

Semantic Aspects of Morphological Typology

Joan Bybee
University of New Mexico

The traditional division of languages into isolating, agglutinative, fusional and polysynthetic has been an extremely useful typology for many generations of linguists. As we have now come to expect, this typology does not yield discrete categories, but does provide some landmarks that serve as prototypes for the organization of our knowledge about language.

The most influential discussion of morphological typology during the twentieth century is found in Sapir's book, *Language*, where he provides a moderate and intelligent interpretation of this typology, informed by the languages of Native America. An important focus of Sapir's discussion is the role that meaning plays in morphological typology. Sapir distinguishes four types of concepts:

- | | | |
|---------------------|---|-----------------------------------|
| (1) <i>Material</i> | { | I. Basic Concepts |
| <i>Content</i> | | II. Derivational Concepts |
| <i>Relation</i> | { | III. Concrete Relational Concepts |
| | | IV. Pure Relational Concepts |

The first two types are more lexical in character, including the concrete and abstract senses of nouns and verbs, and derivational categories with greater lexical content. The third and fourth types are increasingly grammatical in nature, with the fourth type including such purely grammatical concepts as agreement.

Sapir goes on to classify languages according to the types of meaning they express. He thus makes an extremely important point, which has not been taken up in later applications of the typology. The point is this: *languages of different morphological types express different types of meaning.*

Aside from the general neglect of semantics in structuralist and generative approaches to language, perhaps the reason that the semantic aspects of this typology have not been further studied is the difficulty in categorizing concepts. For example, it is difficult to establish exactly what Sapir meant by the distinction between Concrete Relational and Pure Relational concepts. Some linguists even have difficulties understanding the difference between derivational and inflectional concepts, as seen in Anderson's (1982, 1992) repeated claim that there is no difference in the concepts expressed derivationally and those expressed inflectionally.

Recent studies of grammaticization have developed scalar approaches to meaning that position lexical concepts on one end of a diachronic continuum and grammatical meaning on the other (Bybee and Dahl 1989; Bybee, Perkins and Pagliuca 1994; Givón 1975, 1979; Heine and Reh 1984; Heine, Claudi and Hünnemeyer 1991). Such grammaticization paths do not correspond fully to Sapir's classification of meaning, but they do provide a principled universal ordering of meanings in relation to one another that can be used to investigate the semantic consequences of morphological typology. Using cross-linguistically valid grammaticization paths for the classification of concepts, I will document in this paper the fact that Sapir's typology has both a formal and a semantic dimension. This will lead to a restatement of the point above: *languages of different morphological types carry grammaticization out to differing extents.*

1. The Morphological Types

Four parameters are usually mentioned in discussions of morphological typology. None of these parameters is meant to be applied categorically; all are scalar and position languages closer to or farther from a prototype. Here I will mention a fifth criterion and will argue later in the paper for its significance.

1. Affixation: Analytic languages are distinguished from all others by the fact that they lack affixes (Sapir 1921: 128).
2. Number of morphemes per word: Languages which typically or obligatorily include more grammatical categories in a word are considered more synthetic than those that include fewer. Thus if a language has verbs inflected for aspect, tense, mood as well as agreement (as in Latin, see example [3]), it can be considered more synthetic than one that requires only tense (as in English).
3. Degree of fusion: Agglutinative vs. inflectional or fusional languages are distinguished on the basis of the segmentability of words into morphemes.

In an agglutinative language such as Buriat (Mongolian), it is easy to find the morpheme boundaries, and the relation among allomorphs is largely phonologically determined.

- (2) Agglutinative: Buriat (Poppe 1960:57)
jaba-na-gyi-b
 GO-PRES-NEG-1SG
 'I do not go'

In Latin the process of dividing words into morphemes is complicated by the existence of portmanteau morphs, assimilation and fusion at morpheme boundaries, and the existence of lexically-determined allomorphy in the form of conjugation classes and genders.

- (3) Fusional: Latin
port-ā- v- ī
 carry-1CONJ-PERF:1CONJ-1SG:PERF:IND
 'I carried'

4. Incorporation of two or more items of a more lexical nature in a single word. Polysynthetic languages not only have many grammatical categories expressed in the verb, but also typically display more than one lexical stem per word (Fortescue 1994).

- (4) Polysynthetic: Greenlandic (Fortescue 1984)
naja- lisar- puq
 little:sister-bring:along-3SG:IND
 'He has brought along his little sister'

All of these criteria deal very directly with formal aspects of expression, in particular, how morphemes are combined into words. I would like to suggest a fifth criterion of a slightly different nature that is relevant to morphological typology; this criterion — obligatoriness — has to do with the way grammatical categories function in context.

5. Obligatory categories are those which must have some exponent in a given morpho-syntactic context. Thus for English count nouns, number is obligatory, because all noun phrases containing count nouns have number expressed. In this case, the lack of a marker signals singular, which is then said to have zero expression. As another example, determiners, which are not affixal in English, are also obligatory: English noun phrases have grammatically determined interpretations of definiteness and specificity. The lack of a determiner is highly constrained — occurring only with

proper nouns and plurals. In the latter case the absence of the determiner is meaningful — it signals nonspecific or generic meaning.

The criterion of obligatoriness further defines analytic languages, since one property of analytic languages is that they lack obligatory categories. Thus while languages such as Mandarin Chinese have very little affixation, this is only a part of the more pervasive generalization that no grammatical categories are obligatory in Mandarin Chinese. For instance, the expression of pronouns or aspectual categories (whether affixal or not) is never obligatory.

The lack of obligatory categories is not, however, restricted to languages that lack affixes. Some languages, such as Car (Nicobarese) have derivational affixes, but no inflectional ones; that is, they have non-obligatory morphology, but no obligatory morphology. Basically, then, this criterion corresponds to the presence or absence of inflectional morphology.

2. Areal and Genetic Patterns

Languages of these different types are not evenly distributed throughout the world. Instead, we find strong genetic and areal patterns of distribution. In the Gramcats database, as reported in Bybee, Perkins and Pagliuca 1994, all general grammatical morphemes (grams) associated with verbs were coded in a sample of 76 languages. These languages were randomly selected within genetic groups as established by Voegelin and Voegelin (1978).

Those languages with fewer than 30% of their grams affixed to the verb can be considered strongly isolating. The languages in this category, with their phyla are listed in (5):

(5)	Austroasiatic:	Koho Palaung
	Oceanic branch of Austronesian:	Tahitian Tangan Atchin Nakanai
	Sino-Tibetan:	Cantonese Haka Lahu Lao Nung

Niger-Kordofanian: Engenni
Mano

Creole: Tok Pisin

Taking into account other languages in each phyla, we can say that analyticity is a very strong tendency in Sino-Tibetan, Austroasiatic and Oceanic, but only a minor tendency in Niger-Kordofanian. Sino-Tibetan, Austroasiatic and Oceanic are areally contiguous in many places, suggesting that analyticity is heavily influenced by genetic factors and areal contact.

Polysynthesis, defined by Fortescue (1994) as involving both incorporation and a large inventory of bound morphemes, is also areally and/or genetically influenced. Most polysynthetic languages are found in North America and occur in several groups: Eskimo-Aleutian, Algonquian, Iroquoian, Caddoan, Na-Dene, Uto-Aztecan, Wakashan, Salishan, Hokan, Totonac-Tepehua and Mixe-Zoque families.

These two polar types, then, are identified with particular parts of the world. The types in between are much more difficult to pin down, though still, certain strong areal or genetic tendencies are discernible. For instance, certain families tend to have consistent agglutinative structure: Ural-Altai, Bantu and Australian Pama-Nyungan languages. Also, certain families in contiguous areas tend to be highly fusional: Afro-Asiatic, Indo-European, Caucasian and Dravidian. The existence of these genetic and areal patterns attests to the relative stability of morphological typology over time, a point that has consequences for our understanding of the grammaticization process.

3. Grammaticization

Grammaticization is the process by which lexical morphemes in particular constructions become grammatical and then continue their development. It should be seen as a continuous process ongoing in all languages at all times. It involves the parallel development of form, meaning and grammatical behavior. On the level of form, increasing grammaticization corresponds to increasing synthesis (Bybee et al. 1994; Givón 1975, 1979; Heine and Reh 1984; Lehmann 1982.)

As Hopper (1991) has pointed out, languages such as English have constructions that show varying degrees of grammaticization, even within the same semantic domain. In most cases, more synthetic structures have a more grammaticized meaning. Thus, compare the English Perfect, *have waited*, to the simple past. The Perfect involves more components of meaning, including as it

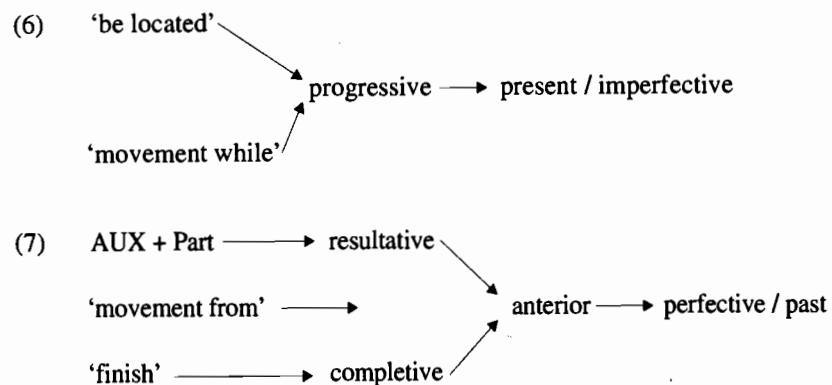
does, not just pastness, but also the sense of current relevance.

If grammaticization is ongoing in all languages at all times, then all languages would contain an even mix of analytic, agglutinative and fusional structures, and there would be no types. Since there are clear types, and since these types are relatively stable over time, there must be some difference in the way grammaticization proceeds in languages of different types.

If grammaticization involves the parallel development of meaning and form, then languages which grammaticize form to a lesser extent (e.g. analytic languages) must also grammaticize meaning to a lesser extent. Cross-linguistic studies of grammaticization provide an avenue for the study of these intriguing issues.

4. Paths of Grammaticization for Tense and Aspect

Cross-linguistic studies of grammaticization patterns reveal, among others, two heavily used paths for the grammaticization of tense and aspect (Bybee and Dahl 1989; Bybee et al. 1994)



In his cross-linguistic survey of tense and aspect, Dahl (1985) noticed that meanings which are earlier on these paths, i.e. progressive and anterior, tend to have periphrastic expression, while the meanings to which they give rise tend to have inflectional expression, i.e. they are both bound and obligatory.

Table 1. Expression of major gram-types in Dahl (1985), Bybee and Dahl (1989)

Periphrastic		Affixed	
anterior	88%	perfective	73%
progressive	95%	past	85%
		imperfective	100%

In Bybee et al. (1994), using a more detailed analysis of form, we show a strong correspondence between the degree of fusion with the verb, the degree of dependence on surrounding material and the shortness of the gram with the meaning expressed by it.

In this study we were able to divide the second path shown above, which we can call the Perfective Path, into five stages, which we called Perfages, because they represent five stages in the semantic development or semantic age of the grams in question:

- (8) Perfage 1 completives
 Perfage 2 young anteriors
 Perfage 3 old anteriors
 Perfage 4 perfective
 Perfage 5 past

Young anteriors have no other uses but anterior, while old anteriors have added other uses, such as perfective past, suggesting that they are farther along in their development. When this semantic classification is compared to the expression types of these grams (a total of 165 in our database) we found that grams with higher Perfages also tended to be:

more fused with the verb, as measured by whether or not they are written bound to the verb, condition allomorphy in the verb and have allomorphy conditioned by the verb ($p < .0001$).

more dependent upon surrounding material in general, as measured by the number of allomorphs, allomorphs conditioned by other grammatical morphemes, the ability to take stress ($p < .0001$).

shorter in terms of the number of consonants and vowels that comprise its longest allomorph ($p < .0001$).

These studies confirm that form and meaning develop in parallel in the process of grammaticization.

5. Consequences for Typology

If some languages do not have affixes or obligatory categories, it would imply that these languages cannot have the more grammaticized tense/aspect meanings of *perfective / imperfective* or *present / past*. That is precisely the finding of Dahl (1985), Bybee and Dahl (1989), and Bybee et al. (1994). Languages that otherwise lack inflection (that is, lack categories that are obligatory and bound) also tend not to have inflectional perfective / imperfective or present / past distinctions.

In the more or less analytic languages listed in (5) above, past and perfective grams are generally lacking, with the possible exception of the Oceanic languages, some of which appear to have nonbound grams expressing past or perfective that are obligatory.

Moreover, grams marking past or perfective are so widespread in the languages of the world, that with very few exceptions, the only languages that do not have them are those that lack all other inflections as well. Languages that lack inflection — analytic languages — lack the more grammaticized types of grammatical meaning. It follows, then, that the meanings expressed by grammatical morphemes in analytic languages are different from those expressed by inflectional languages.

For instance, an analytic language is more likely to have a gram expressing anterior (past with current relevance) or completive (to do something thoroughly and completely) than one expressing past or perfective. Such grams are not likely to be obligatory, which means that they are not used redundantly as often as grams which are obligatory. They are only used where their semantic content is nonredundant and contributing to the new information. Consider example (9) showing the use of the anterior form *awe*, literally meaning 'it is finished', in Sango, a creole language of Central Africa (Samarin 1967: 158–162):

- (9) Sango (Samarin 1967)
kòlì así gígí awe, ála zia lóró da,
 manSUBJ:arrive outside finish 3PL put speed there
akpé na wále bíaní awe.
 SUBJ:run and woman truly finish
 'When the man had come out, they put speed into it, and ran away with the woman.'

Note that this sentence contains three clauses, all in a past context, but only the first and third are marked with *awe*. In the first clause *awe* functions as an anterior, sequencing one event before another to which it is relevant. In the third

clause, *awe* signals completion. In a language with an obligatory past or perfective, the second clause would also require marking. Semantically, *awe* is adding much more to the utterance than a past or perfective gram would.

In Bybee et al. (1994: 115–119), we report a significant association between the overall degree of fusion and dependence in a language and the degree of semantic development of its grams on the Perfective Path. This shows again that languages that are more synthetic and fusional tend to have grams that are more semantically grammaticized. No such association was found between the overall shortness of grams in a language and the degree of grammaticization of meaning. But of course, no one has ever proposed that the length of morphemes is a significant typological parameter.

6. Typology and Grammaticization

Our studies establish a significant relation between morphological typology and the extent to which grammaticization, both of form and meaning, is carried out. We have confirmed Sapir's claim that languages of different morphological types express different types of meaning, although without using precisely his categories. Moreover, we are not viewing typology as synchronic or static, but rather viewing the creation and maintenance of a type as a dynamic matter. So to arrive at an explanation for morphological typology, we must ask what is different about grammaticization in languages of different types.

Grammaticization takes place slowly and gradually as language is actually used. It is driven by the interaction of the speaker and hearer, as they strive to communicate effectively and efficiently using the resources afforded them by their culture. Some of the mechanisms of semantic change we find in grammaticization are generalization of meaning, based often on metaphor and metonymy, and the conventionalization of implicature, by which commonly made inferences become part of the meaning of a form or construction.

We do not yet have enough detail about grammaticization in different languages to know for sure how it differs in analytic vs. synthetic languages. But since the most prototypical analytic languages lack obligatory categories, it follows that they must lack the discourse conventions that cause obligatory categories to develop.

How do obligatory categories develop? Two important factors must be recognized (García and van Putte 1989; Bybee 1994):

- (i) the redundant use of a morpheme, which increases its frequency, and leads to

- (ii) the inference that if X is not present, not-X is meant, which creates a zero and thus an obligatory category.

If discourse conventions eschew the use of redundant material, the prerequisite high frequency may not be achieved and if the inferencing style does not include inferring not-X if X is not uttered, then there is no way to create obligatory categories.

In a study of sentence processing in Chinese, Li, Bates and MacWhinney (1993) found that Chinese speakers use the lexical semantics of Ns much more consistently than grammatical cues to choose which N is the subject of the sentence. The animacy of the Ns proved more important in speakers' choices than word order or the object marker *ba*. Only the passive marker *bei* affected the speakers' choice more than animacy. *Bei*, however, is infrequent in the spoken language and is by no means an obligatory category. In fact, though called a passive marker, it has rather specific meaning, including the adverse notion that something unfortunate or undesired happened.

Compared to languages with inflection or obligatory word order patterns, Chinese sentences lack redundancy and obligatory grammatical cues. A variety of factors, including context and lexical semantics, are used to interpret sentences. In making inferences speaker/hearers juggle a number of factors, none of which rigidly determines a unique interpretation. Speaker/hearers must remain open to a variety of possible interpretations of sentences.

Speakers of languages with obligatory categories, whether word order or inflection, come to points in sentences in which reliable cues close down all but one interpretational option. They become accustomed to a type of inferencing in which presence and absence of particular elements is criterial. Presumably, using this style of inferencing in discourse will lead to the continued creation of obligatory categories. In the absence of this inferential style, obligatory categories are not created.

Chinese has been an analytic language for several millennia. Grammaticization takes place in Chinese: it is simply not carried as far as in other languages. Lin (1991) traces the development of several anterior markers from the earliest written texts in Chinese to the present and finds several markers that have progressed through early portions of the Perfective Path shown in (7). However, these markers do not show the dramatic frequency increase and redundant use characteristic of grammaticization in other languages, nor do these markers progress to perfectives, with the possible exception of the modern marker *-le*, which is a perfective when verb-final.

I suggest, then, that the conditions necessary for the development of obligatory categories are not present in Chinese discourse interpreting conven-

tions. Speakers avoid redundancy, and hearers maintain an open set of inferencing options; in particular they do not elevate grammatical cues over those provided by the lexical semantics and the context.

Of course, this hypothesis requires much further study. Inferencing conventions may be of many different types and they may interact with other typological parameters in a variety of ways. Very little research has been directed at cross-cultural comparison of discourse processing strategies. At least one study has pointed out differences in inferencing patterns between Japanese and Chinese on the one hand, and English on the other hand. Hinds 1987 has argued that in the context of written English, the writer accepts the responsibility for the clear communication of the message, while in the oriental cultures he observed, the reader has the responsibility for inferring the intended meaning. This observation would correspond well to a lack of redundancy in grammatical expression. However, Chinese and Japanese are not morphologically of the same type. What they do have in common, however, is the lack of grammatical agreement, the ability to 'leave out' noun phrase arguments, and tense and aspect systems that are less grammaticized than in more purely inflectional languages. Also in Japanese direct object case marking is optional (Fujii and Ono 1995) and pragmatic factors determine whether the other direct argument of the verb is marked with *wa* or *ga*. Perhaps these properties correspond to the discourse interpreting strategies that leave many open inferences for the hearer, and the presence or absence of affixation is determined by other factors (Bybee et al. 1990).

7. Conclusion

When the meaning expressed in different languages is compared, we find many universals, particularly in the paths of development the meaning takes in the course of grammaticization. However, we also find that languages express different types of meaning, not just lexically, but also grammatically. The basic parameter is the degree of specificity or generality: Some languages generalize grammatical meaning to a greater extent than others do.

Since generalization of meaning is a process that takes place as language is used, it follows that different degrees of generalization of meaning will arise out of different strategies for producing and processing discourse. Such strategies are part of a broader cultural heritage, and may be resistant to change. They are also independent of any particular linguistic forms, which accounts for their influence on the development of new forms and their stability over time.

References

- Anderson, Stephen R. 1982. "Where's morphology?" *Linguistic Inquiry* 13: 571-612.
- Anderson, Stephen R. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press. [Cambridge Studies in Linguistics 62].
- Bybee, Joan L. 1994. "The grammaticization of zero: asymmetries in tense and aspect systems." In *Perspectives on Grammaticalization*, W. Pagliuca (ed.), 235-254. Amsterdam: John Benjamins.
- Bybee, Joan L. and Östen Dahl 1989. "The creation of tense and aspect systems in the languages of the world." *Studies in Language* 13: 51-103.
- Bybee, Joan L., William Pagliuca and Revere Perkins. 1990. "On asymmetries in the affixation of grammatical material." In *Studies in Diachronic Typology for Joseph H. Greenberg*, W. Croft, K. Denning and S. Kemmer (eds), 1-42. Amsterdam: John Benjamins.
- Bybee, Joan L., Revere Perkins and William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.
- Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.
- Fortescue, Michael. 1994. "Morphology, polysynthetic." In *Encyclopedia of Language and Linguistics*, R. Asher (ed.). Pergamon Press.
- Fujii, Noriko and Tsyoshi Ono. 1995. "The occurrence and non-occurrence of the Japanese direct object marker *o* in conversation." Unpublished manuscript, University of Oregon.
- García, Erica and F. van Putte. 1989. "Forms are silver, nothing is gold." *Folia Linguistica Historica* VIII/1-2: 365-384.
- Givón, Talmy. 1975. "Serial verbs and syntactic change: Niger-Congo." *Word Order and Word Order Change*, C. Li (ed.), 47-112. Austin: University of Texas Press.
- Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
- Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: The University of Chicago Press.
- Hinds, John. 1987. "Reader vs. writer responsibility: a new typology." In *Writing Across Languages: Analysis of L2 text*, U. Connor and R. B. Kaplan (eds), 141-152. Redding, MA: Addison Wesley Publishing Company.
- Hopper, Paul J. 1991. "On some properties of grammaticization." In *Approaches to Grammaticalization*, E. Traugott and B. Heine (eds). Amsterdam: John Benjamins.
- Lehmann, Christian. 1982. *Thoughts on Grammaticalization: A programmatic sketch*. Vol. I. Arbeiten des Kölner Universalien-Projekts Nr. 48.
- Li, Ping, Elizabeth Bates and Brian MacWhinney. 1993. "Processing a language without inflections: a reaction time study of sentence interpretation in Chinese." *Journal of Memory and Language* 32: 169-192.

- Lin, Zi-yu. 1991. The development of grammatical markers in Archaic and Han Chinese. Unpublished Ph.D. dissertation, SUNY at Buffalo.
- Poppe, Nicholas N. 1960. *Buriat Grammar*. Bloomington: Indiana University.
- Samarin, William. 1967. *A Grammar of Sango*. The Hague: Mouton.
- Sapir, Edward. 1921. *Language*. New York: Harcourt Brace and World.
- Voegelin, Carl R. and Florence M. Voegelin. 1978. *Classification and Index of the World's Languages*. New York: Elsevier.

ESSAYS ON
LANGUAGE FUNCTION
AND LANGUAGE TYPE

Dedicated to T. Givón

Edited by

JOAN BYBEE

University of New Mexico

JOHN HAIMAN

Macalester College

SANDRA A. THOMPSON

University of California

INDIANA UNIVERSITY

LIBRARIES

BLOOMINGTON

P
204
.E86
1997

JOHN BENJAMINS PUBLISHING COMPANY

AMSTERDAM / PHILADELPHIA

malh



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

RSC4
IAJH 5237

Library of Congress Cataloging-in-Publication Data

Essays on language function and language type : dedicated to T. Givón / edited by Joan Bybee, John Haiman, Sandra A. Thompson.

p. cm.

Includes index.

I. Typology (Linguistics) 2. Grammar. Comparative and general. 3. Functionalism (Linguistics) I. Givón, Talmy, 1936- . II. Bybee, Joan L. III. Haiman, John. IV. Thompson, Sandra A. P204.E86 1997

410'.1-dc21

97-7390

ISBN 90 272 2168 5 (Eur.) / 1-55619-522-2 (US) (alk. paper)

CIP

© Copyright 1997 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Contents

Instead of an Introduction	1
From His Colleagues	3
His Publications	13
The Essays	23
Semantic Aspects of Morphological Typology <i>Joan Bybee</i>	25
The Typology of Predicate Case Marking <i>Bernard Comrie</i>	39
Grammaticalization and the Gradience of Categories: Relator Nouns and Postpositions in Tibetan and Burmese <i>Scott DeLancey</i>	51
A Typology of Argument-Determined Constructions <i>R.M.W. Dixon and Alexandra Y. Aikhenvald</i>	71
Are Grammatical Relations Universal? <i>Matthew S. Dryer</i>	115
Introducing Ergative Word Order via Reanalysis: Word Order Change in the Cariban Family <i>Spike Gildea</i>	145
Living in Three Languages <i>Colette Grinevald</i>	163
Self-Abasement in Language: A Case Study on the Viability of a Metaphor <i>John Haiman</i>	181
The Misumalpan Causative Construction <i>Ken Hale</i>	199