



PROJECT MUSE®

---

*The syntax of anaphora* (review)

Eric Reuland

Language, Volume 85, Number 1, March 2009, pp. 231-237 (Review)

Published by Linguistic Society of America

DOI: <https://doi.org/10.1353/lan.0.0095>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/262995>

**The syntax of anaphora.** By KEN SAFIR. Oxford: Oxford University Press, 2004. Pp. 328. ISBN 9780195166149. \$60.50.

Reviewed by ERIC REULAND, *Utrecht Institute of Linguistics OTS*<sup>1</sup>

Virtually the only sentence that Jespersen (1933:111) devotes to what we now know as binding theory is a statement that English uses a reflexive for the object when the subject and object are identical. Yet, about half a century later, binding had become one of the central modules of the grammar, governing not only the distribution and interpretation of pronominals and anaphors, but also movement (Chomsky 1981). In the minimalist program (Chomsky 1995 and subsequent work) the integration of binding into the grammar is being pursued in a different way. The whole range of intricate facts about binding discovered since Jespersen should be derivable from core properties of the grammar, rather than be stipulated as part of a separate component. Apart from a definition, the theory should contain no special statements about binding.

The core of the canonical binding theory (CBT) was formed by the following conditions on A-binding: (i) an anaphor is bound in its local domain; (ii) a pronominal is free in its local domain; and (iii) an R-expression is free. Binding was defined in terms of coindexing and c-command (Reinhart 1976), and the local domain was defined as the minimal category containing the bindee, a governor, and a subject—with some restrictions I do not discuss here—hence its technical name of GOVERNING CATEGORY.

A striking property of the CBT is its simplicity, expressing complementarity between anaphors and pronominals in the local domain, approaching a scientific ideal. As empirical investigation progressed, however, many languages turned out to have anaphoric systems that do not fit into the CBT. Starting out with Germanic, Dutch, for instance, has a three-way system with pronominals, simplex anaphors (SE-anaphors), and complex anaphors (SELF-anaphors); the Scandinavian languages have a four-way system in the argument domain (pronominals, SE-anaphors, anaphors of the form <PRON SELF> and of the form <SE SELF>). They all have different distributions. SE-anaphors require a subject as an antecedent. SELF-anaphors must be locally bound—with some notable exceptions—whereas SE-anaphors allow a more remote antecedent, with some intriguing crosslinguistic variation (Everaert 1986). Romance languages have reflexive clitics that may or may not be combined with SELF-type elements. The same holds true for a number of Slavic languages.

Many languages have possessive anaphors, but others, including English, do not. Languages pervasively allow 1st and 2nd person pronominals to be locally bound, which is already problematic for the canonical binding condition B, but some (e.g. Frisian) even allow local binding of 3rd person pronominals. Complementarity between pronominals and anaphors faces problems even in English, for example, in locative PPs, as already noted in Chomsky 1981. Furthermore, in the early 1970s Cantrall (1974), Ross (1970), and others had observed that under certain conditions the local binding requirement on English anaphors can be obviated. And, as discussed by Clements (1975), under suitable discourse conditions, certain languages allow SE-anaphors not to be bound (Thráinsson 1991). Subsequently, many more languages were discovered that show unexpected complexities. Yet, CBT remains a surprisingly good approximation. Nevertheless, it became clear that its foundations had to be reconsidered.

The book under review, like its companion (Safir 2004), is an impressive contribution to the overall enterprise of working toward an explanatory binding theory in the face of the *prima facie* bewildering variety of anaphoric systems. It is the first attempt to develop a comprehensive

<sup>1</sup> It should be clear that I am not a standard disinterested reviewer. In developing his approach Safir takes issue with extant alternatives, and extensively discusses and criticizes Reinhart & Reuland 1991, 1993, and occasionally Reuland 2001, which became available at a relatively late stage of the preparation of the manuscript. It would have been tempting to include a reply, rather than just writing a review. A real discussion, however, would have taken up much more than the available space. I decided to leave most of the reply part for another occasion (some of the issues are addressed in Reuland 2010), limiting myself here to what seemed necessary for current purposes. I would like to thank Martin Everaert, Alexis Dimitriadis, and especially Ken Safir for their valuable comments on an earlier draft. All remaining errors are mine.

theory of binding based on one theoretical perspective applied to a wide variety of languages. Thus, it is a required read for anyone interested in binding.

The book is extremely rich in content. A thorough evaluation of each point the book makes would lead me far beyond the scope of a review. What I can do is place its main contribution in the current theoretical debate, guide the reader through the main issues, and point out a couple of cases where the argumentation is less compelling than *prima facie* might seem to be the case.

Safir addresses the problems of the CBT by making *DEPENDENCY* and *COMPETITION BETWEEN ANAPHORIC ELEMENTS* into the cornerstones of his theory. In many respects, as he acknowledges, his proposal draws on earlier insights from Luigi Burzio, Lars Hellan, Howard Lasnik, Pierre Pica, Tanya Reinhart, Eric Reuland, and others, but as he argues, the pie should be cut differently than they have done.

How to eliminate coindexing plays a prominent role in the current debates about the representation of interpretive dependencies. Reinhart (1983) has already shown that the conception of an index in the CBT needed revision. Within the minimalist framework, further rethinking of the way binding is encoded in the grammar was required, since the inclusiveness condition limits computations in (narrow) syntax to elements of a purely morphosyntactic vocabulary. Indices cannot be part of such a vocabulary.

S's proposal, elaborating on Reinhart 1983 and subsequent work, distinguishes between two forms of *COCONSTRUAL: COREFERENCE* (two expressions pick out the same referent in discourse) and *DEPENDENT IDENTITY* (expression A can only have its referential value determined as a function of the interpretive content of expression B) (24). The indices of the CBT are replaced by Higginbotham-style arrows marking dependency relations (not obeying inclusiveness, though).

S's further theoretical claim is that the syntax of anaphora is governed by a general form-to-interpretation principle (FTIP). Simplifying, FTIP is a proposal for capturing what underlies the complementarity between anaphors and pronominals expressed by principles A and B of the CBT. FTIP governs the competition between potentially dependent forms. Potentially dependent forms compete to represent a given dependent identity interpretation (7). Forms are ranked on a dependency scale and from this scale the most dependent form available is chosen to represent the dependency. In Germanic, for instance, the dependency scale is represented as in 1a and in French as in 1b (86, 87).

- (1) a. SIG-SELF >> pronoun-SELF >> SIG >> pronoun >> R-expression
- b. se >> independent clitic pronoun/tonic pronoun >> r-expression

FTIP, then, entails that in a context where SIG and a pronoun compete, SIG will have to be selected, since SIG is more dependent. The dependency scale is based on the following assumptions: first, being an anaphor is a property of a subset of lexical items, namely those that lack deictic potential. The property of being an anaphor—or being a possibly dependent element in a dependent identity relation—is a primitive lexical property (anaphors are a lexically marked subset of the class of forms that lack deictic potential (86));<sup>2</sup> and second, between any two anaphors, the more referentially specified one is more dependent. Among nonanaphors, the more referentially specified one is less dependent (for instance, SIG-SELF is more referentially specified than SIG, and hence will win in a competition context).

The notion of 'referential specification' itself does not receive an independent definition. S himself notes that its contribution is reversed if one crosses the transition point on the scale between +anaphor and –anaphor. This remains stipulative unless one would be provided with a definition from which this could follow.

FTIP can be said to emulate the canonical condition B, chain-formation effects from Reinhart & Reuland 1991, 1993 (henceforth R&R), and condition C effects.

The theoretical contribution of the book as a whole can be summarized in the following claim: 'A theory based on the notion of a dependency as a primitive can cover the ground of theories

<sup>2</sup> Note that conceptually the property *ANAPHORIC* is just the converse of the property [+R], from Reinhart & Reuland 1991, 1993, which characterizes pronominals, distinguishing them from simplex and complex anaphors. This is unlike what S's criticism of the [+R] property suggests.

based on coindexing, and—if supplemented with the notion of competition—with substantially increased empirical coverage’ (40). On the basis of the evidence presented, this claim is indeed substantiated.

As always, a ‘big picture’ has to be elaborated in the form of concrete proposals. In doing so, S introduces the following specific principles. I summarize them and indicate how they relate to principles proposed by others:

- (i) FTIP (see above).
- (ii) Pragmatic obviation (54): if FTIP does not permit *y* to be interpreted as dependent on *x*, then *x* and *y* form an obviate pair. It is intended to capture the effects of Rule 1 from Grodzinsky & Reinhart 1993.
- (iii) Local antecedent licensing (LAL) (148): an anaphor must be *c*-antecedent in domain *D*, where domain *D* for *X* is the minimal maximal extended projection containing *X* (where the verb may extend the projection of a *P* with a dependent complement). It captures the part of condition A that anaphors must be bound.
- (iv) The locally reflexive principle (LRP) (108): an identity-specific anaphor (SELF-form) is dependent on its coargument antecedent if it has one. It sets out to capture condition A (licensing of reflexivity) as formulated by R&R.
- (v) The coargument dependency constraint (CDC) (104): if *A* is identity-dependent on *B* and *A* and *B* are coarguments, then for any distributed interpretation of *B*, *A* depends on every distributed atom of *B* in the same way (captures a residue of condition B not captured by FTIP).
- (vi) The independence principle (40): if *X* *c*-commands *Y*, then *Y* is not an antecedent of *X*.
- (vii) Promotion: a dependent form (anaphor) is promoted to an unbounded dependent form (UD-form), a discourse-sensitive dependent, if it does not participate in a complete thematic complex (179). This principle is intended to capture the exemption condition on SELF anaphors of R&R and Pollard & Sag 1992, 1994.

These principles are largely specific to binding. Although descriptively an advance on the CBT, they do not relate the binding conditions to more elementary properties of the grammar.

In this respect, S does not go as far as Hornstein (2001) and Kayne (2002), who propose to derive (a subset of) anaphoric dependencies from movement, or Reuland (2001), who proposes an encoding of dependencies of simplex anaphors based on morphosyntactic feature checking (Agree in Reuland 2005).<sup>3</sup> S clearly opts for descriptive coverage. Yet, this difference limits the validity of S’s criticism of more ambitious proposals.<sup>4</sup> It would have been worthwhile if the book had provided a perspective on how to move toward further explanation and a more parsimonious theory.

The book takes the reader on a journey through the system. S presents his theory by taking the basic intuition behind FTIP as a starting point, introducing the other principles listed above as needed, and dealing with potentially problematic facts as they arise. As the theory is developed, it is pitted against alternatives, with especially extensive discussion of Lasnik 1989, Grodzinsky & Reinhart 1993, Reinhart & Reuland 1991, 1993, and occasionally Reuland 2001.

<sup>3</sup> It is important to distinguish between such goals in principle and shortcomings of specific choices and implementations (see Safir 2003 for pertinent criticisms of the latter).

<sup>4</sup> As an example, take the case of locally bound 3rd person pronominals in Frisian. Reuland (2001) derives the syntactic encoding of anaphoric dependencies on the basis of the morphosyntactic feature specification of the elements involved. The binding facts of Frisian derive from the properties of the Frisian case system (Reuland & Reinhart 1995) and are embedded in a wider pattern of variation among German dialects discussed there. On p. 244, n. 9, S dismisses this analysis without further argument as being both ‘implausible and unnecessary’. S, however, has to stipulate that the Frisian pronominal *him* (optionally) carries the feature anaphoric, in order to capture its relation to the anaphor *himself*. Thus S just stipulates what Reuland (2001) sets out to derive.

It should come as no surprise that I followed the arguments with keen interest. The ground to be covered is substantial. To keep space for a *pars construens*, many arguments of the *pars destruens* had to stay relatively superficial, though. The overall focus of the argumentation is conceptual rather than empirical. Hence, where the book promises a refutation of existing theories, the argument is often not really conclusive.<sup>5</sup>

Illustrative is S's discussion of 'Lasnik-contrasts' such as \**we voted for me* (distributive) versus *we elected me* (collective). This contrast is important, since it is inconsistent with a competition-based model. It is also important for his discussion of R&R, since R&R argue these contrasts to show that certain disjointness effects hold at a semantic level (their revised condition B is a condition on semantic predicates). S takes the issue up in Ch. 3. He argues that Lasnik-contrasts fall under the CDC (see §2).

CDC, however, is specifically introduced to capture distributive interpretations. This clearly carries a cost, and thus requires empirical motivation. But the summing up just states that principle B (Lasnik 1989 or R&R) 'must be rejected' for three reasons (95):

- A. Most of these distributions [the distributions it covers—*EJR*] are redundantly accounted for by FTIP.
- B. Where there is a general restriction against coargument dependency, the relevant restriction does not appear to be specifically about pronouns or reflexive morphology.
- C. Cases like (3) [Lasnik's cases—*EJR*] that were part of what originally motivated Lasnik's disjoint reference notion do not seem to be sufficiently general to support a predicate blind principle like Principle B or RIP [S's term for R&R's condition B—*EJR*].

Both A and B fail to be compelling, though. They presuppose FTIP rather than present independent support for it. C applies to the canonical condition B, but not to R&R, since the latter's condition is not predicate blind. S (p.c.) notes that the argument shows that FTIP, CDC, and so on can replace principle B effectively. That may be correct. However, it is the 'must be rejected' of the alternative that is not shown. In part this may point to an issue in the exposition. S is presenting a comprehensive system; an endeavor to provide independent support for each of its component parts is quite ambitious. A more feasible alternative might have been to focus on the evaluation of the theory as a whole against its ambitions.

Straw men constitute a pervasive instrument in the exposition, and this is understandable. S's task is daunting, and, laudably, he decided to systematically discuss the alternatives. Without straw men, the book could have easily become double the size. Yet, since straw men tend to be overly simplistic, the result of the discussion is sometimes hard to evaluate.

To give one example, in Ch. 6 (182) S discusses what he calls a rigid internalist position of anaphora, characterized as the thesis that establishing a form's inner nature should directly predict its distribution. He ascribes this position to his own previous work, to Faltz 1977, Pica 1987, and to R&R. He sets out to REFUTE it in favor of the hypothesis that the internal properties of potentially dependent forms do not predict their distribution directly, but only serve to predict their availability to participate in FTIP competitions. Both R&R and Pica (1987, 1991), however, albeit in a different vein, argue that the behavior of potentially dependent forms is governed by their internal feature composition TOGETHER WITH the way these features interact with the syntactic environment. Neither adopts the competition model, but also neither meets the criteria for being

<sup>5</sup>The discussion is occasionally apodictic and less than accurate. For reasons of space, one example should do. For instance, the introduction states that 'the condition on A-chains [from R&R—*EJR*] . . . lacks independent motivation' (18). However, no argument against R&R's claim that it comes for free is presented. The subsequent discussion ignores that it takes care of hierarchy, and so on. I have no doubt that S had some arguments in mind; they are not given here. The point may seem moot by now, since R&R's chain notion is preminimalist, but also in Reuland 2001 and 2005 encoding by chains is argued to come for free.

'rigidly internalist' as defined.<sup>6</sup> Although the argument against the position as defined is convincing, however, it remains open to which actual proposals it applies.

This is not just a quibble, but, just like the Lasnik-contrasts discussed above, it bears on the basis of S's approach, since both involve cases where the complementarity that FTIP predicts breaks down. A Dutch case of the relevant type is given in 2, allowing a bound occurrence of either *zich* or *hem*.

- (2) Max zag het boek achter zich/hem.  
Max saw the book behind SIG/him.

R&R predict that in Dutch bound SE-anaphors and pronominals show no complementarity in positions as 2, from which syntactic chain formation is blocked. For FTIP 2 is problematic if *zich* and *hem* are forms of unequal dependence. S outlines differentiating strategies to be applied if forms of unequal dependence appear to support the same dependent reading in the same structural context: (a) interpretations are distinct, (b) forms tie on the most dependent scale, and (c) numerations are distinct (97). Neither (a) nor (b) applies in this case. For instance, both variants support strict and sloppy readings in *only*-contexts. There is no evidence that Dutch *zich* ever undergoes 'promotion' to an UD form (see below). As to (c), the assumption of an optional PRO in the case of *hem* would help. But I know of no independent motivation, and, for instance, R&R's account does not need it. So 2 does indeed go against S's straw man, but follows from R&R, and raises a problem for FTIP. S's discussion of his example 68, a more complex Danish case of noncomplementarity in object position, stays inconclusive in the end; hence here a problem for FTIP remains as well.

English exempt anaphors illustrate a different issue. In order to account for the noncomplementarity between *him* and *himself* in cases such as 3—a potential problem for FTIP—S invokes the notion of promotion of an anaphor to discourse-sensitive UD form as in 4.

- (3) In Arthur's opinion, Lisa should trust no other than himself. (§5.4, ex. 60c)  
(4) An anaphor can be promoted to a discourse-sensitive dependent if it does not participate in a complete thematic complex.

In R&R's system, which is among the systems that S takes as a starting point, the exemption ('logophoricity') of *himself* follows directly from the way condition A is formulated. The same holds true of the alternative by Pollard and Sag (1992, 1994). Both give explicit definitions of when an anaphor is exempt. By contrast, S gives no explicit definition of the core notions 'participate in' or 'complete thematic complex' in 4. It may well be the case that 'participate in' is to be read as R&R's 'be a syntactic argument of', and 'complete thematic complex' as 'syntactic predicate', but this is not stated. This makes things unnecessarily hard on the reader, since such information is needed for assessing the proposal.

It would have helped if S had presented the reader with a key, relating the terms employed to theories by Pollard and Sag, Reinhart and Reuland, or others, with a clear discussion of convergence or divergence, and differences in predictions.

The statement in 4 is also not general enough to capture the 'promotion' of Icelandic 'logophoric' *sig*. Here, not the thematic complex but the subjunctive mood is involved.<sup>7</sup> It may be possible to accommodate it in S's system, but it remains unclear how it could be done without further stipulation.<sup>8</sup>

The notion of promotion would equally benefit from elucidation. It has the flavor of a process, but besides that, it is not articulated. Does it involve manipulation of features, for instance, eliminating the property 'anaphoric'? If so, it would go against the inclusiveness condition. Does it reflect some other process? This is something one would like to know, and no doubt S will be able to present an answer. But regrettably we have to wait for another book to get it.

<sup>6</sup> Note, though, that how Pica summarizes what he does differs from what he actually does.

<sup>7</sup> Reuland 2001 shows that subjunctive blocks chain formation.

<sup>8</sup> Safrir (2005) argues that Icelandic *sig* can be used as a true logophor. He leaves open, though, what licenses its use as a logophor or an anaphor.

It would have been helpful for the reader if the book had listed all principles and assumptions in an overview section and compared them systematically to the alternatives discussed, not only in empirical coverage, but also in aims and scope (it would already have helped if the index had marked crucial page references in bold face). I did some of this earlier in this review in the hope of helping readers to find their way through this important book.

The book is rich in its coverage of both empirical and theoretical issues. It is impossible to go into the detail needed to do justice to every argument and insight it contains, given the restrictions on a review. I can only recommend reading it, and also assessing its arguments carefully, because even where they are inconclusive, they are worth studying.

We all know that our theories will at some point turn out to be wrong. But some theories turn out to be useful nevertheless, because they help us focus on what we want to understand, and identify patterns that can survive paradigmatic changes. S's is one of these. I feel S took up a real challenge in writing this book, systematically developing a uniquely comprehensive perspective based on an impressive range of facts. As always, progress can only be achieved by constructively and critically evaluating substantive proposals that are presented. I trust my comments contribute to such an evaluation. Despite the problems I noted, S's synthesis is a very important achievement. Studying this book advanced my understanding, as it will advance anybody's understanding of this intriguing subject.

#### REFERENCES

- CANTRALL, WILLIAM. 1974. *View point, reflexives and the nature of noun phrases*. The Hague: Mouton.
- CHOMSKY, NOAM. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- CHOMSKY, NOAM. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- CLEMENTS, GEORGE N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10.141–77.
- EVERAERT, MARTIN. 1986. *The syntax of reflexivization*. Dordrecht: Foris.
- FALTZ, LEONARD. 1977. *Reflexivization: A study in universal syntax*. Berkeley: University of California, Berkeley dissertation.
- GRODZINSKY, YOSEF, and TANYA REINHART. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24.69–103.
- HORNSTEIN, NORBERT. 2001. *Move!* Oxford: Blackwell.
- JESPERSEN, OTTO. 1933. *Essentials of English grammar*. London: Allen & Unwin. [Reprinted, Tuscaloosa: University of Alabama Press, 1964.]
- KAYNE, RICHARD. 2002. Pronouns and their antecedents. *Derivation and explanation in the minimalist program*, ed. by Samuel D. Epstein and T. Daniel Seely, 133–66. Oxford: Blackwell.
- KOSTER, JAN, and ERIC REULAND. 1991. *Long distance anaphora*. Cambridge: Cambridge University Press.
- LASNIK, HOWARD. 1989. *Essays on anaphora*. Dordrecht: Kluwer.
- PICA, PIERRE. 1987. On the nature of the reflexivization cycle. *North Eastern Linguistic Society* 17.483–99.
- PICA, PIERRE. 1991. On the interaction between antecedent-government and binding: The case of long-distance reflexivization. In Koster & Reuland, 119–35.
- POLLARD, CARL, and IVAN SAG. 1992. Anaphors in English and the scope of the binding theory. *Linguistic Inquiry* 23.261–305.
- POLLARD, CARL, and IVAN SAG. 1994. *Head-driven phrase structure grammar*. Stanford, CA: CSLI Publications, and Chicago: University of Chicago Press.
- REINHART, TANYA. 1976. *The syntactic domain of anaphora*. Cambridge, MA: MIT dissertation.
- REINHART, TANYA. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.
- REINHART, TANYA, and ERIC REULAND. 1991. Anaphors and logophors: An argument structure perspective. In Koster & Reuland, 283–321.
- REINHART, TANYA, and ERIC REULAND. 1993. Reflexivity. *Linguistic Inquiry* 24.657–720.
- REULAND, ERIC. 2001. Primitives of binding. *Linguistic Inquiry* 32.439–92.
- REULAND, ERIC. 2005. Agreeing to bind. *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 505–13. Berlin: Walter de Gruyter.
- REULAND, ERIC. 2010. *Anaphora and language design*. Cambridge, MA: MIT Press, to appear.
- REULAND, ERIC, and TANYA REINHART. 1995. Pronouns, anaphors and case. *Studies in comparative Germanic syntax*, ed. by Hubert Haider, Susan Olsen, and Sten Vikner, 241–69. Dordrecht: Kluwer.

- ROSS, JOHN R. 1970. On declarative sentences. *Readings in English transformational grammar*, ed. by Roderick A. Jacobs and Peter S. Rosenbaum, 222–72. Waltham, MA: Ginn.
- SAFIR, KEN. 2003. Anaphora, movement and coconstrual. *Grammar in focus, vol. 2: Festschrift for Christer Platzack*, ed. by Lars-Olaf Delsing, Cecilia Falk, Gunlog Josefsson, and Halldor Sigurdsson. Lund: Department of Scandinavian Languages, Lund University.
- SAFIR, KEN. 2004. *The syntax of (in)dependence*. Cambridge, MA: MIT Press.
- SAFIR, KEN. 2005. Person, context and perspective. *Rivista di Linguistica—Italian Journal of Linguistics* 16.107–54.
- THRÁINSSON, HÖSKULDUR. 1991. Long-distance reflexives and the typology of NPs. In Koster & Reuland, 49–76.

Utrecht University  
 Utrecht Institute of Linguistics OTS  
 Janskerkhof 13  
 3512 BL Utrecht  
 The Netherlands  
 [eric.reuland@let.uu.nl]

**Corpus-based approaches to sentence structures.** Ed. by TOSHIHIRO TAKAGAKI, SUSUMU ZAIMA, YOICHIRO TSURUGA, FRANCISCO MORENO-FERNÁNDEZ, and YUJI KAWAGUCHI. (Usage-based linguistic informatics 2.) Amsterdam: John Benjamins, 2005. Pp. vi, 315. ISBN 9789027233141. \$143 (Hb).

Reviewed by MATTI RISSANEN, *University of Helsinki*

This volume has been produced by the Usage-Based Linguistic Informatics (UBLI) project at Tokyo University of Foreign Studies. The project is a part of the 21st Century Center of Excellence Program, funded by Japan's Ministry of Education, Sports, Culture, Science and Technology. The UBLI program aims to produce 'an overall integration of Theoretical and Applied Linguistics', which 'will be realized on the basis of Computer Sciences' (3). The research concentrates on linguistic usage, which is illustrated with large amounts of textual data. Thus, it is only natural that most of the articles of the volume are based on evidence derived from corpora or other text sources.

The research carried out by UBLI and reported on in the present volume is in many respects related to the area of evidence-based linguistic study that sees language mainly as a means of communication within a speech community, first and foremost between the speaker/writer (SP/W) and audience/reader (AD/R), with due attention paid to discourse and pragmatics (see e.g. Traugott & Dasher 2002). In recent decades this approach, which pays special attention to linguistic variation, has rapidly gained ground, particularly since the introduction of computerized corpora. The study of variation takes into account both extralinguistic factors—mainly sociolinguistic, genre-based, or regional—and language-internal trends of variation, such as metaphoric and grammaticalization. Theoretical issues are, of course, duly included in the analysis and in the conclusions drawn. Many of the articles in the present volume offer interesting starting points for more elaborate and extensive studies of variation. We can also hope that in the compilation of new corpora the extralinguistic factors causing variation will be kept in mind.

The most remarkable feature of the volume, which consists of three introductory chapters and fourteen articles, is the variety of the languages discussed, both non-Indo-European and Indo-European. They include Japanese, Korean, Chinese, Malay, Tagalog, Nuuchahnulth, Turkish, Arabic, Russian, German, French, English, and Spanish. Nuuchahnulth deserves a special mention: it is a Wakashan language spoken by only a handful of people in British Columbia. In a way, it is a pity that no Finno-Ugric languages are included in any of the studies. Finnish, for instance, which contains both original non-Indo-European and contact-based Indo-European syntactic features, might have offered an interesting point of comparison for some of the constructions discussed.