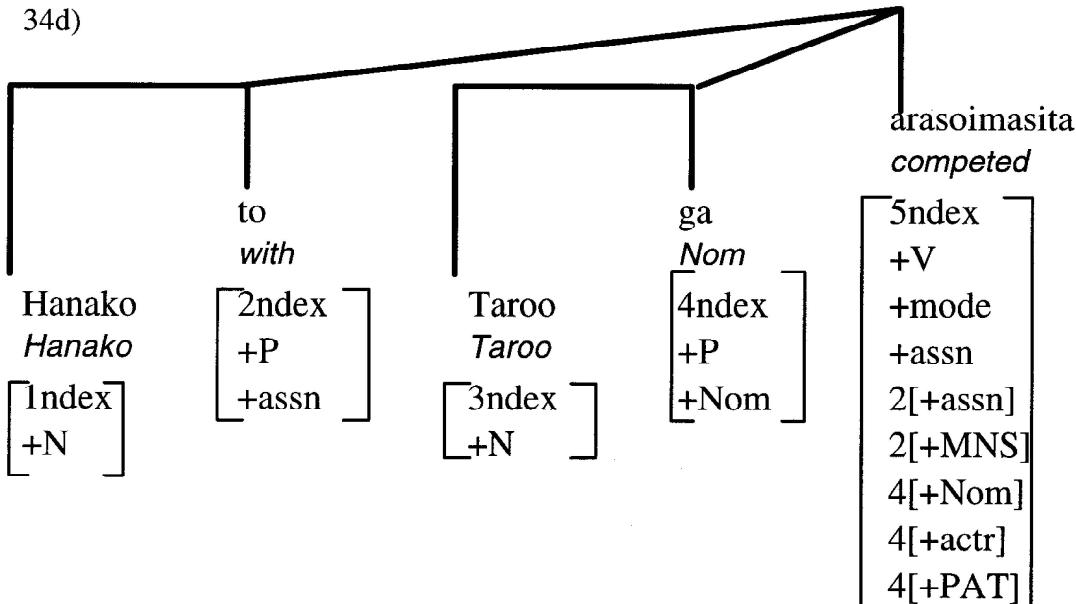
The grammatical status of *Hanako to* in the two examples is also different. Thus, the omission of the phrase *Hanako to* from 34) gives an incomplete and fragmentary sentence 34c), while the omission of the same sequence from 35) yields a grammatical sentence 35c).

34c) *Taroo ga arasoimasita.
‘Taroo fought (with).’

35c) Taroo ga kimasita.
‘Taroo came.’

We may conclude from this that the construction *Hanako to* is a complement of the regent verb *arasoimasita* in 34) while the same construction is an adjunct to the regent verb *kimasita* in 35) (cf. Springer 1993a). The lexicase dependency analysis for 34) and 35) are given below as 34d) and 35d). The difference between complements and adjuncts is shown in this system as a lexical difference, with complements indicated by obligatory contextual features, here ?[+assn], ?[+MNS], and adjuncts indicated by optional contextual features, here ?([+assn]), ?([+MNS]):

34d)



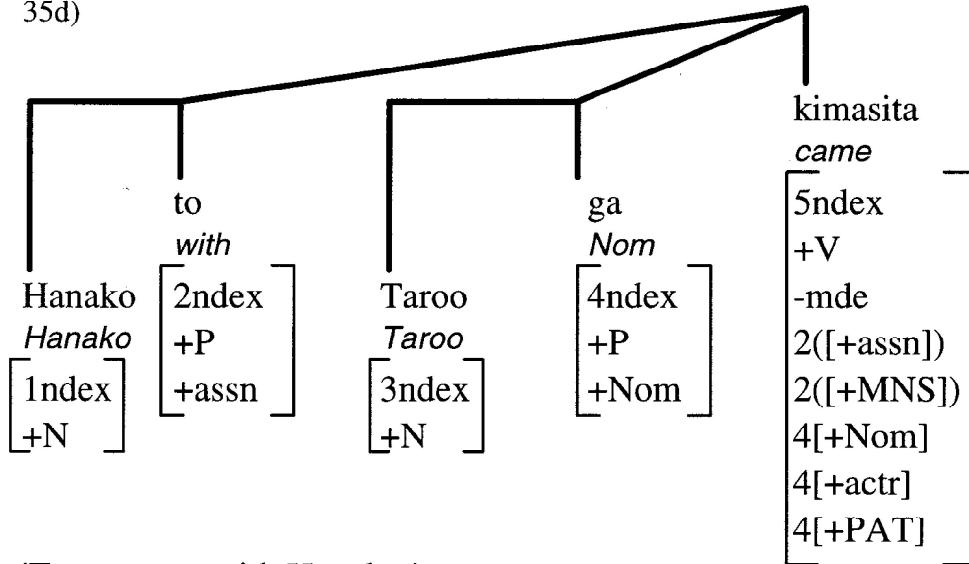
'Taroo competed with Hanako.'

Hudson's tentative claim then turns out to be correct for such examples, in that the putative coordinate NPs in examples 34) and 35) do not form constituents at all, but rather represent sequences of two distinct constructions with the postposition *to* as the head of the PP.

The examples above illustrate putative phrase-level coordination which turns out not to be. Examples like 36) below, like 16) - 22) in the previous section, also support Hudson's suggestion in that they indicate that Japanese may not have coordination at the clause level.

36) Taroo ga ukurere o hiite/hiki, Hanako ga Aroha Oe o utaimasita.
'Taroo played the ukulele and Hanako sang Aloha Oe.'¹¹

35d)



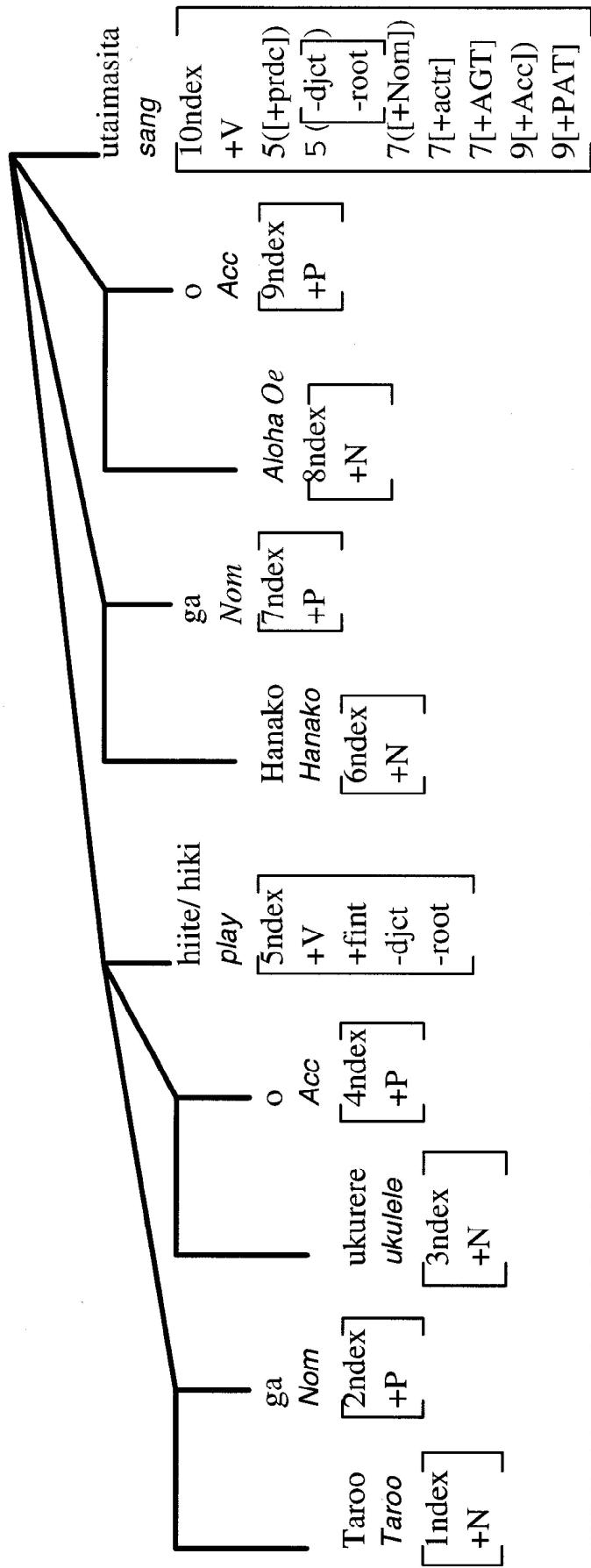
'Taroo came with Hanako.'

The particular morphological forms of the verb for 'to play (string instruments)' in Japanese, *hiite* and *hiki* in 36) both lack tense. The latter form is older and found more frequently in literary works, while the former form is newer and more prevalent in conversation (cf. Watanabe 1971, p. 247; Suzuki 1972, p. 333). In our dependency analysis, we mark these forms as [+V, +fint, -root]. The respective tensed form of the verbs, *hiku* or *hikimasu* for non-past, and *hiita* or *hikimasita* for past tense, cannot appear in the same position, while the final verb form cannot be in the [-root] form without producing a fragment interpretation.. Thus the two supposed conjunct can not be permuted. which counts against a coordination analysis.

In 36) we analyze the sequence that begins with *Taroo* and ends with the [-root] form of the verb 'to play (string instruments)' as an adjunct to the regent verb *utaimasita*. This adjunct is licensed in the matrix of the regent verb *utaimasita* with 5index. The parenthetical enclosures represent the adjuncthood. A structurally more appropriate translation for 36), therefore, is 'Taroo playing the ukulele, Hanako sang Aloha Oe'.

¹¹ *Aloha Oe* is a song composed by Queen Liliuokalani, the last monarch of the Kingdom of Hawaii.

36a)



'While Taroo played the ukulele, Hanako sang Aloha Oe.'

V. CONCLUSION

In the preceding section (Section IV), we examined in detail Hudson's observation that Japanese may not have coordination at all. We argued that there are phrase-level Japanese constructions that warrant a coordination analysis, but that at the clause level, Hudson's suggestion is borne out. More generally, we demonstrated that we can analyze well-formed coordinate constructions in the lexicase dependency framework without relaxing any constraints, and provided preliminary evidence that reduced and gapped constructions can be accounted for as fragments of full coordinate constructions related to their fully grammatical counterparts by statable processing strategies. Bottom line: dependency can do coordination.

VI. REFERENCES

Chomsky, N.A. (1965). *Aspects of the theory of syntax*. The M.I.T. Press, Cambridge, Massachusetts.

Hudson, R.A. (1989). Gapping and grammatical relations. *Journal of Linguistics* 25:57-94.

Hudson, R.A. (1990). *English Word Grammar*. Oxford: Basil Blackwell

Lobin, Hening. (1993). *Koordinationssyntax als prozedurales Phänomen*. Studien zur deutschen Grammatik 46. Gunter Narr, Tübingen.

Mel'čuk, Igor A. (1988). *Dependency syntax: theory and practice*. SUNY series in linguistics. State University of New York Press, Albany.

Milward, D. (1994) Non-Constituent Coordination: Theory and Practice. In: *Proceedings of COLING 94*, 935-941, Kyoto.

Milward, D. (1995?) Dynamic Dependency Grammar. To appear in *Linguistics and Philosophy: Special Issue on Mathematics of Language*.

Sgall, P., J. Panevová. (1989). Dependency syntax - a challenge. *Theoretical linguistics* 15, No. 1/2, 73-86.

Springer, H.K. (1993a). *Perspective Shifting Constructions in Japanese*. Ph.D. Dissertation. University of Hawaii, Honolulu.

Springer, H.K. (1993b). Japanese topicalization: a lexicalist dependency analysis. In *Pan-Asian linguistics: Proceedings of the Third International Symposium on Language and Linguistics, Volume III*. (Sudaporn Luksaneeyanawin (Ed)), 1191-1206. Chulalongkorn University Press, Bangkok.

Starosta, S. (1988). The Case for Lexicase. Pinter Publishers, London.

Suzuki, S. (1972). *Nihon Bunpo Keitairon*. Mugi Shobo, Tokyo.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Librairie C. Klincksieck.

Watanabe, M. (1971). *Kokugo Kobunron*. Tokyo: Haniwa Shobo.