Ergativity in Discourse and Grammar

John W. DuBois

The Oxford Handbook of Ergativity
Edited by Jessica Coon, Diane Massam, and Lisa Demena Travis

Print Publication Date: Jun 2017 Subject: Linguistics, Morphology and Syntax
Online Publication Date: Aug 2017 DOI: 10.1093/oxfordhb/9780198739371.013.2

Abstract and Keywords

This chapter considers how a discourse profile may provide a key piece of the puzzle for explaining the distribution of ergative grammatical structures within and across the world's languages. The ergative discourse profile, isomorphic to the ergative-absolutive pattern of syntactic alignment, is found in a typologically diverse array of languages including ergative, accusative, and active. Speakers tend to follow soft constraints limiting the Quantity and Role of new and lexical noun phrases within the clause. Evidence for the universality of the ergative discourse profile is examined from typology, child language, and diachrony. A conflicting discourse pressure for topicality motivates accusativity, giving rise to competing motivations. As one recurrent resolution of competing demands, ergativity represents an evolutionarily stable strategy realized in grammar. While discourse-pragmatic and cognitive motivations contribute crucially to a functional explanation of ergativity, additional factors must include semantics of verbs, constructions, aspects, and splits; inherited morphosyntax; and more.

Keywords: ergativity, ergative discourse profile, Quantity Constraint, Role Constraint, soft constraints, adaptive selection, absolutive zero, word order, universals

EVERY language provides its users with systematic ways of organizing the core arguments of the clause, establishing a more or less stable and consistent framework for the foundations of its grammar. Remarkably, languages differ even in this most basic level of structural organization. Yet certain configurations of arguments tend to recur, emerging again and again in the grammars of the world's languages. For syntactic alignment what proves pivotal is how the grammar treats the sole argument of a one-place predicate (S), aligning it with one or the other argument of a two-place predicate, with respect to case-marking, agreement, word order, extraction, and so on. Some languages treat the S like the object of a two-place predicate (O), yielding ergative alignment (S=O vs. A), while others treat S like the subject of a two-place predicate (A), yielding accusative alignment (S=A vs. O). Still other languages are sensitive to the semantic variability inherent in the population of one-place predications, aligning the more agent-like subset of S (S_a) with A,

and the more patient-like subset of S (S_o) with O, giving active alignment (S_a=A vs. S_o=O).

Yet languages are not simply ergative, accusative, or active. There is great diversity, as well as convergence, across the world's languages with respect to the various systematizations of basic grammatical relations (Silverstein 1976; Comrie 1978, 2013a, 2013b, 2013c; Dixon 1979, 1994; Mithun 1991; Malchukov 2005; Bickel & Nichols 2009; Bickel 2010; Siewierska 2013). This diversity can penetrate into the grammar of a single language, where a mix of distinct alignments is often found in different parts of the same grammar (Silverstein 1976; Comrie 1978; Mithun 1991; Coon 2010a, 2013a; Malchukov 2014; see also Coon and Preminger, Chapter 10 in this volume). Accounting for diversity and convergence within and across languages remains a compelling yet elusive task for linguistics (Evans & Levinson 2009). Argument structure configurations (Goldberg 1995) represent the very foundations of the clause—what may be considered the "basic body plan" (Mayr 2001) of the utterance. No theory of grammar can be considered explanatory without contributing to an understanding of how such a diversity of basic (p. 24) plans could have emerged in the world's languages. Yet the challenge of accounting for structural variability at a foundational level has proved baffling, such that many leading linguists have postponed the day of reckoning with ergativity. Fillmore considered his principles of subject selection universal, "given certain qualifications for the interpretation of ergative systems" (1977: 61). Dowty acknowledged that "argument selection in ergative languages" (1991: 581) was relevant to his proto-role model, but invoked an "inverse" model of ergativity which "means in effect treating the transitive 'Patient' as a grammatical subject and the transitive 'Agent' as analogous to an object" (1991: 582). Ergative languages are said to employ the same proto-agent and proto-patient roles as in accusative languages, but "merely REVERSE the syntactic association" with subject and object (1991: 582). Often it seems that ergativity is taken up only after commitments to basic theoretical assumptions are set (Ackerman & Moore 2001: 1, fn. 1). But ergativity is unlikely to reveal its secrets to those who approach it superficially or too late, whether with afterthoughts or mere mirror-image models.

This chapter explores the connection between the well-known ergative pattern in grammar and a pattern in discourse that is isomorphic to it, with the goal of providing a functional explanation for ergativity. The ergative discourse pattern holds the key to the grammaticization of ergativity, perhaps—or is just a piece of the larger puzzle. The specific approach presented here is known informally as discourse and grammar (Givón 1979; Hopper & Thompson 1980; Thompson & Mulac 1991; Du Bois 2003b, 2014), which seeks to understand grammar in light of discourse, and discourse in light of grammar. Patterns of grammatical form are linked to communicative functions on the evidence of naturally occurring language use, in order to shed light on why grammars are the way they are. Crucially, "grammars" is plural, inviting attention to typological diversity. Not only must the broad alignment types of ergative, active, and accusative be distinguished, but also such cross-cutting typological features as head-marking vs. dependent-marking, optional vs. consistent, aspect-based vs. person-based splits, and so on (DeLancey 1981; Nichols

1986; Garrett 1990; Bickel & Nichols 2009; Malchukov 2014; Nichols & Bickel 2013a; van de Velde 2014).

The discourse-and-grammar approach accords equal importance to discourse and to grammar. In this chapter, however, I will devote more space to the discourse side of the equation. This is feasible because ergative grammar is well documented in the literature (including this volume); it is necessary because ergative discourse is not. That said, a key task will be to bring together the facts of ergative grammar and ergative discourse, and to clarify the connection between them.

If ergativity is seen as a problem, it's one that is not going away any time soon. That may be a good thing for linguists, who have a lot to gain by taking up the challenge of explaining ergativity. But one group for whom ergativity has never been a problem is the speakers of ergative languages. Ergative speakers do just fine, learning their language with equal ease (Ochs 1982; Narasimhan, Budwig, & Murty 2005; Bavin & Stoll 2013; Brown et al. 2013; Pye et al. 2013), and using it to perform the full range of functions that every language serves (de León 1999, 2000). Linguists seeking explanations might take a cue from the language users, and treat ergative (p. 25) grammar as a system that works. The idea is to see ergativity as the solution, and ask what the problem is. This is not to suggest that there will be easy answers, nor that every aspect of ergative grammar will be transparently motivated or directly functional. Nonetheless, how a grammar works for its users is one of the most productive questions a linguist can ask.

The view that grammars solve problems in new and creative ways accords well with the approach advocated by Evans and Levinson (2009), who, while arguing for massive cross-linguistic diversity, also maintain that languages tend toward "evolutionarily stable strategies," representing "recurrent solutions across time and space." These strategies

> result from myriad interactions between communicative, cognitive, and processing constraints which reshape existing structures through use. A major achievement of functionalist linguistics has been to map out, under the rubric of grammaticalization, the complex temporal sub-processes by which grammar emerges as frequently used patterns sediment into conventionalized patterns
>
> (Bybee 2000; Givón 2009).(Evans & Levinson 2009: 444–445)

From the discourse-and-grammar perspective, the enterprise begins in discourse with the search for "frequently used patterns"; it continues by showing how the "recurrent solutions" resolve universal functional needs; and it ends, if successful, by elucidating the emergence of the "conventionalized patterns" known as grammar—including the grammar of ergativity and its competitors. One seeming paradox is that grammar is already present in discourse from the start. Discourse is never found without its grammatical clothing; but by the same token, grammar is never realized except in discourse. A basic task for discourse-and-grammar research, then, is to tease discourse and grammar apart. With ingenuity and a little typology, the problem is solvable, as will become clear. This

then sets the stage for investigating the connection between the respective ergative patterns in discourse and in grammar.

I begin (2.1) with a look at a stretch of discourse in an ergative language, identifying a recurrent pattern which is isomorphic with the ergative–absolutive pattern of grammar. The next section (2.2) documents this pattern quantitatively as the ergative discourse profile, based on evidence from a typologically diverse array of ergative, active, and accusative languages. I propose that the ergative discourse profile is shaped by a set of soft constraints known as Preferred Argument Structure (Du Bois 1987b, 2006; Du Bois, Kumpf, & Ashby 2003). The next section (2.3) explores whether the ergative discourse profile represents a discourse universal, examining evidence from child language, typology, genre, and diachrony. In the following section (2.4) I take up the functional explanation for ergativity, having introduced the analysis of competing motivations (Du Bois 1985, 2014; MacWhinney, Malchukov, & Moravcsik 2014; Malchukov 2014) that favor either ergative or accusative alignment. Next (2.5) I respond to some common objections to the discourse explanation for ergativity. Finally, I present some directions for future research (2.6) and conclusions (2.7).

(p. 26) 2.1 Ergativity in Discourse

To understand ergativity, it is important to look at how discourse connects to grammar and to meaning. As Dixon says

> The most important task for future work on 'why some languages are ergative in a certain way and others are not' is to investigate the semantic and discourse-pragmatic makeup of each of a sample of languages, and study the way in which this determines (or partly determines) its grammatical profile.
>
> (1994: 219–220)

If discourse has the power to affect grammar, it is because "discourse, clause structure, and verb semantics are all intimately interwoven" (Foley & Van Valin 1984: 373). But to move beyond broad generalities about the interdependence of form and function, it is necessary to tease apart these three forces, if only to show how they come together again to shape the fundamental grammar of the clause.

To make good on the promise that "recurrent patterns" shape grammar (Evans & Levinson 2009), it is necessary to do the empirical work to document the specific "discourse profiles" (Du Bois 2003a: 40–44) that are linked to the grammatical constructions of interest, and their functional niches. The relevant work on discourse profiles focuses on "discourse inside the clause" (Du Bois 2003a: 13; 2003b: 83), seeking to identify the distinctive discourse correlates of the clause, its arguments, and other aspects of structure. While researchers sometimes speak broadly of the discourse profile of a language, akin to whole-language typology (Nichols & Bickel 2013b), it is often more useful to target the discourse profile at a more specific level—on a par with a specific argument structure

construction, for example. Thus one can ask about the discourse profile of the intransitive, transitive, or ditransitive clause; or the agentless vs. agentive passive, and so on.

One challenge in working with discourse is its evident variability, born of freedom. Seemingly, speakers exercise the absolute liberty to construct their utterances as they wish, within the broad limits circumscribed by the rules of grammar. The result appears, to some, as unpredictable variability. Yet a closer look reveals recurrent regularities in discourse, including some which are not strictly required by any grammatical rule. Understanding ergativity depends on sorting out how argument structure constructions balance the multiple demands of information processing, anaphoric reference, topic continuity, event semantics, and other factors, including the inherited morphosyntax of the language at hand.

To make these matters more concrete, it will be useful to examine a sample of discourse from an ergative language. The following narrative is in Sakapultek, an ergative, head-marking, verb-initial language of the Mayan family, spoken in highland Guatemala (Du Bois 1981). The narrative was produced as a telling of a short film, the Pear Film (Chafe 1980; Du Bois 1980). To highlight the grammatical elements that matter for the discourse profile, the following conventions are used: underscore represents a (p. 27) referent mention marked by a reduced referential form (pronoun, agreement, or zero);[1] **boldface** indicates a lexical noun phrase; *italics* in the free translation indicates a verb or preposition (head of its phrase). The distribution of lexical vs. reduced forms is also indicated schematically, with capital letters (A S O) indicating a lexical noun in the designated grammatical role, while small letters represent a reduced form (a s o). A clause-by-clause analysis of this narrative's ergative discourse profile is published elsewhere (Du Bois 2006); here I present just the first 13 lines.

(1)

Pear Story (Sakapultek)

1	… (H) Ee x-Ø-inw-il-anh,	oaV
	FOC CP-3.ABS-1.ERG-see-TV	
	… (H) What I *saw* was,	
2	.. x-Ø-aq'an **jun achenh**,	VS
	CP-3.ABS-ascend one man	
	.. **a man** *climbed up*,	
3	.. ch-u' **chee'**,	PX
	at-top tree	
	.. *in* **a tree**,	
4	… (H) x-Ø-a-r: … ch'up-o' **niky'aj péera-s.**	aVO
	CP-3.ABS-LAT-3.ERG pick-DEP some pear-PL	
	… (H) he went and … *picked* **some pears.**	
5	… Tik'ara' Ø-Ø-qaaj-uu: l,	sV
	then CP-3.ABS-descend-hither	
	… Then he *came* dow: n,	
6	… Ø-Ø-r-su' **r-iij juu: n**,	aVO
	CP-3.ABS-3.ERG-wipe 3.ERG-back one	
	… he *wiped* **the surface of one**,	
7	… (H) ii despwee: s,	
	and then	
	… (H) and the: n,	
8	… (H) x-Ø-r-ya'[2] qaj p l **chikech**,	oaV PX
	CP-3.ABS-3.ERG-put down at the basket	
	… (H) he *put* it *in* **the basket**,	

2

(p. 28)

9 Ø-Ø-r-alsa-aj p l r:- … m komo ber gabaacha. oaV PX
CP-3.ABS-3.ERG-remove-TV at the 3.ERG um like see apron
he *removed* it *from* his: … um like apron.

10 … Despwees tik'ara',
then then
… Then,

11 … Ø-Ø-pee jun aj-laab', VS
CP-3.ABS-come one DIM-boy
… a little boy *came*,

12 .. ch-ij bisikleeta, PX
at-back bicycle
.. *on* a bicycle,

13 … (H) xaq x-Ø-a-r"-k'am-a' jun chkech peera. aVO
just CP-3.ABS-LAT-3.ERG-take-DEP one basket pear
… (H) he just came and *took* a basket of pears.

Viewed in grammatical terms, the data exhibit the hallmarks of ergative alignment in the grammar of verbal agreement (pronominal clitics). Focusing on third person singular referents, transitive subjects (A) are cross-referenced with *r-* '3rd person singular ergative' (lines 4, 6, 8, 13). In contrast, intransitive subjects (S) are unmarked, i.e. cross-referenced with Ø- '3rd person singular absolutive' (lines 2, 5, 11). Transitive objects (O) receive the same treatment as S (lines 4, 6, 8, 13). The ergative–absolutive pattern holds throughout the agreement paradigm (Du Bois 1981, 1987a: 210; 1987b: 809–810), as in virtually all Mayan languages (Larsen & Norman 1979; England 1983; Robertson 1983; Kaufman & Norman 1984; Law 2009).

Viewed in functional terms, the data illustrate some common patterns in the realization of basic discourse functions. For example, a new human referent is introduced using a full lexical noun phrase in the S role (lines 2, 11), and is subsequently tracked through the discourse with reduced forms (e.g. zero). The tracking of the most topical referents occurs most often in reduced A's (lines 4, 6, 8, 13), but also in a reduced S (line 5). Inanimate referents are introduced here with a lexical mention in O role (lines 4, 6, 13), and tracked in subsequent discourse using a reduced O (lines 8, 9). Locative prepositional phrases also serve for the introduction of lexical and New inanimates, which may occur as adjuncts in separate intonation units (lines 3, 12), or in more tightly integrated verbalizations within the same intonation unit as the verb (lines 8, 9).

The point of this exercise is to show what discourse looks like when viewed through the lens of grammar. To generalize from this tiny sample, lexical mentions occur mostly in absolutive argument positions (S and O roles), expressing New information. A's are mostly reduced forms, and express given information. In this verb-initial language, the preferred order of overt lexical nouns is VO and VS (in the four-way typology of Dryer 1997, 2013a,

2013b), which can be generalized as V-Lex word order: verb followed by a (p. 29) lexical noun.[3] The attentive reader will have noticed that the discourse distribution of lexical arguments (and of new information) corresponds to the absolutive category in the grammar of ergative languages, while topically continuous elements are found in what would be the subject in accusative languages. The latter reflects, perhaps, the common wisdom that subject is a grammaticization of agent and topic (Givón 1983a; Comrie 1988).

The ergative and accusative discourse patterns coexist in one and the same stretch of narrative, in a language whose grammar is basically ergative. The tension between these two discourse patterns will prove fruitful for understanding the role of competing motivations (Du Bois 1985, 2014) in the discourse motivation of ergativity (1987a; Du Bois 1987b, 2006); see section 2.4. But we are getting ahead of ourselves. First we must ask: Do these observations characterize the discourse of one speaker, or do they represent something broader? This calls for a quantitative perspective, which is addressed in terms of discourse profiles.

2.2 The Ergative Discourse Profile

The first systematic study of information structure to distinguish A, S, and O was by Du Bois (Du Bois 1981, 1985, 1987a, 1987b, 2003a, 2003b, 2006; Du Bois, Kumpf, & Ashby 2003). Previous studies had commonly been framed in terms of subjects, documenting contrasts in information structure that naturally seemed to favor accusative languages. But as long as S and A were collapsed within the all-too-familiar subject category, the differences between them remained effectively hidden. Adopting the grammatical terms A, S, O (Dixon 1979) or P (Comrie 1978), originally designed for typologically neutral comparisons encompassing ergative and accusative languages, made it possible to recognize a new kind of pattern in discourse. Early evidence from Sakapultek, illustrated in example (1), revealed a skewed distribution of lexical arguments across A, S, and O. A similar skewing was discovered for New information. Specifically, most lexical mentions occur in absolutive argument positions (S or O), but are avoided in the ergative (A) slot, which is mostly restricted to reduced forms (pronoun, agreement, zero). Correspondingly, most new mentions occur in S or O, with few occurring in A. Many more languages have since been investigated; see Table 2.1 for a selection. In most cases the findings reported tend to fit the generalizations of Preferred Argument Structure. This holds true whether the language is ergative, active, or accusative. Of course, not all scholars agree on the interpretation of the findings; for a discussion of controversial issues, see 2.5.

It is useful to distinguish between discourse profiles, understood as observable patterns of linguistic behavior in discourse, and Preferred Argument Structure, understood as the functional constraints that govern them. The ergative discourse profile (p. 30) (p. 31) can be observed in a corpus as a skewed distribution of new and lexical noun tokens across the argument slots of the clause. Preferred Argument Structure represents the constraints or generalizations which—as a first approximation—appear to govern the skewed distribution of utterance tokens. Whether the four constraints should be considered mere

generalizations about linguistic practices, or whether there are deeper cognitive motivations behind them, is a question that remains open to alternative theoretical (p. 32) interpretations (see section 2.5). Though closely related, the two perspectives are useful to distinguish, at least for the moment.

Table 2.1 Selected studies of Preferred Argument Structure

Language	Family	Region	Type	Genre	Studies
Sakapultek	Mayan	Guatemala	ergative	narrative (PF)	(Du Bois 1987b)
Mam	Mayan	Guatemala	ergative	narrative	(England & Martin 2003)
Tektiteko	Mayan	Guatemala	ergative	narrative	(England & Martin 2003)
Mochó	Mayan	Mexico	ergative	narrative	(England & Martin 2003; Martin 2003)
Q’anjob’al	Mayan	Guatemala	ergative	narrative	(England & Martin 2003)
Itzaj	Mayan	Guatemala	ergative	narrative	(Hofling 2003)
Tsotsil	Mayan	Mexico	ergative	narrative	(Martínez Álvarez 2012)
Hieroglyphic Maya	Mayan	Guatemala, Mexico	ergative	dynastic history	(Mora-Marín 2004)

Inuktitut	Eskimo-Aleut	Canada	ergative	child	(Allen & Schröder 2003)
Hindi	Indo-Iranian	India	ergative	child	(Narasimhan et al. 2005)
Nepali	Indo-Iranian	Nepal	ergative	narrative	(Genetti & Crain 2003)
Hawrami	Indo-Iranian	Iran	ergative	narrative (PF)	(Mahand & Naghshbandi 2014)
Gooniyandi	Australian	Australia	ergative	narrative	(McGregor 1998)
Roviana	Austronesian	Solomon Is.	ergative	monologue	(Corston-Oliver 2003)
Chamorro	Austronesian	Guam	ergative?	narrative	(Scancarelli 1985)
Acehnese	Austronesian	Indonesia	active	narrative	(Durie 1988, 1994, 2003)

Chol	Mayan	Guatemala	active	narrative	(Vázquez Álvarez & Zavala Maldonado 2013)
Yagua	Yaguan	Peru	active	narrative	(Thomas Payne 1993)
Mapudungun	Araucanian	Chile	hierarchical	narrative	(Arnold 2003)
Tohono O’odham	Uto-Aztecan	Arizona		narrative	(Doris L. Payne 1987)
Jarawara	Arauan	Amazon	accusative	narrative	(Dixon 1994: 209)
Japanese	Japonic	Japan	accusative	conversation	(Matsumoto 1997, 2000)
Korean	Koreanic	Korea	accusative	child	(Clancy 2003)
To’aba’ita	Austronesian	Solomon Is.	accusative	narrative	(Lichtenberk 1996)
Finnish	Uralic	Finland	accusative	conversation	(Helasvuo 2003)

Hebrew	Semitic	Israel	accusative	narrative	(Sutherland-Smith 1996)
English	Indo-European	United States	accusative	narrative (PF)	(Kumagai 2001, 2006)
English, di-achronic	Indo-European	England	accusative	drama	(Shibasaki 2006)
French	Indo-European	France	accusative	conversation	(Ashby & Bentivoglio 1993)
French, di-achronic	Indo-European	France	accusative	epic poetry	(Ashby & Bentivoglio 2003)
Spanish	Indo-European	Venezuela	accusative	conversation	(Ashby & Bentivoglio 1993)
Spanish, di-achronic	Indo-European	Spain	accusative	epic poetry	(Ashby & Bentivoglio 2003)
Portuguese	Indo-European	Brazil	accusative	narrative	(Dutra 1987)
Portuguese	Indo-European	Brazil	accusative	various	(Brito 1996, 1998; Everett 2009)

multiple	various	various	various	various	(Haig & Schnell 2016)

Notes

The languages are arranged by alignment type, then by family. The designation as ergative or accusative is not meant to dichotomize, as its necessarily glosses over details of split ergativity, etc.

“PF” indicates Pear Film narratives (Chafe 1980).

Specifically, Preferred Argument Structure comprises four soft constraints, which collectively influence the discourse distribution of grammatical expressions and pragmatic statuses. In the grammatical dimension, the Quantity constraint limits the number of lexical arguments in the clause core to a maximum of one. The Role constraint specifies where in the clause the single lexical argument may appear, excluding it from the A role. Paralleling this pair of constraints is a second pair in the pragmatic domain. Here the Quantity constraint limits the number of new information arguments in the clause core to a maximum of one. The Role constraint specifies where in the clause this new argument will appear, again excluding it from the A role.

In contrast to the generalizations of Preferred Argument Structure, the ergative discourse profile represents the empirically observable pattern of recurrent linguistic behavior, characterized by a statistical skewing in the distribution of lexical argument and new information tokens in a population of utterances. Broadly speaking it is the difference between rules and consequences—or generalizations and facts on the ground. Table 2.2 summarizes the relation between the four constraints of Preferred Argument Structure and the corresponding consequences observable in the ergative discourse profile.

To be more precise, the pragmatic constraints should be described as applying, not to new information, but to low accessibility referents (Ariel 1990, 2001), i.e. to referents whose cognitive status motivates a verbalization with a relatively informative (high surprisal) form. For present purposes, the looser formulation in terms of new information is adequate. But the more precise theoretical framing has important implications for research design, including the need for a continuous scale of accessibility/surprisal (Ariel 2001) and, correspondingly, a continuous variable for the size of linguistic forms. For future research, such terms of analysis are to be preferred.

Preferred Argument Structure is neither a syntactic structure nor a discourse structure, but a preference in discourse for the use of certain configurations of grammar. All four constraints are soft constraints (Givón 1979; Bresnan, Dingare, & Manning 2001), which can be violated without producing ungrammaticality. Yet in spontaneous language use, speakers tend to follow them. Together they influence the shape of discourse productions, yielding the distribution of argument structure constructions recognizable as the ergative discourse profile.

Figure 2.1 [4] presents findings from several languages regarding lexical argument quantity, that is, the frequency of clauses with zero, one, or two lexical arguments. (For Figures 2.1–2.4, the languages are sorted in the same sequence as Table 2.1, with ergative languages presented to the left, and accusative languages to the right.) While there is considerable variability, the key finding here is that clauses with two lexical arguments are consistently rare across languages of different typologies, regions, and genealogies.
(p. 33)

Table 2.2 Preferred Argument Structure and the ergative discourse profile

Constraint	Domain	Preferred Argument Structure	Ergative discourse profile
QUANTITY	GRAMMAR	Avoid more than 1 lexical core argument	lexical arguments ≤ 1
	PRAGMATICS	Avoid more than 1 New core argument	New arguments ≤ 1
ROLE	GRAMMAR	Avoid lexical A	free use of lexical S & O
	PRAGMATICS	Avoid new A	free use of New S & O

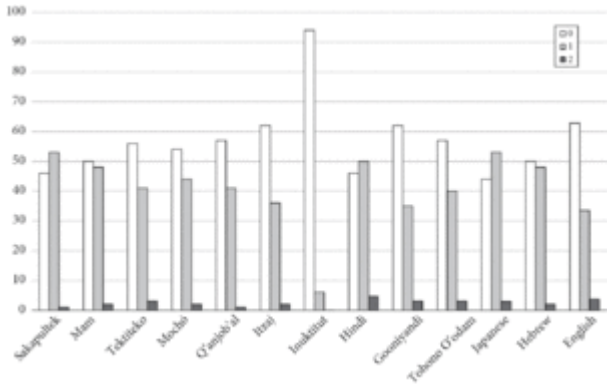

Figure 2.1 Lexical argument quantity: frequency of clauses with 0, 1, or 2 lexical core arguments

Noun phrases are not produced in a functional vacuum, of course, but are tied to cognitive-pragmatic processes of information management. The use of a lexical noun is linked to, but far from equivalent to, the presenting of new information. Thus the Quantity constraint on lexical core arguments is paralleled by a similar constraint on new core arguments. Figure 2.2 presents findings across several languages regarding New argument quantity, that is, the frequency of clauses containing zero, one, or two new arguments. Some language samples show no instances at all of clauses containing two new core arguments (Sakapultek, Roviana), while others show very few (Inuktitut, English). (p. 34)

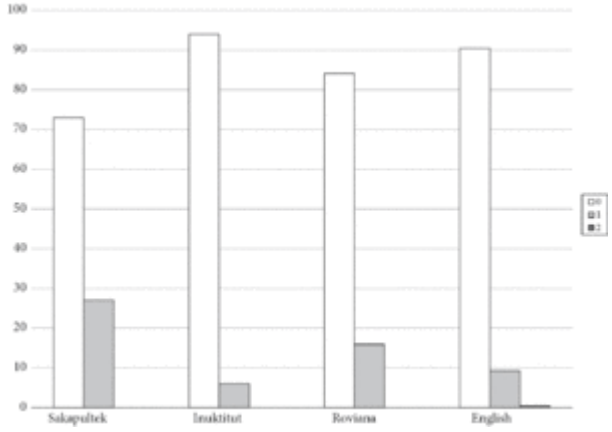

Figure 2.2 New argument quantity: Frequency of clauses with 0, 1, or 2 new core arguments

While the Quantity constraint allows speakers to freely deploy One Lexical Argument per clause core, not all syntactic positions in the clause are equally good candidates for realizing this mention. This is where the Role constraint comes in, constraining the use of lexical arguments in A role, while freely allowing their use in S or O roles. Figure 2.3 summarizes findings from a number of languages regarding where lexical arguments are realized within the clause, showing how lexical core arguments are distributed across A, S, and O. While there is considerable variation in some dimensions (e.g. the relative preference for S or O), again what matters most is what is specifically constrained: the A role, which shows relatively low frequencies of lexical mentions as a recurrent pattern across many of a typologically diverse array of languages. Nonetheless, variation here is substantial, raising interesting typological questions (see below).

For one-place predicates, it is easy to satisfy the Quantity constraint, since the S role is unconstrained, freely allowing the introduction of a new referent. Indeed, this may be one reason language users select a one-place predicate over a transitive alternative—for its pragmatic, rather than semantic, affordances. But for two-place predicates, the Quantity constraint on new information is more directly consequential. Thus the Role constraint establishes an opposition within the transitive clause between A and O, constraining new arguments in A, while freely allowing them in O. Figure 2.4 presents the distribution of new core argument realizations across A, S, and O.

A close examination of the frequencies in Figures 2.1–2.4 shows considerable cross-linguistic similarities in some dimensions, along with substantial variability in others. **(p. 35)** This is why it is important to evaluate the discourse evidence in light of a theoretical framework, such as Preferred Argument Structure, which offers specific hypotheses about which aspects of argument realization are constrained, and which are not. One misunderstanding that often arises in the literature involves an attempt to compare raw frequencies between two languages (one of which is usually Sakapultek). But whether frequencies match or not is beside the point (and no one language of the many surveyed has a privileged position). What matters instead is the testing of specific, theoretically moti-

vated hypotheses about constraints on syntactic positions in argument structure constructions, e.g. with respect to information structure. Another common temptation is to attribute the difference in frequencies found in two studies to the structural attributes of the languages in question. This may prove to be correct in some cases, and is certainly a question worth asking. But there are other candidates that should be considered as well in accounting for variance, especially genre differences. As the field of Preferred Argument Structure studies develops further, these issues are likely to become more visible, and we can anticipate new findings that tease apart the subtle variables involved. Now that so many typologically diverse languages have been studied, new questions arise about universality and variability of Preferred Argument Structure. Especially interesting is the potential for two-way interactions between the grammar of a language and its discourse profiles, occasioned by the never-ending cycle of grammaticization and language use. Are there different Preferred Argument Structures for (p. 36) different languages, and does this correlate with the ergative–active–absolutive typological contrast (Durie 1988, 2003)? The way is open for a broader theoretical framing of questions about why these recurrent patterns arise across such a broad typological array of languages, yet vary within seemingly well-defined limits. There is much that remains to be discovered.

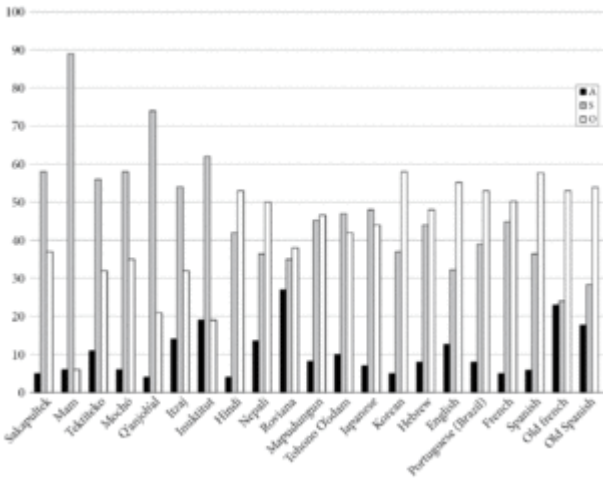

Figure 2.3 Lexical argument role: Distribution of lexical arguments across A, S, and O

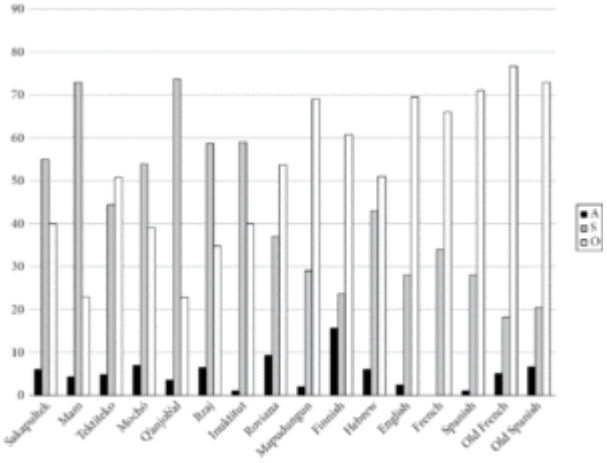

Figure 2.4 New argument role: Distribution of new arguments across A, S, and O

2.3 A Discourse Universal?

Is the ergative discourse profile universal? To answer this question, I begin with the child, and go on to examine the evidence from typology and diachrony.

2.3.1 Child Language

As might be expected, children who are exposed to ergative languages exhibit the ergative discourse profile. This has been shown for Tzeltal (Mayan, Brown 1998, 2008; de León 1999), Inuktitut (Eskimo-Aleut, Allen 2000; Allen & Schröder 2003), and Hindi (p. 37) (Indo-Iranian, Narasimhan 2013; Narasimhan et al. 2005), and similar observations have been made for Samoan child language (Austronesian, Duranti, & Ochs 1989; Ochs 1982).

But if the question is about universals, what is more telling is what happens in the discourse of children learning accusative languages. Clancy's extensive studies of Korean child language show that even when exposed to an accusative grammar, children produce the ergative discourse profile (Clancy 1993, 1995, 1996, 1997, 2003). For Spanish, Bentivoglio argues for late acquisition of the ergative discourse profile (1998), while other studies of Spanish and French show a more complex picture (Khorounjaia & Tolchinsky 2004; Salas 2010). Older English-speaking children with autism show the ergative discourse profile (Weber 2003), as do adults with aphasia (Kohn & Cragnolino 2003).

How deep does the ergative discourse profile go? For Goldin-Meadow this leads her to ask what the child's discourse can reveal about the mind, and she goes to great lengths to find out. Seeking out "people who have had no exposure to any conventional language whatsoever" (2003: 493), Goldin-Meadow works with profoundly deaf children who have

had no access to either spoken language or sign language. What these children do is invent their own gesture systems, with revealing implications for cognition:

> The thoughts of these individuals cannot possibly have been shaped by language. As a result, whatever categories they express reveal thoughts that *do not* depend on language—thought before language.
>
> (Goldin-Meadow 2003: 493)

One of Goldin-Meadow's most compelling discoveries is that in the invented gesture systems, discourse tends to "pattern like ergative languages: intransitive actors and patients are treated alike (produced), whereas transitive actors are treated differently (omitted)" (2003: 502). She observes:

> We have found the ergative pattern to be robust in communication situations. Deaf children of hearing parents who are inventing their own gesture systems tend to organize their gesture sentences around an ergative pattern. Equally striking, we found that when asked to describe a series of action vignettes using their hands rather than words, English-speaking adults invented an ergative structure identical to the one developed by the deaf children, rather than the accusative pattern found in their spoken language. These findings suggest that the ergative pattern may reflect a robust solution to the problem of communicating information from one mind to another, be it an adult or a child mind.
>
> (Goldin-Meadow 2003: 516)

While Goldin-Meadow finds evidence to support the absolutive distribution of lexical arguments, she does not find evidence for the role of new information (2003: 516; Schulman, Mylander, & Goldin-Meadow 2001). Pointing to eye-tracking research by Griffin and Bock (2000), Goldin-Meadow suggests that "focusing on patients may be a default bias found in both processing and acquisition tasks" (2003: 517). While their proposed "patient focus" motivation must be balanced against other (p. 38) factors—competing motivations—deriving from pragmatics, semantics, etc., Goldin-Meadow and associates have made a compelling case for the presence of an ergative discourse profile at the earliest stages of communication. This begins even before exposure to language, arguing for a deep-rooted cognitive basis for "thinking ergative" (Goldin-Meadow, Yalabik, & Gershkoff-Stowe 2000; Goldin-Meadow 2003).

2.3.2 Typology

From a typological perspective, the first question is whether the ergative discourse profile is restricted to languages with ergative grammar, or rather represents a property of the discourse of all languages, independent of syntactic alignment. If it is restricted to ergative languages, it could be a consequence of ergative grammar. But if it is a discourse universal, this is no longer a tenable position, and it becomes a potential factor in the functional motivation for ergativity (2.4).

As already noted, studies of Preferred Argument Structure have been carried out for a number of languages (2.2, Table 2.1, and Figures 2.1–2.4). While a truly random typological sample is not possible, the available languages appear to be reasonably diverse in typology, region, and genealogy, covering all major alignment types. Based on these languages, the preponderance of evidence supports the view that most languages, whether ergative, active, or accusative, tend to display an ergative–absolutive patterning in the discourse distribution of lexical (or "heavy") nouns, and of new (or low-accessible) information, in the core argument positions of the clause. To be sure, the interpretation of the discourse patterns, and even their existence, remains controversial for some scholars; I address these issues in 2.5. I believe that when the evidence from discourse, typology, child language, and diachrony are all taken into account, the picture that emerges is of an ergative discourse profile pervasive across languages, independent of the typology of syntactic alignment.

That said, one of the most interesting questions raised by cross-linguistic evidence like that in Figures 2.1–2.4 is whether ergative, accusative, and other language types may have their own distinctive variants of Preferred Argument Structure, maintaining some universal aspects while also fine-tuning to the grammar at hand, through cyclic interactions between grammar and use.

2.3.3 Diachrony

If the ergative discourse profile represents a universal, found in modern languages regardless of their type, the uniformitarian principle (Hock 1991)[5] predicts that older (p. 39) stages should share this feature as well. While few languages can provide direct textual evidence for older stages, some do. Diachronic studies conducted so far, though few, have consistently found the ergative discourse profile in both earlier and later stages of the same language. Among accusative languages, diachronic studies of Old Spanish, Old French, and their modern descendants show that "despite a gap of seven centuries … the medieval and modern forms of French and Spanish are remarkably similar in their manifestations of Preferred Argument Structure" (Ashby & Bentivoglio 2003: 73; see also Bentivoglio 1994). In a study of English spanning six periods from Old English to Modern English, Shibasaki concludes that, despite higher lexical densities[6] in older stages, each period largely conforms to the constraints of Preferred Argument Structure (Shibasaki 2006).

Among ergative languages, the Mayan family offers especially rich historical implications, given the available comparative and philological evidence. With 30 modern descendants, almost all of the languages have remained consistently ergative over four millennia (Norman & Campbell 1978; England 1983, 1991, 1990; Robertson 1983, 1992; Kaufman & Norman 1984; Law 2009, 2014).[7] In the domain of discourse, the ergative discourse profile has been documented in detailed case studies covering four of the six major branches of Mayan, extending across a wide expanse of the Mayan territory: Mamean (Mam and Tektiteko, England & Martin 2003); Q'anjob'alan (Q'anjob'al and Mochó, England & Martin 2003); K'ichean (Sakapulteko, Du Bois 1987b, 2006); and Yucatecan (Itzaj, Hofling

2003).[8] While at least one nineteenth century scholar claimed to reconstruct the text of a proto-Indo-European myth, such feats were ultimately deemed beyond the reach of the comparative method. But a discourse profile is not a text. Its wide distribution across the Mayan family makes the ergative discourse profile a plausible candidate for reconstruction to older stages. Support comes from early textual evidence in yet a fifth branch of Mayan (Ch'olan-Tzeltalan). Mora-Marín shows that Maya hieroglyphic inscriptions, written in an early Ch'olan language, display the ergative discourse profile (Mora-Marín 2004). While areal diffusion is a theoretical possibility, the family-wide distribution and textual evidence taken together point to a (p. 40) more likely conclusion: that the ergative discourse profile was present in early stages of Mayan, even Proto-Mayan. This conclusion is also consistent with the uniformitarian principle, given the typological evidence for universality.

The wealth of evidence for the Mayan family offers unique opportunities for tracing the history of ergativity in grammar and discourse. The implications are worth dwelling on here, as they are relevant not only for discourse-and-grammar reconstruction, but also for tracking the development of ergativity down through the daughter languages. One point of intersection between studies of comparative Mayan and the ergative discourse profile comes in the work of Nora England and Laura Martin (England 1983, 1991; England & Martin 2003). Drawing on their own research on the grammar and discourse of four languages from two branches of Mayan, they find the Given A constraint in both Mamean and K'ichean, and consider how discourse tendencies may become grammatical rules:

> It may be the case that, in K'ichean languages in particular, a grammatical restriction against indefinite subject NPs exists or is developing. This would presumably be a grammaticalization of the discourse constraint noted by Du Bois (1987b), that agents (ergators) are typically not used to convey new information. His analysis and analyses on other languages by England and Martin (1989) [=(England & Martin 2003)] show that, in texts from five different Mayan languages, lexical new mentions in agent (transitive subject) role typically occupy about three percent of the total lexical new mentions. Therefore, it is very rare to encounter an indefinite transitive subject noun, for discourse reasons. K'ichean languages appear to be creating a syntactic rule that reflects the same constraint.
>
> (England 1991: 484)

Thus the ergative discourse profile, present in Mamean and K'ichean languages and presumably in their common ancestor, is undergoing incipient grammaticization in K'ichean but not Mamean. This means the K'ichean change cannot be considered deterministic. Nevertheless, in moving from soft constraints to hard constraints, it traverses a path from ergative discourse to ergative grammar.

2.3.4 Interim Conclusions

Is Preferred Argument Structure universal? The evidence so far paints a complex picture (2.5), and a full answer must await further inquiry (2.6). But what seems clear is that something like the ergative discourse profile represents a "recurrent pattern," appearing again and again in the discourse of children just learning their language—or exposed to no language at all—and remaining constant in languages widely separated by space, time, genealogy, and typology. Whether the complexity and variation evident in the research record can be subsumed under a single set of generalizations remains to be seen. For now, it seems productive to explore the view that the ergative discourse profile, (p. 41) and the Preferred Argument Structure constraints that motivate it, reflect enduring properties of language in use, with consequences for cognitive processing, acquisition, typology, diachrony, and grammaticization.

Having introduced several parts of the problem, it is time to see how they fit together, and to assess what it would take to provide a functional explanation for ergativity.

2.4 Explaining Ergativity

Preferred Argument Structure is claimed to be universal, and to motivate the grammaticization of ergativity. So far the functional explanation for ergativity seems on track—until we stop to think about accusative languages. If the ergative discourse profile is present as a motivating force in all languages, why aren't all languages ergative? "A system-external functional force, once appealed to, cannot simply be turned off at will" (Du Bois 1985: 353). To explain ergativity, one more piece of the puzzle must be introduced: competing motivations (Du Bois 1985, 2014; MacWhinney et al. 2014; Malchukov 2014). In the arena of discourse, the two top competitors are powerful and pervasive, even universal: topic continuity and new information.

The grammaticizing power of topicality is widely acknowledged. Subject is said to be a grammaticization of topic, or topic-cum-agent (Givón 1983a; Comrie 1988). But what about languages that don't have subjects, or at least don't grammaticize them in the central way that accusative languages do? Speakers of ergative languages have been claimed to differ from speakers of accusative languages in having "different conceptions of prototypical agenthood" and "different basic topicalizations" (Plank 1979: 28). Curiously, these psychological claims were made in the absence of any experimental or corpus evidence. To put such speculations to the test, a corpus-based study of topic continuity was devised (Du Bois 1987b: 842–843), which showed that in the ergative language Sakapultek, topical referents (appearing in two successive clauses) overwhelmingly favor S=A continuity over S=O continuity, 80 percent to 20 percent. The idea that ergative speakers think differently about topicality is debunked. But this leaves us with a quandary. Admitting that ergative speakers track topics just like accusative speakers do, it is still only accusative languages that grammaticize topic in an all-encompassing syntactic and morphological subject. If speakers of all languages track topic continuity in the S=A groove that moti-

vates accusative alignment, why aren't all languages accusative? The typological question comes full circle.

Both questions receive the same answer: competing motivations. Ergativity and accusativity are both motivated, each with its own dedicated motivation operating at all times in all languages. But only one motivation can prevail at a time, in organizing the basic structure of a grammar (or part of a grammar) in a given language. The discourse profiles that drive this eternal competition were hinted at already, in the discussion of the discourse excerpt in (1) (2.1). Within a single short stretch of discourse, two (p. 42) recurrent patterns are observed to coexist: first, introduction of new referents is managed in O and S (absolutive); and second, topic continuity is managed in A and S (subject). Quantitative cross-linguistic evidence shows the two discourse profiles are widely distributed, found together in diverse genres and across languages of all types. As usual, S is Janus-faced: Variable in discourse, it becomes pivotal in grammar.

Having seen some key pieces of the ergative puzzle, while acknowledging the competition from accusativity, we can now ask how it all fits together—to articulate the discourse motivation for ergativity, at least, if not a complete explanation for ergativity. The Quantity and Role constraints of Preferred Argument Structure set broad limits on the information structure of the clause. In effect they define a gross template for any argument structure construction, including intransitive, transitive, and ditransitive clauses. In the arena of discourse, utterance tokens realizing such constructions undergo selection to satisfy the constraints, producing the "facts on the ground" of the distribution of argument structure configurations in the utterance population. The ergative discourse profile represents a generalization of the statistical distribution of recurrent patterns, as candidates for grammaticization in the grammar of a given language. At the same time, the accusative discourse profile, lurking in the very same utterance population, presents its own patterns as alternative attractors for grammaticization. In general, the distribution of New information (reflecting cognitive processing demands) motivates a discourse pattern isomorphic to ergative–absolutive grammar. The distribution of topic continuity motivates a discourse pattern isomorphic to nominative–accusative grammar. Both functional pressures are present in the discourse of all languages, but at any given point only one can grammaticize, determining the syntactic alignment of a specific argument structure construction. If just one syntactic alignment is to prevail, the competition must be resolved. And this is just what grammaticization does: It resolves competitions, converting functional motivation into normative motivation (Du Bois 2014: 280). What emerges is a grammar that may seem arbitrary in its specific forms and normative rules, but that works for its users, serving as a unified framework for communicative practice and cognitive affordance.

Consider the Mayan case, which shows what can happen when the ergative discourse profile interacts with word order, setting up the conditions for the emergence of ergativity. First, in line with the ergative discourse profile, the single lexical argument (reflecting the Quantity constraint) typically occurs in either O for transitives, or S for intransitives (reflecting the Role constraint). Second, for Mayan languages going back to Proto-Mayan,

the dominant word order (Dryer 2013a) is verb-initial. Taken together, these two factors place both lexical arguments to the right of the verb, producing a structurally consistent V-Lex$_{\{S/O\}}$ or V-Absolutive word order (an ergative order in the sense of Dryer 2013b). This constitutes the gross structure of verb and noun, maintained consistently in the discourse profiles of most if not all Mayan languages from Proto-Mayan to now. But Mayan languages also have a fine structure of pronominal clitics, implemented in head-marking of both ergative and absolutive on the verb. How does the fine structure of agreement morphology interact with the gross structure of word order syntax?

(p. 43) As noted earlier, different parts of the cross-referencing paradigm behave differently. First and second person referents are given information, and thus are regularly marked by overt (non-zero) pronouns/clitics, in both ergative and absolutive. For 3rd person, referents in A are typically given, and thus reduced, and potentially cliticized. But S and O are often new, hence lexical, and so are not likely to be reduced or cliticized. The absolutive mention, being expressed overtly in the clause with a lexical rather than pronominal mention, is not itself cliticizable. The absence of a clitic may then be interpreted as absolutive zero agreement. Once the pronoun-to-agreement shift takes places (presumably in pre-Proto-Mayan), the features that define the Mayan ergative complex (V-Lex word order, head-marking, absolutive zero, and the ergative discourse profile) create a kind of V-Lex lock-in. This proves to be an evolutionary stable strategy, resistant to change over four millennia.

Along with the gross structure of information, there is also the fine structure of inflectional detail, semantic nuance, and the specificity of pragmatic interpretation. The gross structure defined by Preferred Argument Structure sets broad constraints, but leaves speakers and languages plenty of room to maneuver at the level of fine structure. This is where the precise details of morphosyntactic analysis and historical development become critical, as processes of utterance production, interpretation, analogy, reanalysis, and grammaticization deploy and reconfigure the fine points of grammar to serve the exacting needs of language users. The grammarian’s, semanticist’s, and historical linguist’s attention to detail comes into its own here, analyzing the fine structure to elucidate the precise accounting of grammar and meaning, which is indispensable to the functionality of language. But the gross structure matters too. In the end there is no need to choose between them. Gross and fine work together in all of human action, and language is no exception.

Languages often show an apparent harmony between gross structure (e.g. information structure, topicality, word order) and fine structure (e.g. inflection, agreement). But what if a discord or disruption arises between gross structure and fine—how is it to be bridged? Must it be? To locate the critical arena where the crisis comes to a head, look to word order in use. This is where discourse-and-grammar research may contribute to resolving certain mysteries about the grammaticization of syntactic alignment. Preferred Argument Structure motivates the gross configuration of lexical nouns, e.g. the ergative discourse profile. If these nominal elements are arrayed on the same side of the verb as in most Mayan languages, perhaps motivated by a formal and/or functional analogy, the

result is a recurrent word order configuration (e.g. V-Lex$_{\{s/o\}}$ or Lex$_{\{s/o\}}$-V) that constitutes a potential model for ergativity. If another historical development (say a phonological merger or analogical leveling) collapses key distinctions between elements in the agreement system, this may undermine the fine structure of morphological agreement. Yet what remains is the gross structure. Now the ergative distribution of gross structure becomes a potential model upon which to rebuild a fine structure, this time along new lines, perhaps ergative. Whether such a trajectory can be confirmed or disconfirmed in a language family with a long written history remains an open question. But the only way to find out is to follow the patterning of language in use. This (p. 44) means documenting the discourse profiles that define the gross structure of the relevant utterances. It may be rare to find suitable conditions for inquiry into both gross and fine structure at the required level of detail, but it is well worth seeking out.

The analysis of the Mayan ergative complex in relation to the ergative discourse profile provides an important case study illustrating certain general principles of functional explanation, insofar as it combines universal generalizations about gross structure of the ergative discourse profile with locally specific generalizations about fine structure of the inherited morphosyntax of a particular language family.

Extending this approach to other language families, each with its own unique history, one should expect three things. First, the layer of gross structure produced by principles like the ergative discourse profile has a discourse-based coherence of its own, and is likely to remain relatively stable; yet word order developments can introduce fundamental changes even here. Second, the layer of fine structure (inflection, agreement, and so on) linked to the language-specific inherited morphosyntax has its own logic of continuity and change, which operates in part independently of that of gross structure. Third, interactions between gross structure and fine structure may trigger a dynamic of change that disrupts the ecology of grammar, setting in motion events that lead to restructuring the system of grammar. The story of the interaction among gross and fine must be discovered anew in the history each new family. Yet even here, analogy, reanalysis, and adaptive selection (based on cognitively motivated information processing constraints, for example) offer general principles for a theory of grammaticization.

The roster of motivations with the power to shape ergative and other syntactic alignments is not exhausted by the two considered here. A more complete account will have to incorporate interactions with competing and converging factors such as event structure, causal chains, verb semantics, tense and aspect, voice, the constructional repertoire, evidentiality, analogy, and more (DeLancey 1990; Croft 1998; see also the chapters in this volume). Many of these involve morphosyntactic, semantic, and pragmatic fine structure, which is essential to complete the functional explanation of ergativity. A critical task for future research is show how multiple layers of fine structure interact with layers of gross structure to shape the grammaticization of ergative and other argument structure configurations.

2.5 Objections and Refutations

The claim of a discourse basis of ergativity (Du Bois 1987b) has generated a certain amount of controversy, which I address in this section. Some objections question the existence of the ergative discourse profile, or try to explain it away, while others accept it but doubt it supports a functional explanation for ergativity. While some studies raise useful points that warrant attention in future research, others reveal a misunderstanding of the nature of discourse-based explanation. Common conceptual errors are essentialism, reductionism, and epiphenomenalism, all of which involving idealization of the (p. 45) facts of language use. In the following I will try to distinguish the useful critiques from the dead ends; and, for the latter, suggest an alternative approach. A recurrent theme is that generalizations about language must be grounded in linguistic realism rather than idealization.

2.5.1 Diversity

One of the earliest and most productive critiques was developed in a series of discourse-and-grammar studies by Mark Durie (1988, 1994, 2003). Preferred Argument Structure studies had adopted the A–S–O grammatical categories (Dixon 1979) for their neutrality with respect to typological diversity in syntactic alignment, to avoid imposing alien categories on ergative languages. But Durie argued they were doing just that to active languages. He rejected the view of S as a universal category for linguistic analysis, arguing that it obscured important diversity in both discourse and grammar, in active languages like Acehnese (Austronesian, Indonesia). Here, intransitive subjects are internally differentiated, both in grammar and in discourse profile, between S_a and S_o. A better analysis of Acehnese discourse could be achieved by respecting the alignment typology evident in Acehnese grammar, which distinguishes Actor (S_a=A) and Undergoer (S_o=O). By investigating categories relevant to the grammar at hand, Durie was able to show that each discourse profile was both internally consistent and distinct from its counterpart. Subsequent studies in ergative, accusative, and especially active languages have confirmed and extended these findings. For example, Chol, a Mayan language once characterized as “split-ergative” but now recognized as active, makes a similar distinction between Actor and Undergoer in both grammar and discourse profiles (Vázquez Álvarez & Zavala Maldonado 2013). These studies challenge the assumption that S is uniform in language use, and provide a useful corrective to the A–S–O schema. More generally, they remind us that the categories of the language being investigated are a potent guide to new discourse profiles waiting to be discovered. This raises the question of whether there are different Preferred Argument Structures for active, ergative, and accusative languages, or even finer-grained distinctions; and, if so, whether they can be interpreted as principled variations on a single unifying theme.

Durie nevertheless saw value in framing his analysis in terms of Preferred Argument Structure, once the necessary adjustments were made to accommodate the diversity of grammar and discourse profiles of the language being analyzed. Durie recognizes a key point that befuddles some critics, regarding the role played in discourse-functional expla-

nation by the discourse profile (Du Bois 2003a: 40–44), what I have previously called "recurrent clusterings in parole" (Du Bois 1985: 357). Discourse profiles are generalizations about recurrent patterns of linguistic behavior, including recurrent co-occurrences of pragmatic, semantic, and grammatical features. As statements about language use, they are not to be confused with the grammatical categories they motivate, which have a normative status in the linguistic system. Durie nicely points up a (p. 46) common misconception about discourse-functional motivation of grammatical categories in a response to Herring (1989) which remains relevant today:

> Herring (1989), misunderstanding the concept of motivation, regards the mismatch in Sakapulteko as a logical flaw in Du Bois's argument. A further distortion in Herring's argument derives from her demand that the grouping of S and [O] should involve functional specialization in terms of cognitive or semantic factors. In Du Bois's account, the cognitive/semantic motivations do not themselves directly impact upon coding choices, but only *proximately*, through the "recurrent clusterings in parole" (Du Bois 1985: 357) which they produce. This is of course a key difference between the Du Bois account of ergativity and more directly cognitively or semantically based approaches.
>
> (Durie 2003: 192, fn. 193)

Discourse profiles are general statements about the "facts on the ground" of language use. They arise as a result of multiple factors, including factors such as cognitive processing, salience, and so on. As system-external motivations they interact with system-internal factors to shape the emergence of grammar. But the process is not deterministic. While discourse profiles influence the grammaticization of linguistic categories, the profile is not the category.

2.5.2 Essentialism

Less productive are critiques which, while bringing no empirical research to bear on the question at hand, translate originally statistical observations into the language of categorical statements, the better to draw logical deductions—or contradictions. Needless to say, the meaning of a discourse profile may be severely compromised if its empirical variability is disregarded. When soft constraints are restated as hard constraints, the result is unlikely to be faithful to the original.

Such a confusion mars the critique of Harris and Campbell (1995), who paint a portrait of the discourse basis of ergativity (Du Bois 1987b) that is almost unrecognizable to someone familiar with the theory. They confidently appeal to the "spirit" of Preferred Argument Structure constraints, which they feel are "blatantly violated" by the passive origin of ergativity (1995: 253). They enthusiastically assert that transitive "A [is] prohibited almost totally from introducing new information" (1995: 253), then quietly delete the "almost" as they substitute a categorical rule for the original soft constraint. They perceive "an unpleasant circularity in Du Bois' picture" (1995: 254), based on their questionable reframing of Nora England's findings of incipient grammaticization of certain Preferred

Argument Structure constraints into grammatical constraints on the ergative in K'ichean (England 1991; England & Martin 2003). Their story of the passive origins of ergativity begins with a proto-language agented passive, represented as containing two proper nouns (*Mary was.hit by Jane*) (1995: 252). This contrasts with the attested (p. 47) examples they themselves cite from Old Persian to illustrate the passive to ergative change, which contain a pronoun rather than a lexical noun in the agentive *by*-phrase (1995: 244, 255). Though somewhat murky, the reasoning behind their argument seems to depend on the following assumptions and inferences:

1. To get from accusative to ergative alignment, a well-known diachronic pathway is via the reanalysis of an intransitive passive construction as a transitive active construction.
2. To express a two-place predication fully explicitly, the passive sentence cannot be agentless, but must include the agent.
3. Thus the agented passive construction is used, with the agent expressed in an oblique role, e.g. a prepositional phrase adjunct (*by*-phrase).
4. Obliques (prepositional phrase adjuncts) often contain new information and lexical nouns (by Preferred Argument Structure).
5. Therefore the oblique agent must have been new and lexical.
6. The two-participant predication, being a passive, is intransitive.
7. Therefore the subject of the two-participant predication (the semantic Patient) must be an S.
8. Intransitive S arguments are often new and lexical (by Preferred Argument Structure).
9. Therefore the S role argument must have been new and lexical.
10. By reanalysis, the *by*-phrase (originally an optional adjunct) is reinterpreted as an obligatory core argument (transitive subject A), and the S becomes O.
11. The formerly rare agented passive intransitive construction undergoes a massive change in frequency to become the new normal for transitive constructions, but speakers make no changes in their use of pronouns vs. nouns, or given vs. new information.
12. Now there are two new and lexical nouns in the core arguments of the clause (violating Preferred Argument Structure).
13. Now there is a new and lexical A (violating Preferred Argument Structure).

The problems with this account are several, arising on multiple levels. First, the logical flaws. The reasoning depends on converting a statistical tendency to a categorical rule. This must be done twice (in deriving 5 from 4, and 9 from 8), in order to generate the desired contradiction. Second, language users are presented as being sufficiently creative to introduce major changes to the structure of their grammar, reanalyzing an optional oblique as an obligatory core argument (step 11), but they were apparently too timid to begin using a pronoun in place of a lexical noun—during the decades or centuries it must have taken for the frequency shift and ergative reanalysis to be completed. Ignored is the fact that there are no constraints against using pronouns or given information: Not only are the Preferred Argument Structure constraints always soft constraints, the limits are

always upper limits, not lower (Du Bois 1987b: 834; 2003b: 73). (p. 48) Thus there has never been any minimum requirement to fill a syntactic slot, whether argument or adjunct, with either new information or a lexical noun.

One begins to understand why it was necessary for Harris and Campbell to appeal to the "spirit" of Preferred Argument Structure in making their argument, instead of employing the actual constraints. The result is unfortunately fairly typical of essentialism, which, faced with statistical evidence of diversity in argument realization, substitutes a categorical stereotype, and then uses the stereotype to work out the logic of its reasoning. But surely this is antithetical to the "spirit" of Preferred Argument Structure, if it must have one.

Is there an alternative? Actually, very little needs to be changed in the above account to make it clear that not only is Preferred Argument Structure compatible with the grammaticization pathway in question, it actually facilitates it. Moreover, the only changes needed involve replacing the rigidly idealized conjectures of Harris and Campbell with common-sense observations on how speakers talk—as confirmed in corpus-based studies of naturally occurring language use. Preferred Argument Structure allows speakers the flexibility to use a pronoun in any slot, whether argument or adjunct, and speakers routinely do just that (Du Bois 1987a, 2003b). Moreover, as Ariel et al. (2015) have shown, speakers use pronouns when innovating a new argument slot, precisely because pronouns fly under the radar of the Quantity constraint. This leaves speakers free, when extending the use of a structure like the passive, to choose a pronoun in the *by*-phrase. In fact this is precisely what the textual evidence from Old Persian shows:

(2)
ima tya manā kartam Parθavaiy
"This is what was done by me in Parthia"
(Darius the Great) (John R. Payne 1980: 151) (cited in Harris & Campbell 1995: 255)

The genitive pronoun (*manā* 'me.GEN') expresses the agent in a *by*-phrase, initially an optional adjunct. But as this structure comes to be routinely used to express two-place predications, it undergoes reanalysis (step 11) as an ergative core argument. The new transitive structure easily satisfies the Quantity and Role constraints of Preferred Argument Structure.

2.5.3 Reductionism

One attraction of essentialism is that it feeds reductionism. By replacing a complex empirical generalization with a handy rule of thumb, a theoretical economy is achieved. If the rule is not only categorical but universal, the conclusion may be drawn that empirical investigation of the language at hand is unnecessary. But when statistical generalizations are replaced with categorical statements, the likely result is a false economy.

Such a chain of essentialist substitutions is found in the reductionist proposal of Haspelmath (2006). Where I had pointed out the functional motivation linking the avoidance of lexical nouns with avoidance of new information (Du Bois 1987b: 829–830), Haspelmath wants to go further, suggesting that "since new information is **mostly** coded (p. 49) by lexical NPs, … the [lexical] tendencies … could be **reducible** to the [new information] tendencies." Moreover:

> [T]he 'quantity' tendencies … [may] **follow straightforwardly** from the 'role' tendencies.
>
> (Haspelmath 2006: 910)
>
> [I]t appears that the strong **tendency** to avoid clauses with two new/lexical core arguments can be **reduced** to the strong avoidance of new/lexical As … .
>
> [W]e simply like to talk about human beings and their actions, so animates **tend** to be topical. … Since the A argument is **almost invariably** animate, it **follows** that it is **typically** topical and **hence** given.
>
> (Haspelmath 2006: 911, emphasis added)

Haspelmath concludes that "the whole notion of preferred argument structure may be **reducible** to other, well-established tendencies and generalizations" (2006: 911, emphasis added); (see also Dahl 2000: 50; Everett 2009; Haig & Schnell 2016).

This style of argument slides easily from most to all, from it tends to it follows, and draws inferences accordingly. While it overlooks the fact that the Quantity constraint probably has a better independent motivation as a cognitive limitation on information processing resources, Haspelmath's proposal to dispense with the Quantity constraint, and indeed all of Preferred Argument Structure, makes a certain sense—if we accept a series of inferences about language use, each apparently well motivated on its own:

1. Humans are interesting to humans.
2. Therefore, humans are topical.
3. Therefore, humans are given.
4. Therefore, humans are expressed by reduced forms (pronouns or zeroes).

Leaving aside the inferential gaps, this is still not enough to make viable predictions about argument realizations in syntactic argument slots. If you add the further assumption that humans are agents, and that the subject role expresses agent-cum-topic, you can, seemingly, derive the inference that subjects are expressed by pronouns. In contrast, direct objects are said to encode inanimates. Inanimates are less interesting, therefore not topically continuous, therefore not accessible, therefore new, therefore lexical. But do these conclusions actually follow?

Goldberg evaluates a similar reductionist proposal and rejects it: "[T]he Given A constraint does not follow directly from the prevalence of animate topics. … [T]he Given A

constraint is not simply epiphenomenal" (Goldberg 2004: 431). The problem with the essentialist–reductionist line of reasoning is that the conclusions follow only if each of a long chain of assumptions is valid—specifically, only if each generalization is categorical. But none of them are. And because each step in the chain represents at best a statistical tendency, the inferential failure compounds with each step. Flawed logic aside, the most serious problem is empirical: the observable facts of discourse do not confirm the logical deductions about what discourse "should" look like. The fact is, only some (p. 50) humans are topical, and therefore pronominalizable, and this makes a critical difference for grammaticization pathways, e.g. in the role of pronouns in the emergence of innovative argument structure constructions (Ariel 2000; Ariel et al. 2015).

Is there another way? A recent corpus-based study of Hebrew datives (Ariel et al. 2015) offers a relevant model, even if the construction involved is different. Ariel et al. compare datives (syntactic arguments) with adjuncts headed by the preposition *bishvil* 'for.' Both introduce mostly humans, and both express the same thematic role (roughly, benefactive). Yet they part company in their discourse profiles: Only 5 percent of the datives are lexical, but as many as 23 percent of the *bishvil* adjuncts are. The difference cannot be attributed to animacy, but only to the difference in syntactic status: argument vs. adjunct. This accords with the predictions of constructional Preferred Argument Structure (Ariel et al. 2015: 270–272; Du Bois & Lester in progress), which freely allows lexical mentions in adjuncts, but not in core argument slots other than S/O. More generally, pronouns sidestep any problems with the One Lexical Argument constraint, and thus are exploited as pioneers in an incipient grammaticization introducing an additional argument position into the clause, allowing a second human participant (in addition to the agent) to be expressed in a benefactive-like role.

What are the implications for ergativity? While the alignment types and constructions in question differ, interesting parallels can be drawn nonetheless. In both cases, speakers modify an existing argument structure construction, adding a new argument role (ergative or dative). And the strategy they adopt to accomplish it without violating Preferred Argument Structure constraints is essentially the same: cognitive containment (Ariel et al. 2015: 270–272; Du Bois & Lester in progress). The safe strategy is to use a pronoun in the innovative argument slot, to avoid violating the One Lexical Argument constraint—or, to put it in cognitive–functional terms, to avoid overloading limited processing resources.

2.5.4 Historical Accident

Claims of functional motivations have implications for language change, but are sometimes at odds with the findings of historical linguists (Cristofaro 2014). For example, Næss (2015) points to a series of seemingly random historical changes in rejecting, not only the competing motivations analysis of ergativity (citing Du Bois 1985; Du Bois 1987b), but any functional explanation for the ergative structure of Äiwoo (Oceanic, Solomon Islands). After detailing the complicated historical processes that produced the ergative verb phrase, he states:

> What the [ergative] VP structure of Äiwoo demonstrates is … that any linguistic system at any point in time is the outcome of a number of interacting factors which do not add up to either a universal structural template or a set of functionally-based competing motivations.
>
> (Næss 2015: 102)

(p. 51) This statement about grammar as "the outcome of … interacting factors" is fine up to the conclusion, which does not follow. The assumption that the claimed VP merits its own "set of functionally-based competing motivations" seems premature, in the absence of the relevant discourse profile information. Be that as it may, it is surely unwise to conclude that the randomness of historical changes precludes an adaptive account of the historical evolution of linguistic structure. From the perspective of linguistic as well as biological evolution (Beckner et al. 2009; Lane 2015: 172–204), there is no principled contradiction between the fact that a system is subject to historical processes, which may randomly introduce partial arbitrariness, and the applicability of the evolutionary processes of adaptive selection, which yield partial motivation.

The origin of ergativity seems particularly prone to invite frustration and despair, leading some historical linguists to go so far as to question whether there is any functional motivation for ergativity:

> The absence of a clear case of extension creating ergativity argues against a clear functional motivation unique to the ergative pattern. With the exception of the passive > past/perfective ergative, it appears that ergativity is a historical accident that has come up again and again in many parts of the world.
>
> (Gildea 2004)

To conclude that "ergativity is a historical accident," while acknowledging that it is a recurrent pattern arising independently in languages around the world, seems a contradiction. Anything in language can be made to look like a historical accident—even the grammar of accusative languages. But such a stance appears valid only if one restricts the inquiry to tracing etymological sources and describing the arbitrary signs that result. What looks like a "historical accident" may well turn out to involve adaptive selection, given an evolutionary account.

What is the alternative? The evolutionary development of any aspect of language can be seen as the result of many small, locally motivated actions, taken by speakers who lack an overarching view of the system (Keller 1994). But in this respect linguistic evolution is no different from biological evolution (Dediu et al. 2013; Lane 2015; Mayr 2001). Evolutionists don't give up on adaptive explanation just because random mutations introduce one accident after another. On the contrary: Such accidents (mutations) provide the necessary variability (Bybee 2007) for selection to act on. On one interpretation, functional constraints play a role in grammaticization by acting as selective processes that winnow the variability of naturally occurring discourse. The forms and constructions that survive

and reproduce in the longer term will be those that satisfy cognitively motivated constraints, like Quantity. In general, constraints on the selection of functionally viable linguistic structures can act over time to constrain the possible outcomes of grammaticization processes, leading to functionally motivated structures (Hopper & Traugott 2003; Traugott & Trousdale 2014).

(p. 52) 2.5.5 Epiphenomenalism

Epiphenomenalism may seem an arcane philosophy, originating as it does in nineteenth-century mind–body dualism (Walter 2009: 1137). But despite its abandonment in most modern sciences, in linguistics (and in sociobiology, Searle 2013) the epiphenomenon is invoked surprisingly often (Hopper 1987; Jaeger & Snider 2008; Malchukov 2008). This includes in claims about ergativity (Everett 2009; Haig & Schnell 2016). In practice, labeling an empirical observation as epiphenomenal typically prefigures a move to dismiss it as inconsequential. "An epiphenomenon is a *secondary symptom*, a mere "afterglow" of real phenomena" (Walter 2009: 1137). While the supposed epiphenomenon is acknowledged to have a real cause, it is claimed to have no further causal consequences in the world (Walter 2009: 1137). By setting up a disconnect between mind and materiality—or langue and parole—epiphenomenalism inherits the problems of dualism (Searle 2013). Not surprisingly, attempts to partition facts into real phenomena and epiphenomena tend to suffer from a lack of consensus about criteria for deciding which is which: One scholar's epiphenomenon is another's phenomenon. But the real problem lies in the idealization that is introduced into otherwise empirical questions: a dualism, not of mind and body, but of grammar and use.

Appeals to epiphenomenality often arise in response to claims about language use as an influence on grammar. An empirically observable pattern in use is said to be epiphenomenal—in effect, not a phenomenon, but only illusory—to the extent that it can be explained away as caused by other factors—in a word, reduced (Haspelmath 2006). But this neglects the first question that should be asked: Does the observable pattern in language use have downstream causal consequences? Specifically, the epiphenomenal charge has been made regarding discourse patterns identified by Du Bois (1987b) as consequential for the grammaticization of ergativity (Everett 2009; Haig & Schnell 2016). In a study otherwise notable for its careful analysis and impressive multilanguage database,[9] Haig and Schnell, speaking of the Given A Constraint, maintain that:

> [T]he apparently marked behavior of the A role, another cornerstone of the ergativity claims, … is an epiphenomenal by-product of two well-documented and robust tendencies: the pervasive tendency for transitive subjects to be [+hum], and the pervasive tendency for all subjects (S or A) to be topical, hence given information.
>
> (2016: 612)

Their claim to distinguish some patterns as epiphenomenal (while others, presumably, are not) leads them to conclude, somewhat surprisingly, that ergativity itself is epiphenomenal:

(p. 53)

> In Iranian for example, a shift from accusative to ergative alignment (restricted to past tenses) … was a particular, and highly contingent, combination of … changes that conspired to yield ergative alignment … . These and similar diachronic developments speak of a more contingent approach to ergativity, according to which ergativity arises as an epiphenomenal and construction-specific constellation, through the combination of essentially independent morphological and phonological processes.
>
> (Haig & Schnell 2016: 614–615)

Here the idea of epiphenomenality begins to collapse on itself. Against claims of functional explanation, the epiphenomenalist proposes an absence of explanation: "ergativity is a historical accident" (Gildea 2004) or "epiphenomenal" (Haig & Schnell 2016: 615). True, a grammatical pattern may seem accidental, when viewed exclusively in terms of its etymological source materials. But some historical accidents are accidents waiting to happen. As a typologically recurrent pattern, ergative syntactic alignment must be recognized as a powerful attractor state, that is, an evolutionary stable strategy. Accusativity too is a potent attractor. But there is no contradiction here. Both recurrent patterns are attested worldwide, and each has its motivations. Indeed, the two motivations coexist within the discourse of every language. The eternal tension between ergative and accusative motivations, evident in split ergativity, is best understood in terms of the theory of competing motivations (Du Bois 1985, 2014; MacWhinney et al. 2014). But to dismiss a linguistic pattern—even ergativity—as epiphenomenal is to cut off inquiry prematurely. It would be better to drop the epiphenomenal stance altogether, and take up the very real challenges of building theory and method capable of accounting for recurrent discourse patterns and their downstream consequences for grammar. In the study of ergativity, the facts on the ground of discourse hold much promise for understanding split ergativity (Silverstein 1976; Comrie 1978; Dixon 1979, 1987; DeLancey 1981; Durie 1988, 2003; Malchukov 2005, 2014; Law et al. 2006; Mahand & Naghshbandi 2014; see also Laka, Nash, Coon & Preminger, and Woolford Chapters 7, 8, 9, and 10, this volume).

2.5.6 Interim Conclusions

I have considered various objections to Preferred Argument Structure and the ergative discourse profile. The most productive tend to come from researchers who combine in-depth first-person research on the grammar in question with detailed and sensitive empirical investigations into the discourse distribution of grammatical elements (Durie 1988, 1994, 2003; Vázquez Álvarez & Zavala Maldonado 2013). Others were found wanting: mired in essentialism, reductionism, and epiphenomenalism.

What is the problem? Not generalization per se, which is indispensable for understanding and explaining grammar. Rather, the problem lies in idealization, cutting the system of language off from the reality of its use. Idealization begins with a misplaced essentialism that reifies categories, obscuring the variability that characterizes (p. 54) populations of naturally occurring utterances. It continues (sometimes) with a reductionism that creates an illusion of economy, without testing to see if the reduced principles can in fact reconstitute (predict) the facts of the world it claims to have reduced. On the other hand, sometimes scholars dwell on a maze of intricate historical details—no lack of empirical facts here—but when the time comes for an explanation they come up empty, claiming a historical accident. Finally, the apotheosis of idealization is reached in epiphenomenalism, which dismisses certain facts as not phenomena, banishing them to the black hole of epiphenomena—from which no downstream causal consequences can escape.

What is the alternative? First, linguists must commit to linguistic realism. The facts on the ground of discourse are not to be dismissed, lest their consequentiality in the world of language be overlooked. Nor are they to be shunted aside as epiphenomena. Rather, language use is firmly located in the world. Here it coexists and interacts with the practices, norms, and knowledge of language, even if all these have somewhat different ways of being in the world. Linguistic realism urges documenting the empirical generalizations about language use that define its capacity to shape grammar: the discourse profile. Second, it is equally important to follow through on the theoretical end. Discourse profiles have downstream causal consequences. Identifying the consequences serves at the very least to verify that a meaningful discourse profile has been identified. The combination of linguistic realism and theoretical generalization is critical for explaining ergativity, and for all questions of the interinfluence of discourse and grammar.

While some objections have proved lacking, even so they serve to elicit clarification of issues left murky in previous formulations, and point to gaps in our knowledge that call for further research. Important issues have been broached regarding the relation of language use to grammar, and of linguistic realism to functional explanation. Certainly some aspects of the discourse approach to ergativity are likely to remain controversial, until resolved by further research. It remains for new collaborations among researchers, bringing together corpus-based methodologies, multifactorial statistical techniques, grammar, typology, competing motivations, and functional theory, to advance our understanding of the outstanding questions.

2.6 Directions for Future Research

What new possibilities does discourse research bring to the study of ergativity? Whether the topic is pragmatics, syntax, semantics, constructions, splits, morphology, prosody, cognition, typology, diachrony, or grammaticization, all can benefit from incorporating a focus on language in use. The study of discourse integrates well with a wide range of research questions, bringing new perspectives on how a targeted set of grammatical resources serves the communicative goals of its users.

For ergativity the biggest payoff is likely to come from “discourse inside the clause” (Du Bois 2003a: 13; 2003b: 83), an approach that seeks to map out the distinctive functional (p. 55) correlates of the various structural components of the clause. A useful research strategy is to follow the trail of difference, in both discourse and grammar. Differences in linguistic structure are shaped by the multiplicity of functional needs, modulated by complex interactions among competing motivations, and driven by the dynamics of grammaticization. Interesting structural differences arise between contrasting elements in split grammar, variation, typology, child language, diachrony, and grammaticization, all potentially linked to distinctive functions. The world’s languages offer a rich set of natural experiments (Evans & Levinson 2009), where each case represents a potentially novel way of linking contrasts in form to contrasts in function. Integrating a discourse methodology into cross-disciplinary research can bring a key piece of the puzzle, helping to trace out the similarities and contrasts that manifest in such fundamental differences as, for example, ergativity vs. accusativity. For discourse to become an integral and valued component of such research, it must identify the specific discourse profiles that are relevant to the problem at hand.

One ongoing challenge is to tease apart grammar and discourse: grammar with its seemingly static structure, and discourse with its free plasticity. The challenge is compounded, yet all the more important, because grammar and discourse are so intimately intertwined. For untangling the multiplicity of factors that impinge on the discourse distribution of referential forms across syntactic roles, a much-needed development is the application of newer multivariate statistical techniques (Du Bois & Lester in progress). The increasing popularity of soft constraints (Bresnan et al. 2001) as a way of describing morphosyntactic and functional variation makes the development of statistical models all the more urgent for corpus-based ergativity research. Such techniques are needed to address questions that have been raised about functional correlates of variation in ergative marking, for example in exploring Duranti’s observations about how the so-called “optional ergative” is exploited in discourse to index social power and access to agency (Duranti & Ochs 1990; Duranti 2004). Similarly, Dixon, observing that discourse profiles differ in detail from one ergative language to the next, asks how this may reflect differences in cognitive style between their respective cultures (Dixon 1994: 219–220).

Yet perhaps the most productive questions, offering a combination of the most effective methodological purchase and the greatest theoretical rewards, will be those that explore the diversity of constructions that characterize the grammar of ergative and other languages. The constructional approach is particularly well positioned to shed light on issues of split ergativity, with large theoretical implications. The exploding interest in the grammar of constructions provides a useful framing for new investigations of the discourse profiles of a wide variety of constructions, each potentially revealing some new aspect of the structural and functional diversity of language. Constructions combining aspects of argument structure, valence, reference, person, voice, tense, aspect and so on are promising topics for new research on syntactic alignment that incorporates a discourse-and-grammar dimension. For the targeted argument structure construction, it will be important to map out its information structure, as revealed in its distinctive discourse

profile, operating within the relevant functional niche (Du Bois 2014). (p. 56) A further step is to work out how such construction-level discourse profiles interact with the semantics of the verbal repertoire; inherited morphosyntax; functional strategies for utterance processing; and other factors that combine to motivate the grammaticization of ergative and other alignments. A particularly incisive version of this problem concerns the three-way interaction between split ergativity, split function, and split discourse, where each new language's "natural experiment" differentiates contrasting elements within the scope of a single language. Insights gained from split ergativity can extend even to languages that do not themselves overtly exhibit these splits, insofar as they provide clues to concealed patterns and discourse profiles that reveal the incipient fault lines of potential change.

For linguists who are prepared to use the combined tools of discourse and grammar to investigate the diversity of natural experiments in the world's languages, the potential for discovery is open-ended. A valuable research strategy is to (1) document a construction, or better, a contrasting set of grammatical constructions; (2) identify their respective discourse profiles; (3) clarify the connection of each to its corresponding functional niche; and (4) explore the ramifications for grammaticization. Such a research agenda is well positioned to expand and refine our understanding of the dynamic processes that motivate the grammaticization of the structures of ergative languages in all their diversity, with broad implications for understanding all forms of grammar.

2.7 Conclusions

Amidst increasing interest in the complexity, diversity, and heterogeneity of grammar in all its forms and functions (Beckner et al. 2009; Page 2010; Evans 2013), the challenge of coming to terms with ergativity takes on special significance. Ergativity has long revealed an uneasy fit with the conventional categories of standard theories of grammar, including the supposedly fundamental categories of subject and object, or even A, S, and O. Ergativity calls into question the universality of such preconceived categories; yet neither can its own indigenous categories of ergative and absolutive be set up as universal structures in their stead. Ergativity and its variations and competitors challenge the linguist's desire for easy answers.

One way to engage with this complexity is by attending to the variability, and the recurrence, of patterns of language use. For example, the ergative discourse profile has been documented in a typologically diverse array of languages around the world: ergative, active, and accusative. To be sure, it does not stand alone: competing discourse forces of topicality strongly motivate ergativity's main competitor, the accusative type. Yet whether expressed in grammar or only in discourse, the robustness of the ergative discourse profile remains, attested in its continuity across grammatical typology, historical change, child language, and genre. Its presence, sometimes variable and often submerged, is nevertheless revealed in small ways, shaping the child's earliest productions and the occasional grammaticized construction even in accusative languages. While it (p. 57) cannot

constitute a complete account of ergativity on its own, neither can any explanatory account of ergativity be complete, realistically, without incorporating the “facts on the ground” of the ergative discourse profile, with all its consequences. As a counter to reductionism, essentialism, and idealization, the empirical trend of modern linguistics invites attention to the explanatory power of corpus-based evidence. Typologically aware research on discourse profiles documents the complexity and diversity of vast populations of utterances—but also the consistency that gives them power to shape the adaptive evolution of grammars.

Surely the functional, structural, and historical basis of ergativity is more complex than is envisioned in any one current analysis. A full explanation of ergativity and its variants and alternatives will require the corpus-based identification of multiple competing and converging motivations, and their integration into a larger explanatory framework. Relevant forces include the distribution of given and new information across clausal arrays of argument slots, the semantics of force dynamics in the clause, the lexicon of verbs and argument structure constructions, the learning and reanalysis of inherited morphosyntax, recurrent pathways of change and grammaticization, and more. All impact the dynamics of the discourse profile, and all come together to shape the emergence of the system of grammar.

Ergativity is too important to be left to the specialists of ergative languages. It is a problem for all linguistics, and a useful one at that. Ergativity invites linguists to investigate the most fundamental structures of grammar, not only in ergative languages but in every language; and to revisit questions of function and structure, of universality and diversity, that were once thought to be settled. As the field of linguistics turns more and more to evidence-based analysis, traditional methods of elicitation are increasingly supplemented by the empirical tools of corpus-based and experimental methods. There is much work to be done to document the discourse profiles that shed light, not only on the syntactic alignments that broadly define ergative, accusative, and active languages, but also on the details of the rich constructional repertoires of more specialized argument structure constructions, such as passives, antipassives, perfectives, and nominalizations, that may serve as bridges for diachronic realignment from one structural type to another. Ergativity has the potential to disrupt conventional thinking and existing explanations for grammar. Let the disruption begin.

Acknowledgments

I wish to thank Jessica Coon and Shobhana Chelliah for comments on an early draft that have greatly improved this chapter. For helpful input on specific points I thank Matthew Dryer and Danny Law. For conversations that have deepened my thinking on this and other questions of linguistic explanation, I am profoundly indebted to Mira Ariel.

(p. 58) Abbreviations

The following symbols and abbreviations are used in glosses: CP, completive aspect; DEP, dependent; DIM, diminutive; ICP, incompletive aspect; FOC, focus; LAT, lative; PL, plural; TV, transitive verb; 1.ERG, ergative 1st person singular; 3.ERG, ergative 3rd person singular; 3.ABS, absolutive 3rd person singular. In addition to the standard A–S–O symbols introduced in the first paragraph, the following are used in the syntactic schemas: X, oblique/adjunct NP; V, verb; P, preposition.

Transcription symbols (Du Bois 2014b; Du Bois et al. 1993) include: ; [semi-colon] speaker label; … pause; .. micro-pause; : [colon] prosodic lengthening; (H) in-breath; , [comma] continuing intonation; . [period] final intonation. Intonation units are indicated by a carriage return (one IU per line).

Sakapultek data are presented in a practical orthography (Du Bois 2006: 198), in accordance with standards of the Academy of Mayan Languages, with the following special values: *x* voiceless laminopalatal fricative; *j* voiceless velar fricative; *q* voiceless uvular stop; *tz* voiceless apicoalveolar affricate; *nh* voiced velar nasal stop (engma). Apostrophe (’) following a consonant represents a glottalized consonant; following a vowel it represents a glottal stop. Double vowels represent phonemic length.

Notes:

([1]) Note that this annotation focuses on referent mentions and how they are expressed. Thus a referent expressed with a lexical noun phrase plus a cross-referencing affix in the same clause is treated as one mention, not two (Du Bois 1987b: 813). Here it always the heavier form (noun phrase) that is marked (with boldface).

([2]) The underlying *r-* is devoiced in this phonological environment, coalescing with preceding voiceless fricative *x-* and effectively disappearing in the surface form, yielding *xya’* ‘he gives it.’

([3]) The implications of V-Lex as a kind of ergative–absolutive word order (Dryer 2013b) are developed below in “Typology” and “Diachrony” (2.3).

([4]) Figures 2.1–2.4 are based on selected studies in Table 2.1 (see also tables 2.1–2.4 of Du Bois 2003b: 62–69).

([5]) “The general processes and principles which can be noticed in observable history are applicable in all stages of language history” (Hock 1991: 630).

([6]) It bears noting that in several language families, older texts show higher lexical densities than modern texts (Ashby & Bentivoglio 2003), while still conforming generally to Preferred Argument Structure. This may reflect differences in genre of the older texts (epic poetry for Old French and Old Spanish, dynastic history for hieroglypic Maya), for which the modern languages lack a common counterpart.

([7]) The main exception is Chol, which has innovated an agentive system (Law, Robertson, & Houston 2006; Coon 2013; Vázquez Álvarez & Zavala Maldonado 2013). A few Mayan languages show some split ergativity, but this is a late and partial development (England 1983; Law 2009, 2014), which coexists with ergative morphosyntactic alignment. Yet one line is never crossed: No Mayan language has ever been accusative, from Proto-Mayan till today.

([8]) Discourse evidence from a fifth branch is provided for Chol (Vázquez Álvarez & Zavala Maldonado 2013), which seems to have a discourse profile similar to that identified by Durie for Acehnese (Durie 1988, 1994, 2003), with $A=S_a$ opposed to $S_o=O$ in both grammar and discourse profile. Whether Durie's conclusion, that Acehnese is compatible with an (extended) Preferred Argument Structure account, can be applied as well to the Chol data is beyond the scope of this chapter.

([9]) The massive study by Haig and Schnell (2016), encountered online as this chapter was going to press, makes a valiant effort to raise the bar statistically, and merits a more complete assessment than can be given here. Nevertheless, it introduces conceptual problems of its own, touched on here. Note that they cite different numbers of subjects and objects for transitive clauses (e.g. for English they report 1,111 transitive objects but only 422 transitive subjects; see their appendix 2). The gap reflects their omission of 1st and 2nd persons, making it difficult to compare with the findings reported here.

John W. DuBois

John W. Du Bois is Professor of Linguistics at the University of California, Santa Barbara. A specialist in functional linguistics, discourse, Mayan languages, and sociocultural linguistics, his work centers on the interaction between discourse and grammar. He has long been interested in the complex functional competitions that drive the emergence of grammar as a complex adaptive system, yielding the extraordinary typological diversity of argument structure constructions and syntactic alignments in the world's languages. His publications include Competing Motivations (1985), The Discourse Basis of Ergativity (1987), Preferred Argument Structure (2003), Discourse and Grammar (2003), Motivating Competitions (2014), and Towards a Dialogic Syntax (2014).