



## Kobe Shoin Women's University Repository

Title	On the Conservatism of Embedded Clauses
Author(s)	松田 謙次郎 (Kenjiro Matsuda)
<i>Citation</i>	Theoretical and applied linguistics at Kobe Shoin, No.1 : 1-13
Issue Date	1998
Resource Type	Bulletin Paper / 紀要論文
Resource Version	
URL	
Right	
Additional Information	

# On the Conservatism of Embedded Clauses<sup>1</sup>

Kenjirô Matsuda

---

Embedded clauses have been claimed to show syntactic conservatism against incoming linguistic change (Givón 1979, Hock 1986, Matsuda 1993). While all of the cases show the conservative nature of dependent clauses whether it is a past change or a change in progress, it is not clear what causes such a tendency. Three possible explanations — syntactic (Emonds 1970, Ross 1973), speech stylistic, discourse-pragmatic (Givón 1979, Hooper and Thompson 1973), and processing-based (Townsend, Ottaviano and Bever 1979, Klender and Kutas 1993) — are proposed and closely examined. It is then suggested that the most reasonable scenario is that the conservatism is caused by both discourse-pragmatic and processing-based factors.

## 1. Introduction

An interesting property of embedded clauses across languages is their apparent “conservatism” in grammatical changes. There have been numerous reports on cases where a historical change is found to have lagged behind in such clauses in comparison with main clauses (Hock (1986), Givón (1979), Matsuda (1993,

---

<sup>1</sup>This is a paper read at the XIIIth International Conference on Historical Linguistics Düsseldorf, Germany, Aug. 10–17th 1997. I am indebted to John Haig and Michael Wescoat for their comments on the earlier version of this paper, and to Susan Herring, Jack Hoeksema, Fusa Katada and Selikoko Mfwene for their generous comments at the presentation. Usual disclaimer applies here.

1995), among others). The point is best summarized in the following quote from Hock (1986), where he remarks on a case of word order change in Old English:

In the development outlined above, dependent clauses tended to lag behind [...]. It took several additional centuries before the word order of the main clause was obligatorily extended to dependent clauses. This relative **resistance to change** is consonant with a widely noted tendency for dependent clauses to be more ‘conservative’ than main clauses in syntactic change. (There is less agreement concerning the reasons for that conservatism).

(Hock 1986: 332, emphasis in original)

A typical example of such conservatism is found in German word order, where the conservative SOV sequence is preserved in embedded clauses while an innovative Verb-second order is found in main clauses.

As Hock mentions above, there have also been attempts to explain this peculiarities from different points of view. In this paper, I will examine four possible theories that may look plausible as accounts of the conservative nature of embedded clauses: syntactic, stylistic, discourse-pragmatic and processing-based explanations. Then I will show that in view of available evidence, the conservatism is best ascribed to discourse-pragmatic and processing-related causes.

## 2. Review of Preceding Data

First of all, a very brief survey of past studies is in order. Hock (1986) considers word order change in Old English, which is responsible for the current English SVO order. From what used to be a *main verb + auxiliary verb* order in the sentence final position in early Germanic developed an AUX-second order in *Beowulf*. Then, the main verb was attached next to the AUX, making an AUX-main verb order in the second position. As he mentions, dependent clauses showed a clear conservative trend with respect to this change.

Givón (1979) cites several cases of word order change. The first one is the retention of VS syntax in many subordinate clauses in Biblical Hebrew contrasting with the innovative SV syntax main clauses. A similar example is found in Kru, where the conservative SOV order survived in infinitival complements and related modal embedded environments. The innovative order, SVO, is found in all other clause types.

Another type of conservatism that Givón discusses is the frozen nature of infinitival verb complements. Such VP compounds as *street cleaner*, *garbage disposal*, or *man-hunt* result from the freezing effect of nominalization which

preserved the older OV order. In Yoruba, an old OV order survived in cleft-predicate phrases, while the main clause predicate now is VO. The same freezing effect is also observed in Bantu WH-questions and cleft constructions, in which the older copula *ni* is still found.

The cases reviewed so far consist of historically documented changes in the past. A more vivid picture can be obtained from sociolinguistic studies of language change in progress. One such study concerns an ongoing analogical leveling of inflectional suffixes conveying the meaning of potentiality in Tokyo Japanese (Matsuda 1993). The process, whereby a conservative suffix *-are* is replaced by an innovative one *-e* (e.g. *mi-r-are* vs. *mi-r-e* 'can see') *tabe-r-are* vs. *tabe-r-e* 'can eat') is found to be more advanced in main clauses than in embedded clauses in natural speech data ( $p < .008$ ):<sup>2</sup>

(1) **Independent/Main clause:**

Atasi-mo [kimono]            ki   -r -e r -u   -yo, hitori -de  
I too            [kimono (ACC)] wear POT   PRES FIN by myself  
'I can put on a kimono by myself, too' [KM/FYD/8806-0-250]

(2) **Semi-embedded clause:**

Okyoo-ga age   -r -are -reba oyome-ni ik   -are -ta   -no  
sutra ACC recite POT   PROV bride   to go POT PAST FIN  
'If you could only recite a sutra, you could get married' [HIS/FOD/8602-1-670]

(3) **Embedded clause (Noun complement clause):**

Yappari itiban tadasi-ku iki   -r -are -ru   -no   -wa  
after all most righteously live POT   PRES COMP TOP  
saibankan -kana  
a judge   FIN  
'After all, it is a judge who can live in the most righteous manner' [KM/MYU/8931-B-257]

Thus, here we have evidence from synchronic data on ongoing change that suggests conservatism, and also we see here that conservatism can also affect a morphological change in progress, in addition to the syntactic changes that have constituted the bulk of the cases examined in the literature to date. Another sociolinguistic study of linguistic variation concerns the accusative case marker

<sup>2</sup>Examples from speakers are identified with their initials, their social background (encoded with three attributes: Male/Female, Young/Old, Uptown/Downtown), and the tape-side-counter number.

	<i>Probability</i>	<i>% Zero</i>	<i>N</i>	<i>Significance</i>
<i>Main Clause</i>	0.52	60	2,270	N.S.
<i>Semi-Embedded Clause</i>	0.51	56	3,222	$p < 0.001$
<i>Embedded Clause</i>	0.43	48	2,037	
<i>TOTAL</i>			7,529	

Table 1: The Effect of Embeddedness on Zero-Marking of *-o* (Matsuda 1995: 215)

in Tokyo Japanese, which alternates between an overt form (*-o*) and a zero form in natural speech (Matsuda 1995: 214, 216-7):

- (4) Are -  $\emptyset$  itizikan kiiten -no -yo  
that ACC one hour listen FP FP  
‘I listened to it for an hour, you know’ [IH/FOU/9119-1-694]
- (5) Hutari-site karee raisu -o tabe -te -sa  
two of them curry and rice ACC eat GER FP  
‘They ate curry and rice alone, you know’ [TY/YMD/8839-0-058]

Under the assumption that the zero form is introduced by a rule deleting the case marker *-o* attached to the object NP, Matsuda checks whether the type of the clause in which the object NP appears makes a difference in the rate of occurrence of the zero-form. When the clauses were divided into three categories by their respective degree of embedding — main, semi-embedded (including gerundive and temporal clauses) and embedded (which includes multiple embeddings as well) — there was a statistically significant difference in their relation to the zero-form, as seen in Table 1.

That is, as expected, application of the rule is suppressed to a significantly greater degree in embedded clauses than in main or semi-embedded clauses. Here again, the embedded clauses function as a conservative environment for the zero-marking.

### 3. Some Possible Explanations and Their Inadequacies

Given the evidence above, one may well ask why embedded clauses exhibit such a property. Here I will give a short overview of the major proposals so far.

#### 3.1 Syntactic Explanation

From a synchronic syntactic point of view, the diachronic conservatism of embedded clauses is ascribed to the relative difficulty in applying syntactic rules in

those clauses. Emonds (1970) and Ross (1973) offer two major proposals that attempt to capture this generalization in syntactic terms. Emonds (1970) begins by making a major distinction among transformations between Root Transformations and Structure-Preserving rules. Covering such rules as VP-preposing (*Mary plans for John to marry her, and marry her he will*), Preposing around be (*More significant is the development of a semantic theory*), Topicalization (*This book you should read*) and Right Dislocation (*You should go to see it, that movie*), Root Transformations apply to the highest S or an S immediately dominated by the highest S. Thus, he predicts that these rules do not apply in embedded clauses.

For Ross, it is the Penthouse Principle, formulated as follows (Ross 1973: 411), that captures the systematic differentiation of main and subordinate clauses with respect to permissible transformational rules:

**(6) The Penthouse Principle:**

All syntactic processes that apply in prenominal clauses also apply in other subordinate clauses. All syntactic processes that apply in subordinate clauses also apply in main clauses. [...]

Properly translated into variationist terms, both Ross's and Emonds' frameworks may indeed predict the observed conservatism. As one can easily see, however, their main problem lies in the fact that they are simply formal expressions of a descriptive generalization, and as such, they do not explain *why* subordinate clauses permit fewer transformational rules than main clauses do. In other words, as historical linguists or variationists, one would like to know how such a principle or rule-division is motivated in explanatory terms.

Naturally, the best syntactic explanation here would predict the conservative effect as a result of an interaction of the transformational rules with other independently motivated principles, rules, or conditions. But then, one would face the vast diversity of syntactic processes we saw above where such conservatism was observed; it is simply too hard to come up with any syntactic mechanism or a structural feature that is responsible for word order, inflectional morphology and case marking. The only common denominator for these cases is that they are all from embedded clauses, and nothing more. Hence, one can only offer the descriptive formulation reviewed above.

Another problem with the syntactic approaches are the counterexamples with which other researchers have raised against both Ross's and Emonds' proposals. Hooper and Thompson (1973) make the important observation that Root Transformations do apply in embedded clauses, and claim that what is really behind the regularity is to be captured by a semantic/pragmatic prin-

ciple, rather than a syntactic one as Emonds (1970) proposes.<sup>3</sup> Against the Penthouse Principle, Andersson and Dahl (1973) observed that in Swedish the auxiliary *ha* 'have' can be optionally deleted only in embedded clauses and not in main clauses, a case that counts as a legitimate counterexample to Ross's proposal.

It does not, then, seem promising to look for the cause of the conservatism in syntactic structures; the approach is at best descriptive, and even then there are a number of counterexamples that cannot be captured in syntactic terms.

### 3.2 A Discourse-Pragmatic Explanation

Surveying a wide range of sentence types and their functions, Hooper and Thompson demonstrate that Root Transformations actually can apply to certain embedded clauses, and all those clauses share the characteristics of being asserted. On the other hand, those transformations are not applicable to presupposed clauses, "because it is not appropriate to emphasize elements of a sentence whose proposition is already known, whose truth is presupposed, and whose content is relegated to the background." (Hooper and Thompson 1973: 495). By extending their theory, the diachronic projection of the synchronic conservatism can be traced back to a discourse-pragmatic cause.

The same point was made most clearly by Givón (1979) who posited a direct connection between synchrony and diachrony via discourse-pragmatic motivation. His theory is built upon the notion of discourse presupposition, namely "the degree of presupposed background upon which a sentence is used." (Givón 1979: 48-9). Measured against this criterion, main, declarative, affirmative, active sentences have the lowest presuppositional complexity. According to his theory, any syntactic variants can be graded according to this criterion, and more presuppositionally complex variants exhibit the following four major characteristics: (1) syntactic complexity, (2) distributional restrictions, (3) later grammaticalization and (4) syntactic conservatism. Here, then, syntactic conservatism is thus directly connected to a discourse-pragmatic factor.

However, a discourse-pragmatics-based explanation faces a difficulty wherever a given linguistic change/variation does not have any apparent discourse-pragmatic function but still exhibits a similar tendency. One such example is the innovative potential case mentioned above. Since this variation is purely

<sup>3</sup>Much later, Baltin (1982) noted the following examples that show that topicalization is not entirely impossible in embedded clauses (Baltin 1982: 17):

- (a) He's a man to whom liberty we could never grant \_\_\_\_.
- (b) It's obvious that Mary, he can't stand \_\_\_\_.

inflectional in nature, with no functional pragmatic difference involved, it is hard to see how Hooper and Thompson's theory could explain the observed conservatism. When we turn to Givón's version of the theory, we do see that morphological innovation is covered as the freezing phenomenon in nominalized or frozen clauses. However, as he admits, since "[T]he innovation of grammatical morphology is not [...] constrained by considerations of discourse markedness" (Givón 1979: 85), it is no longer clear whether or not we can consider the effect as deriving from a discourse-pragmatic factor. The same holds true with the accusative case marking variation, where no pragmatic factors were found significant in the multivariate analysis of natural data, and neither variant seems to carry any obvious pragmatic meaning in discourse (Matsuda 1995). Given these limitations, the discourse-pragmatic approach does not seem to be the best explanation of conservatism in subordinate clauses.

### 3.3 Stylistic Explanation

Speech style may look like a promising candidate for the factor responsible for subordinate-clause conservatism. Since embedded clauses are more likely to be found in formal or careful contexts than in casual ones, it may not be surprising to find that the observed difference by clause type is actually a product of a difference in speech style. According to this theory, then, there is no such thing as "conservatism of embedded clauses"; what remains is the simple classical sociolinguistic fact that people tend to use conservative forms in careful speech. In other words, the conservatism would be just an artifact of stylistic alternation.

Given that speech style is a possible confounding factor for embedded clause conservatism, what is needed here is a controlled comparison of main and subordinate clauses with respect to the variable. That is, we need to look at data containing sizeable syntactic variation, code each sentence containing the linguistic variable by style and clause type, and check the result closely to see which of the two factors is the true determinant of the variation. Fortunately, my own data on accusative case marking variation in Tokyo Japanese is of precisely this type (Matsuda 1995). The corpus used there is made up of 7,529 tokens of sentences containing the overt *-o* and the zero marker, and it is coded for a number of linguistic and extralinguistic factors including the clause type and the speech style. The original coding for speech style is done according to the scheme developed by Labov and Sankoff (Labov and Sankoff 1985 — see Figure 1), but in what follows, I simply divide tokens into casual and careful styles.

Figure 2 displays the distribution of three types of clauses: main, semi-embedded and embedded clauses. As expected, the relative frequencies of main



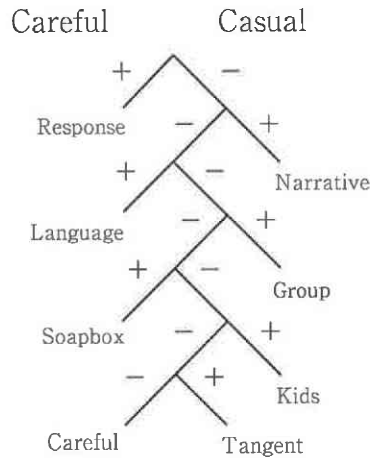


Figure 1: Style Taxonomy Tree by Labov and Sankoff (1985)

and embedded clauses are reversed in the two styles, so that there are more main clauses than embedded ones in casual speech, while the reverse holds for careful speech. The difference observed here is highly significant at the  $p < .0001$  level according to the chi-square test. Then, it is indeed true that one finds more embedded clauses in careful speech. But does it then follow that the style causes the conservatism?

In Figure 3, we see a graphic representation of the interaction of speech style and clause type, with reference to the case-marking alternation. It most eloquently shows that even when the confounding factor, speech style, was controlled, the effect of embedding is still present in the expected direction. Thus, even though the ratio of main/embedded clauses varies considerably by speech style, the conservative effect of the embedded clauses stays the same. Consequently, the style-artifact theory is empirically falsified.<sup>4</sup>

### 3.4 A Processing-based Explanation

As noted several times in this paper, diachronic conservatism is presumably reflected in a synchronic difference in rule-applicability between main and subordinate clauses. Under the assumption that all diachronic changes arise first

<sup>4</sup>It is worth noting that Hill (1973: 40ff) did a similar examination in her Luiseño and Cupeño study, and found that although the frequency of subordinate clauses differ considerably by style, their decline in use across time was still coherent.

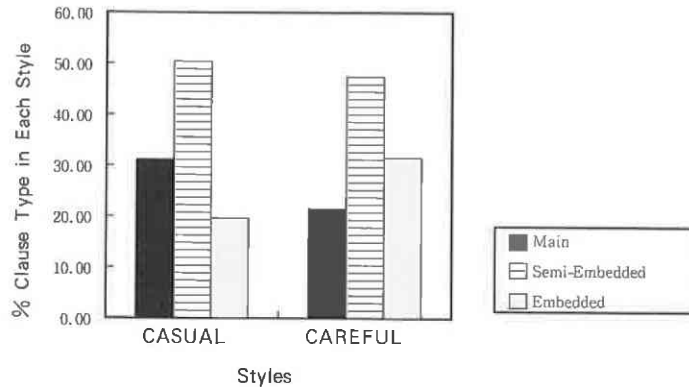


Figure 2: Distribution of Clause Types in Each Style

in synchrony as optional application of linguistic rules, the reasoning here goes that those rules applying in the main clause cause the change earlier there than in embedded clauses, where, through the Penthouse Principle, they do not apply.<sup>5</sup> Assuming that the application of those grammatical rules would cause an increase in processing load, we might also look for a cognitive or processing-related reason for the conservatism.

Indeed, there are a number of psycholinguistic studies that are strongly suggest the cognitive nature of morpho-syntactic conservatism. First of all, there is a widely known fact concerning acquisition: embedded clauses are acquired by children at a much later stage than main clauses are. Then, there's a curious mirror-image fact from language death and aging as documented by Hill (1973) and Kemper (1986). Hill reports a case from Luiseño and Cupeño, where the speakers of the two dying languages are found to use fewer subordinate clauses than they used to do.<sup>6</sup> Kemper's case comes from an experimental result where elderly adults show difficulty in imitating or paraphrasing long sentences, particularly those with sentence-initial embedded clauses, presumably because of memory limitation. These two facts — i.e., that embedded clauses are acquired later and that they are lost earlier — are reminiscent of Jakobson's classical study of phonological acquisition and its loss in aphasia, and they point to em-

<sup>5</sup>Note that it is earlier/later, and not faster/slower, following Kroch's Constant Rate Hypothesis (Kroch 1989).

<sup>6</sup>Hill's case actually makes the term "conservatism of embedded clauses" an indisputable misnomer, since, in those dying languages they are among the first to disappear.

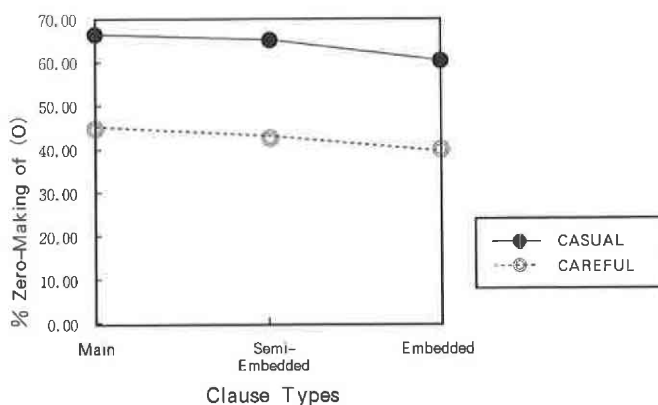


Figure 3: Interaction of Speech Style and Clause Types

bedded clauses as a grammatical constructs which are cognitively difficult for language users.

Experimental psycholinguistics offers more results supporting the processing-based explanation of conservatism. In their report on word-recognition latencies in main and subordinate clauses between adults and children, Townsend, Ottaviano and Bever (1979) advances the hypothesis that the listener interprets a main clause more quickly than a subordinate clause. Kluender and Kutas (1993) employ rich experimental data to attribute Subjacency effects to limits on human sentence processing. Based on a condition whereby “[M]ovement cannot cross more than one bounding node, where bounding nodes are IP and NP” (Haegeman 1991: 365), the Subjacency effect in its essence prohibits any dependency between constituents across deep embedding. Kluender and Kutas convincingly demonstrate that the effect is best understood as resulting from a processing difficulty, instead of being a part of innate universal grammar. All of these psycholinguistic studies provide compelling evidence of the fact that embedded clauses are more difficult to process than main clauses.

The processing-based explanation has definite advantages over the others reviewed so far. First, it obviates the need to look for any structural reason for the differentiated behavior of elements in embedded clauses. Secondly, it correctly predicts that the conservatism is uniformly present over different speech styles. Third, it follows naturally that the effect is present even if the relevant morpho-syntactic rule does not have any obvious discourse-pragmatic function,

simply because the theory does not refer to such factors. Of course, unlike the Penthouse Principle, the notion of Root Transformations or Givón's characterization of frozen clauses, processing-based approach is not merely descriptive but offers an independent explanation of why one would expect fewer rules to be applicable to subordinate clauses.

Does this all mean, then, that the processing-based explanation wins over all the alternatives? There remains at least one fact that is best explained by Hooper and Thompson's discourse-pragmatic theory, rather than by the cognitive approach, and that is the non-isomorphism between the main/subordinate distinction and the pattern of rule-applicability. Hooper and Thompson's examples include not only cases where Root Transformations are applicable in subordinate clauses, but also some where they are inapplicable in main clauses that convey presupposed information (Hooper and Thompson 1973: 493):

(7) \*In came Jerry because it was raining.

(8) \*Sitting in the corner was Tom because he'd hidden grandma's teeth.

Breaking the near-perfect correlation between sentence structure and rule-applicability, their data above effectively suggest the involvement of pragmatic factors and seriously undermine any attempts to solve the issue with a processing-based explanation alone. One must then admit that there are cases that can only be accounted for by means of a discourse-pragmatic-based notion.

#### 4. Conclusion

Through a close examination of possible and actually proposed explanations, we have reached the conclusion that conservatism is most reasonably understood as a product of both discourse-pragmatic and processing-based factors. There is obviously a considerable overlap between the two (Givón 1979), but it seems that one could point to several clues to tell which factor is at work in a given linguistic phenomenon.

First and foremost, there *must* be a clear pragmatic effect involved if the conservatism is to be explained by a discourse-pragmatic factor. The case of Root Transformations addressed by Hooper and Thompson falls into this category, and it is likely that word order changes also belong to this group, so long as some obvious discourse meaning can be isolated for the marked ordering. Change in inflectional suffixes or any other diachronic change devoid of any clear pragmatic differentiation, on the other hand, must be relegated to a processing-based explanation.

Secondly, there seems to be a difference in the degree of conservatism depending on which factor causes it. Looking over Hooper and Thompson's data, one is struck by a clear distinction between the acceptable and non-acceptable sentences. Turning our attention to the case marker variation across three clause-type in Figure 3, we see that the effect is much weaker, if not absent. A paucity of other comparable data deters us from any further meaningful conjectures, but this line of inquiry appears to be a promising avenue for distinguishing the two factors.

Lastly, a note on future study is in order. Among a number of possibilities, here I touch on an examination of the effect of typological differences across languages (Kimball (1975), Babyonyshev and Gibson (1995)). Some embedded clauses may not be so psychologically difficult to process and thus may not be so conservative as others, and the distribution of such differences is expected to correlate closely with typological characteristics. For example, doubly embedded complement clauses in Japanese are known to be more difficult for the speakers of the language to process than the corresponding English sentences. On the other hand, because the head final nature of Japanese, relative clauses in that language are claimed not to cause processing overload, as the corresponding forms do in English (Babyonyshev and Gibson (1995: 4ff.)). The processing-based theory advanced in this paper predicts that these differences should be reflected in language change and variation, but its verification must be left for further investigation.

## References

- Andersson, Anders-Böje and Dahl, Östen. 1974. "Against the Penthouse Principle." *Linguistic Inquiry* 5 (3): 451-3.
- Babyonyshev, Maria and Edward Gibson. 1995. "Processing overload in Japanese." Schüze et al. (eds.) *Papers on language processing and acquisition (MIT working papers in linguistics vol. 26)*, 1-35. Cambridge, MA: MIT Press.
- Baltin, Mark. 1982. "A landing site theory of movement rules." *Linguistic Inquiry* 13 (1): 1-37.
- Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Doctoral Dissertation, MIT.
- Givón, Talmy. 1979. *On understanding grammar*. Orlando: Academic Press.

- Haegeman, Liliane. 1991. *Introduction to government and binding theory*. Oxford: Basil Blackwell.
- Hill, Jane H. 1973. "Subordinate clause density and language function." Claudia Corum et al. (ed.) *You take the high node and I'll take the low node: Papers from the Comparative Syntax Festival*, 33-52. Chicago: Chicago Linguistic Society.
- Hock, Hans Henrich. 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.
- Hooper, Joan and Thompson, Sandra. 1973. "On the applicability of root transformations." *Linguistic Inquiry* 4: 465-497.
- Kimball, John. 1975. "Predictive analysis and Over-the top parsing." In Kimball (ed.) *Syntax and semantics*, vol. 4, 155-179. New York: Academic Press.
- Kluender, Robert and Kutas, Marta. 1993. "Subjacency as a processing phenomenon." *Language and Cognitive Processes* 8 (4): 573-633.
- Kroch, Anthony. 1989. "Reflexes of grammar in patterns of language change." *Language Variation and Change*. 1 (3): 199-244.
- Matsuda, Kenjirô. 1993. "Dissecting analogical leveling quantitatively: The case of the innovative potential suffix in Tôkyô Japanese." *Language Variation and Change* 5 (1): 1-34.
- Matsuda, Kenjirô. 1995. *Variable zero-marking of (o) in Tôkyô Japanese*. Doctoral Dissertation, University of Pennsylvania. [Available as Technical Report IRCS-96-20 from Institute for Research in Cognitive Sciences; [ftp://anonymous@ftp.cis.upenn.edu/pub/ircs/tr/96-20part1, part2, part3](ftp://anonymous@ftp.cis.upenn.edu/pub/ircs/tr/96-20part1,part2,part3)].
- Ross, John Robert. 1973. "The Penthouse Principle and the order of constituents." Claudia Corum et al. (ed.) *You take the high node and I'll take the low node: Papers from the Comparative Syntax Festival*, 397-422. Chicago: Chicago Linguistic Society.
- Townsend, David J., Ottaviano, David and Bever, Thomas G. 1979. "Immediate memory for words from main and subordinate clauses at different age levels." *Journal of Psycholinguistic Research* 8 (1): 83-101.