

GRAMMATICALIZATION IN A SPEAKER-ORIENTED THEORY OF CHANGE

Henning Andersen
University of California, Los Angeles

In this paper I descend the conceptual ladder from Bernd Heine's definition of grammaticalization and the change types of which it is comprised to the innovation types that give rise to such changes and the innate grammatical premisses—the 'laws of language'—that condition speakers to make such innovations. Along the way, (1) I aim to present grammaticalization in the context of other types of macro-change, (2) to survey and illustrate types of change in content, content syntax, expression, and morphosyntax, (3) to show how individual changes can be analysed into complexes of a few types of basic innovations, and, finally, (4) to suggest the special importance in change of innovative reanalysis and the role played in reanalysis and actualization by principles of markedness.

0. Introduction

Throughout the twentieth century the study of grammaticalization (Gzn) has been divorced from the reality of the transmission of language traditions among speakers.¹ To Meillet (1912), Tauli (1958), and Kurylowicz (1965) alike, language was a supraindividual, social phenomenon with no clearly defined locus. To most scholars that have been concerned with Gzn in more recent decades, whether functionalists or formalists, the role of speakers in the development of Gzns seems not to have been important or even relevant. The majority view among both categories of scholars, it seems, is that it must be good enough to study diachronic correspondences—stages along the 'Gzn cline'—or grammar fragments. Yet no aspects of language change can be explained unless they are grounded in an explicit understanding of the way language patterns are passed among individuals.²

¹ This paper was prepared at the Centre for Advanced Study at the Norwegian Academy of Sciences in Oslo during academic 2004–2005. I am pleased to acknowledge the support of the Centre and my gratitude to the Centre's staff for a perfect work environment. I also wish to thank my colleagues at the Centre and participants in the meetings at Rosendal and Lysebu, Norway for comments on preliminary versions of these pages.

² Besides Gzn (grammaticalization), the following abbreviations are used: ACC (accusative), AOR (aorist), Br. (Belarussian), CS (Common Slavic), Da. (Danish), DAT (dative), dial. (dialect(al)), dist.. (distant), DU (dual), E (Early), EModE (Early Modern English), Eng. (English), Est. (Estonian), ex. (example), F (feminine), Fr. (French), FUT (future), GEN (genitive), Gk. (Greek), Gm. (German), IMPF (imperfect), IMPV (imperative), INST (instrumental), It. (Italian), L (Late), La. (Latvian), Lat. (Latin), Li. (Lithuanian), LOC

In this paper I will descend the conceptual ladder from the Gzn Schema most recently described by Heine (2003) (chapter 1) and the change types it includes and some it does not (chapter 2) to the innovation types that initiate all such changes (chapter 3) and, finally, to the innate premisses (the ‘laws of language’) that condition speakers to make such innovations (chapter 4).

My aim is to make clear the distinct levels of observation represented by (i) change schemas (sec.1), (ii) historical changes (sec.2), (iii) innovations and change scenarios (sec. 3), and (iv)—as examples of the bases of grammar formation—the role played by iconicity and markedness in the origin and transmission of language variation, and hence, indirectly, in change (sec. 4).

Along the way, particularly in section 2, it will become clear, I hope, that if we examine the changes that actually occur in language histories, we will find that there is a great deal more that demands our attention than Gzn and Gzn chains alone would suggest.

1. Grammaticalization and other macro-changes

Since the 1800s, historical linguists have observed, described, and endeavored to explain grammatical changes on two levels of observation—as individual changes and as long-term developments. These two levels have not always been distinguished, but in a general discussion of Gzn it is important to recognize (i) that Gzn is one of a number of types of long-term development (macro-changes) and (ii) that the individual changes that together form a Gzn chain represent a small, though important, subset of actually occurring types of grammatical change.

There is no tradition for defining different types of long-term development, but it may be useful to distinguish types on the basis of the observer’s wider or narrower focus. Here I will define three kinds of such ‘macro-types’ in these terms.

1.1 The whole-language view

The whole-language view focuses on major typological properties of language structure. The earliest attempt to view language developments in this

(locative), M (masculine), M (modern), ME (Middle English), m-ness (markedness), Mid (Middle), Mod (Modern), NOM (nominative), NT (neuter), NW (northwest), O (Old), OE (Old English), ON (Old Norse), PERF, PF (perfect), PL (plural), PLUP (pluperfect), Pol. (Polish), PRES (present), Pre-S (Pre-Slavic), R, Russ. (Russian), sec. (section), SG (singular), SW (southwest), Ukr. (Ukrainian).

perspective is perhaps the application of August W. Schlegel's (1818) three language types by Humboldt (1822), who sketched a development from isolating through agglutinating to inflectional morphology and assumed that this reflected the cultural progress of the Indo-European peoples. This simple idea was left in the dust by Whitney (1870), who understood the role of linguistic creativity and the constant interaction of decay and growth and introduced the notion of *renewal*, the pivotal element in any understanding of long-term developments. In the whole-language view, this made it possible to form the more plausible theory of the *Morphological Cycle*, described by Hodge (1970). This cycle is thought to be universal. It is almost fully attested in the development of Chinese and can be posited for Egyptian by extrapolating from the morphology of the verbal categories. In table 1, s = syntactic expression, m = morphological expression, and capitals represent the predominant technique, syntactic (analytic) or morphological (synthetic), in the expression of verbal categories.

Table 1. The Morphological Cycle

proto-Afroasiatic	Old Egyptian	Late Egyptian	Coptic
*Sm	sM	Sm	sM

1.2 The subsystem view

A focus on subsystems serves to capture alternating periods of elaboration and simplification of individual systems such as categories as cases or tenses, aspects, and moods.

The Indo-European system of morphological cases, for instance, appears to have gone through a development of elaboration through four (Hittite, Greek, Germanic) to the classic six (Sanskrit, Slavic) to nine (East Baltic) and subsequent simplifications in Greek, some of Slavic, much of Romance and Germanic, Latvian, and Lithuanian.

The Russian system of tenses and aspects, as it appears through internal reconstruction and the textual record, was elaborated through the development of new aspect distinctions from maybe two tense paradigms some time in prehistory through three (new Imperfect) to eight (periphrastic retrospective tenses and Future) in Old Russian through five in Middle Russian to three (present, past, periphrastic future) in the modern standard language; it has more recently been elaborated to five in some northwest Russian dialects; Andersen MS.b. In table 2, periphrastic tense–aspect paradigms are separated from inflectional paradigms by a line. Morphologizations can be observed between

CS-2 and CS-3 (IMPF) and between OR-2 and MidR-1 (PAST). The table ignores two aspect distinctions which are derivational, the all-pervasive perfective vs. imperfective and the minor determinative vs. indeterminative aspect, which is relevant only to verbs of locomotion.

Table 2. Slavic tense–aspect systems. Elaboration and simplification

	Pre-S	CS-1	CS-2	CS-3	OR-1	OR-2	MidR-1	MidR-2	St. R	NW-R
Simple tenses	PRES	PRES	PRES	PRES	PRES	PRES	PRES	PRES	PRES	PRES
	IMPF	IMPF	AOR	AOR	AOR		PAST	PAST	PAST	PAST
	AOR	AOR		IMPF	IMPF					
	PERF									
Com- pound tenses			IMPF		FUT	FUT	FUT	FUT	FUT	FUT
					PERF	PAST	PLUP	PLUP		DIST. PAST
					PLUP 1	PLUP	FUT PF			PLUP
					PLUP 2	FUT PF				
					FUT PF					

1.3 The single element view

The single-element view allows for two possibilities, a focus on content or on expressions.

1.3.1. The focus on changing expressions for individual content categories can be called onomasiological. Consider the content “animal” in OE *deor* > ME *beast* > EModE *animal*. Change in the words for one content naturally raises questions about other contents, e.g., “deer” in OE *heorot* > ME, ModE *deer*.

In grammatical change, there are cases like 1.3.1.1 and 1.3.1.2, where a given content category alternates (toggles) between having separate morphological expression and having none. And there are cases where one can trace changing expressions for a seemingly stable category that is repeatedly renewed through time (1.3.1.3).

1.3.1.1. The vocative in Russian. In Old Russian, vocative and nominative are distinct. In Middle Russian, vocative merges with nominative. In the 1900s a new, distinct vocative form develops for some given names in one declension.

1.3.1.2. Definiteness in East Slavic. In Pre-Slavic there is no definiteness category. In Common Slavic, definiteness is expressed by attributive adjectives only. In Middle Russian, definiteness is lost through reanalysis. In North Russian dialects, an enclitic demonstrative is reanalysed as definiteness marker. In some North Russian dialects, definiteness is lost as this enclitic is reanalysed as a focus particle.

1.3.1.3. In Romance, the Future tense is renewed, changes from analytic to synthetic, and is renewed again, e.g., Lat. *dīcam* > LLat. *dicere habeo* > Fr. *je dir-ai* > *je vais dire*.

1.3.2. The alternative focus in the single element view is on changing content combined with a continuity of expression; this can be called semasiological. Consider ME *bede* “prayer” > ModE *bead* “globule, bead”. The continuity of expression does not preclude expression changes. Hodge (1970) dubs such continuous, changing expressions ‘chronomorphs’.

In grammatical change, the most obvious examples of semasiological etymology are Gzn chains, developments in which the content of an expression changes from lexical to grammatical or from grammatical to more grammatical (Kuryłowicz 1965) with attendant change in content syntax and typically (but not necessarily) entailed changes in expression and morphosyntax.

1.3.3 *Grammaticalization*. A great deal of progress has been achieved in Gzn studies during the last quarter of a century, not least in quantitative terms, as great numbers of examples has been amassed from languages around the world (Heine and Kuteva 2002). They document both that our predecessors—from Humboldt to Meillet and Kuryłowicz—understood Gzn correctly and, on the other hand, that there is more to it than they suspected. At the same time, the simple-minded ‘Gzn cline’ of the 1990s (1.3.2.1), which blithely confuses content change and morphosyntactic change, has been superseded by advances by both functionalists and formalists. Among the former there are statements such as Heine’s (2003:578–579) in 1.3.2.2, which by its form acknowledges that Gzn concerns linguistic signs with content, content syntax, expression, and expression syntax (morphosyntax). Formal syntactic analyses of Gzns such as those by Roberts and Roussou (2003) and van Gelderen (2004) focus on changes in syntactic function, but reveal the distant relation between such macro-changes and strategies of grammar formation. Van Gelderen (2006) takes up the development of single category cycles and describes a number of these, formulating largely homological schemas for the Negation Cycle, the Aspect Cycle, and the CP Cycle and motivating these cycles by reference to three principles of universal grammar that are manifested as learners’ strategies.

1.3.3.1. Lexical word > grammatical item > clitic > affix > Ø³

³ Somehow many historical linguists who accepted this ‘cline’ did not notice that it confuses content, morphosyntax, and expression: lexical > grammatical is a change in content, word > clitic > affix is a development in morphosyntax, and item > Ø refers to phonological attrition (unless zero content is held to be more grammatical than grammatical content).

1.3.3.2. “[T]he grammaticalization of linguistic expressions involves four interrelated mechanisms: ...

(i) desemantization (or ‘bleaching,’ semantic reduction): loss in meaning content;	[HA: content
(ii) extension (or content generalization): use in new contexts;	content syntax
(iii) decategorialization: loss in morphosyntactic properties characteristic of the source forms, including the loss of independent word status (cliticization, affixation);	morphosyntax
(iv) erosion (or ‘phonetic reduction’), that is, loss in phonetic substance.”	expression]

The role of strategies of grammar formation in linguistic change is so obviously important that it is worth taking a minute to reflect on the different approaches in van Gelderen (2004, 2006) and Roberts and Roussou (2003).

Van Gelderen subscribes to the idea, otherwise particularly widespread among functionalists, that speakers prefer their languages economical. This perennially attractive idea is incompatible with the actual synchronic complexity, not least in fine-grained detail, of every language that has so far been investigated. And although it appears many historical changes illustrate this idea, it appears to be contradicted by the rest.

Roberts and Roussou infer from their examples of Gzn and other changes that language acquirers are cautious (“conservative”) of nature and preferentially avoid complexity in favor of defaults. Their account imagines that acquirers produce abrupt changes, and it completely ignores the observable fact that in every language the new coexists with the old in the form of synchronic variation. It is based on the mistaken identification of diachronic correspondences and the changes that give rise to them—of ‘before and after’ relations with the actual historical developments that connect them.

If the macro-types were all there is to grammatical change, our work would soon be done. Alternatively we could change our focus from schematic representations to actual historical changes, which present a somewhat more complicated picture, but may be a better source for an understanding of grammatical change. Chances are that we might gain greater insight not only into the multitude of factors that condition historical change, but especially into the nature of the acquirers—the analytic procedures, expectations, and preferences they bring with them, and which mature in their minds, during the process of language acquisition.

2. Types of grammatical change

2.0 Type and token

I will begin by urging the importance of the traditional distinction of type and token. Let us accept, at least for the moment, Heine's characterization of Gzn in 1.3.3.2 as a description of a type of change or rather, since it comprises a chain of changes, a macro-type. I will refer to this as the Gzn Schema. Now, if 1.3.3.2 characterizes a macro-type, actual historically attested or reconstructed Gzn chains are tokens or instantiations of this type.

Making the type–token distinction gives us several advantages.

First of all, it enables us to compare the Gzn Schema with other change schemas that are similarly instantiated in historical developments. This may not seem very exciting, but comparing such macro-types of change is useful if we want to understand what kinds of change are possible in linguistic systems, how they arise, and what drives them. In phonology, for instance, there is an established dichotomy between change types in which (phonetic) information is, over time, reduced and other types in which information is, over time, increased. Lenition and Monophthongization fall in the former category, Fortition and Diphthongization in the latter. It is then interesting to note that the Gzn Schema as a whole, as well as its constituent parts, as described in 1.3.3.2, fall into the reductive category of change types, and one can wonder if there are grammatical changes that fall into the opposite category.

There is a second advantage to making explicit the type vs. token distinction: it allows us to recognize both full and partial instantiations of the Gzn Schema, that is, to recognize as Gzns some changes that do not fully correspond to the ideal type the Gzn Schema represents. This is already being done in current literature and is, in fact, explicit in Kuryłowicz's (1965) definition of Gzn, which covers both changes from lexical to grammatical content and changes from a grammatical to a more grammatical element. There are several varieties of 'incomplete' Gzn chains. There are (i) potential chains of which only fragments are attested (2.0.1), (ii) developments that are in progress but have not run their course yet (2.0.2), and possibly (iii) tokens of arrested development, a phenomenon well known from historical dialectology. In addition to partial instantiations that call for a historical explanation, such as these, there are instantiations that may be partial for typological reasons. For instance, in a language that does not have inflectional morphology, we can expect Gzns to yield full words or clitics, but perhaps not inflectional affixes. Or, again, in a language where a certain grammatical category is consistently expressed by full words, new members of that category that arise through Gzn

may remain full words and not develop into clitics or affixes (2.0.3). In short, the type–token distinction enables us to form a flexible and realistic, but still principled understanding of Gzn.

An additional advantage the type vs. token distinction gives us in a sense derives from the second. The recognition that there are tokens of Gzn that are not full Gzn chains encourages us to look more closely at the individual kinds of change comprised by the macro-type the Gzn Schema defines. If we do this, we discover, first, that they can occur relatively independently of one another, and, secondly, that there are other kinds of grammatical change in the histories of languages that apparently have nothing to do with Gzn, but obviously cannot be ignored.

To put it differently, there is a whole world of grammatical change beyond the confines recognized in the Gzn Schema. This will be explored in chapter 2.

2.0.1. *Partial Gzn chains.* The notion ‘partial Gzn chain’ may be difficult to accept for some. It may be useful to compare it with the phonological macro-type Lenition. Complete lenition chains are easy enough to document, e.g., ON $-k >$ MidDa. $-g >$ ModDa. $-\text{ŋ} > -j > -\emptyset$ (ON *ek* ‘I’ $>$ ModDa. *jeg* [jaɛ] \sim [ja]). But while the Lenition Schema describes a sonority scale extending from voiceless tense plosive to voiced vocoid to zero, a token of lenition is defined as any change by which a segment on this scale changes into a segment to its right. Hence historical phonologists readily recognize any partial chain such as $k > g$ or $g > j$ as an instantiation of lenition. A parallel approach in the study of Gzn speaks of such chain fragments as demonstrative $>$ determinative or auxiliary $>$ particle simply as Gzns; see, for instance, Roberts and Roussou (2003), van Gelderen (2004).

2.0.2. *Incomplete Gzn chains.* Consider the contrast between Polish and Serbian–Croatian. In Polish, the auxiliary “be” of the compound past tenses became an enclitic, assigned to Wackernagel’s position, in the Middle Ages; but then its forms were reanalysed as mere person and number markers, and over the last six hundred years these have gradually shifted to the position after the erstwhile main-verb participles, which have long since been reanalysed as finite forms (2.1.3.2; 2.4.1). In Serbian–Croatian, by contrast, the auxiliary “be” of the compound past tenses became an enclitic and was assigned to Wackernagel’s position in the Middle Ages, where it continues to be placed to this day (Browne 1993: 345).

2.0.3. *Partial Gzn chains.* Consider the Gzn of OE *in deed* to ModEng. *indeed* (Traugott 2003) Through this change *indeed* becomes a member of the paradigm of sentence adverbs. These all typically occur contiguous to major sentence constituents, are full words, and carry main stress. *Indeed* shares these characteristics. Although it has been grammaticalized, there is no reason to expect it will undergo the morphosyntactic and expression changes described in the Gzn Schema.

In the traditional approach to Gzn, there is a consensus that the central part of a Gzn, its crux, is the change in content, from lexical to grammatical, or from grammatical to more grammatical content, as Kuryłowicz (1965) put it.

Kuryłowicz's definition implies a simple dichotomy of categories of content, lexical and grammatical. Both lexical and grammatical categories form paradigms (2.0.0.1), and content changes can easily be divided into (i) changes into, (ii) changes within or among, and (iii) changes out of lexical or grammatical paradigms.

The crux in a Gzn is a change by which an element enters a grammatical paradigm or a change within or among grammatical paradigms. In order to be able to discuss such changes independently of the other kinds of change involved in the Gzn Schema, I propose to call these content-change types *grammations* and *regrammations*. Changes by which an element shifts out of a grammatical paradigm, that is, changes in which grammatical content is lost, I will call *degrammations* (2.0.0.2). We will see several subtypes of each of these major types of grammatical content change below (2.1.1–2.1.3).

In Gzn, the content syntax of the grammatized or regrammatized constituent typically changes: its scope increases, it is *upgraded*. *Downgrading* occurs as well in grammatical change, in connection with both regrammation and degrammation. In some instances, the upgrading of one constituent is accompanied by the downgrading of another, what you might call a *content-syntactic permutation* (2.2.1–2.2.3).

What is often called 'phonetic erosion' is better understood more generally as *expression reduction*, for in actual fact, there are several ways in which expressions are commonly reduced in grammatical change. On the other hand, cases of expression *elaboration* are not uncommon. Other notable expression changes are expression *substitution* and *doubling* (2.3.1–2.3.4).

Morphosyntactic changes by which free forms change in the direction of affixes, *bond strengthening* changes, typically occur subsequent to grammation and regrammation. Changes in the opposite direction, *bond weakening* (*emancipation*) may reflect other conditions (2.4.1–2.4.2). Besides these, there are changes of grammatical indexes (2.4.3) and of element order (2.4.4).

2.0.4. *Paradigm*. The term *paradigm* is used here and elsewhere not in the narrow sense of 'inflectional paradigm', but in the general sense of "selectional set", a usage that has been traditional since Saussure.

2.0.5. The triple *grammation*, *regrammation*, *degrammation*, which will be defined and exemplified below, is broadly analogous to the terms *phonologization*, *rephonologization*, *dephonologization*, introduced by Jakobson (1931 [1971]). In phonologization, a sound type becomes phonemic, that is, part of a phonological paradigm; in rephonologization, a sound type changes its relation to other phonemes, that is, changes from one phonological paradigm to another; in a dephonologization, a sound type loses its distinctive status and becomes a variant (positional or stylistic) of another phoneme. For lexical change, the term *lexicalization* is standard; what could be called *relexicalization* (changes within or among lexical paradigms)

is simply called *semantic change* (the hypernym for “meaning change”); so, too, is *delexicalization* (loss of lexical content), which includes ‘bleaching’ as a special case.

2.1 Content changes

2.1.1. *Grammation*. I define grammation as a change by which an expression, through reanalysis, comes to carry grammatical content. (Change from any other, including zero, content to grammatical content.) This definition covers both changes that are naturally associated with the Gzn Schema and changes that are not.

A change such as the one in 2.1.1.1 is indistinguishable from the initial change in many Gzn chains. In Danish, modal verbs undergo little or no phonological reduction other than the loss of stress that is characteristic of auxiliaries and of lexical verbs that incorporate a subject, a direct object, or an adverb. But by its loss of stress, the modal does become a clitic. In due course it may become subject to contraction with a following constituent (in effect, stem shortening), that is, expression reduction, as some modals are.

Grammations such as this are reversible (2.1.3.1). It is important to understand that the reversibility of such changes logically has no bearing on the unidirectionality of Gzn. A degrammation, such as the change in 2.1.3.1 is not a Gnz. Hence it is neither an example of, nor a counter-example to, the unidirectionality of Gzn.

Greenberg (1991) has suggested the distinction between grammations from above, such as the one in 2.1.1.1 and from below, as in 2.1.1.2 (Greenberg uses the term *grammaticalization*.)

2.1.1.1. *Grammation*. Lexical > modal verb. Danish. In the 1700s, the lexical verb Da. *burde* “ought” is apparently reanalysed as a modal verb. In this grammation its content changes little, but the verb ceases to combine with direct-object clauses and instead—like other modal verbs—combines with entailed clauses. As a consequence, it becomes able to incorporate infinitive clauses without the infinitive-marker *at*; the incorporation is expressed by its losing its lexical stress.

E.g., *Hvis I ikke er Hanreeder, saa burde I at være det* “If you are not a cuckold, you ought to be one” vs. *I burde [Ø] være noget andet [end kandestøber]* “You ought to be something different [from tinker]”; examples from plays by Ludvig Holberg (1684–1754), written ca. 1722. By the mid-1800s, the use of infinitive-*at* with *burde*-*bør* “ought” is archaic-poetic: *Hvor ingen Plov kan gaa/ og ingen Le kan slaa/ der bør et Træ at staa* (ascribed to Enrico M. Dalgas, 1828–1894) “Where no plow can go,/ and no scythe can mow,/ there ought a tree to grow (lit.: stand)”. Examples from *Ordbog over det danske Sprog*, s.v. *burde*.

Content: lexical > modal verb. Content syntax: object infinitive clause > entailed infinitive clause. Expression: full stress > weak stress. Morphosyntax: *at*-infinitive > bare infinitive and incorporation. Note the reverse change in 2.1.3.1.

2.1.1.2. *Grammation from below*. Morphophonemic alternant > grammatical morpheme. Russian. Middle Russian consonants are consistently palatalized before the tense mid front vowel /ê/ (a result of historical change which is irrelevant here). In the paradigm of *dv-a* “two; M-NT”, *dv'-ê* “two; F” this ‘automatic alternation’ creates an ambiguity. The desinence *-ê* can be taken as the primary expression for NOM-ACC.F, and the palatalization of the preceding stem-final consonant is then a phonologically conditioned alternant; this seems to have been the persistent reanalysis in the majority of Russian dialects, including the standard; table 3. Alternatively the palatalization can be taken as the primary expression for the feminine and the *-ê* as a conditioned alternant; this seems to have been the reanalysis in southwest Russian dialects and in dialects of Belarus. This alternative reanalysis is tantamount to a grammation of stem-final palatalization in NOM-ACC.F; subsequently, palatalization is extended to all case forms of “two”, which thereby come to show gender concord in these dialects; Avanesov and Orlova (1965); Karskij (1956: 242).

Table 3. Morphophonemic alternant > grammatical expression in Russian dialects

	MidRuss.		ModRuss.		SW Russ. dial., Belarus	
NOM-ACC	<i>dv-a</i> M-NT	<i>dv'-ê</i> F	<i>dv-a</i> M-NT	<i>dv'-ê</i> F	<i>dv-á</i> M-NT	<i>dv-í-é</i> F
GEN-LOC	<i>dv-ux</i>		<i>dv-íx</i>		<i>dv-íx</i> M-NT	<i>dv-í-íx</i> F
DAT	<i>dv'-êma</i>		<i>dv-ím</i>		<i>dv-ím</i> M-NT	<i>dv-í-ím</i> F
INST	<i>dv'-êma</i>		<i>dv-um'á</i>		<i>dv-umá</i> M-NT	<i>dv-í-umá</i> F

The resulting paradigm is unique in these dialects in two regards: it is the only one in which gender concord is expressed in plural forms, and it is the only one in which a nominal gram is expressed by the single distinctive feature of palatalization; the feature can be regarded as a stem suffix as shown in the table.

Content: unchanged (F gender). Content syntax: unchanged (concord marker). Expression: *-ê* “F” > *-í-* (palatalization feature) “F”. Morphosyntax: (i) permutation: desinence > allomorph; automatic alternant feature > stem suffix; (ii) stem suffix extended to oblique forms.

2.1.2. *Regrammation*. A change by which a grammatical expression, through reanalysis, is ascribed different grammatical content. (Change within or among grammatical paradigms.) There are many kinds of regrammation. The examples below show a split (2.1.2.1), whole recategorizations, nominal (2.1.2.2) and verbal (2.1.2.3), and a ‘double’ regrammation (2.1.2.4). Several other examples are discussed in Askedal (2006), Haug (2006).

2.1.2.1. *Regrammation*. Collective > plural (derivation > inflection). Russian. Some Old Russian collective nouns with the suffix *-j-* > ModR plurals. E.g., *brat'-j-a* COLL.F.SG “brothers; brethren” > *brát'-j-a* M.PL, *zub'-j-e* COLL.NT.SG “teeth, cogs” > *zúb'-j-a* M.PL “cogs” vs. M.PL *zúb-y* “teeth”. Other collectives with the *-j-* suffix remain (e.g., *bab-j-ó* “womenfolk”); this is technically a lexically conditioned split.

Content: “collection of X” > “X-PL”. Content syntax: upgrading, lexical > grammatical pluralization (both with PL agreement). Expression: unchanged *-j-*. Morphosyntax: productive derivational *-j-* “collective” + SG desinences > lexically constrained stem suffix, *-j-* “plural” + PL desinences.

2.1.2.2. *Regrammation*. Genitive case > determiner-phrase marker (inflection > derivation). English. The so-called *s*-genitive'. E.g., OE *dæs cyninges sweoster Ecgfrides* "king E.'s sister" > ModE *the king of Spain's daughter, someone else's cat*.

Similar change in Scandinavian languages, e.g., Danish: ... *et af Degnens Ordsprog i Ulstrup* (S. S. Blicher) "one of the church warden's sayings in Ulstrup" vs. *for Gud i Himlens Skyld* (S. Kierkegaard) "for God in Heaven's sake"; examples from Skautrup (1964).

Content: genitive case > DP marker. Content syntax: downgrading, case inflection > DP derivation; Heltoft (2001, 2005). Expression: OE *-(e)s* and other allomorphs > ModE */=z/*, etc. Morphosyntax: emancipation (2.4.2.5), word affix > phrase clitic (2.4.4).

2.1.2.3. *Regrammation*. Present perfect (retrospective present) > past. Russian. OR *ja.NOM.SG jes-m'*.PRES.1SG *da-l-ŭ.SG.M* "I have given" > *ja.NOM.SG jes-m'*.PRES.1SG *da-l-Ø.SG.M* "I gave"; at the same time, simple preterites are lost; Andersen (MS.b). Similar changes in other North Slavic languages and in German, French, etc.

Content: marked aspect (retrospective) > marked tense (past). Content syntax: no change. Expression: no change; later auxiliary > pronoun allomorph > Ø. Morphosyntax: no change; later participle > finite form, auxiliary > subject pronoun (2.1.3.3).

2.1.2.4. *Regrammation*, double. Russian. MidR pluperfect participle > finite past. Auxiliary > "nonresultative" procedural marker; e.g., *oná ujéxala bylá* "she had left" > *oná ujéxala=bylo* "lit.: she went away, but didn't" (i.e. she was about to leave, but didn't; or she set out, but came right back and didn't go); the finite verb is perfective and asserts a completed event; the procedural marker asserts the absence of a resultant state; Andersen (MS.bd).

Content: retrospective past > "nonresultative" past. Content syntax: upgrading (participle > finite past), downgrading (auxiliary > procedural adverb) (permutation). Expression: morphological and prosodic reduction ("nonresultative" loses agreement inflection and stress). Morphosyntax: bond strengthening (word > clitic).

2.1.3. *Degrammation*. A change by which an expression, through reanalysis, loses grammatical content. (Change from grammatical content to other, including zero, content.) Degrammation may be lexicalization (2.1.3.1), or it may be a change by which a grammatical element loses part of its grammatical content (2.1.3.2, 2.1.3.3) or all its content (2.1.3.4)

2.1.3.1. *Degrammation*. Modal verb > lexical verb. Danish. In the 1900s, Da. *turde* "dare" is reanalysed as a lexical verb. It comes to compete with the lexical *vove* "dare", which it tends to replace. As a lexical verb, it comes to combine with *at*-infinitives and to receive full stress. Example from a current style manual: "*Vi tør ikke helt at udelukke muligheden. Bedre: Vi tør ikke helt [Ø] udelukke muligheden*" (Jakobsen and Jørgensen 1999) "We dare not entirely exclude this possibility. Better without *at*".

Content: "dare" changes from modal to lexical paradigm. Content syntax: entailed infinitive clause > object infinitive clause. Expression: weak stress > full stress. Morphosyntax: bare infinitive > *at*-infinitive

Note the opposite change in 2.1.1.1. On Eng. *dare*, Beths (1999). Similar changes due to ellipsis (i) Penn.Gm. *wotte* "would" (past of modal *wellen*) > "wish, desire" (lexical) (Burridge 1998). (ii) Gm. *müssen* "must" > *müssen*₁ "must", *müssen*₂ "have to go to the toilet".

2.1.3.2. *Degrammation*. Auxiliary > participant marker. Polish regrammatizes present perfect as general past tense and loses imperfect and aorist. With the past tenses gone, the present-

tense auxiliary is degrammatized for tense. Its forms retain their content as person and number markers and drift from second position to become past-tense person and number endings. Contrast 2.1.3.3.

E.g., OPol. *kiedy=m.1SG niosł-a.SG.F, kiedy=m.1SG niósł.SG.M* > ModPol. *kiedy niosł-a-m.SG.F-1SG, kiedy niosł-em.1SG.M* “when I carried; F, M” (Andersen 1987).

Content: (auxiliary’s) tense, person, number > person, number (e.g., PRES.1SG > 1SG). Content syntax: unchanged. Expression: insignificant. Morphosyntax: enclitic > desinence; terminal verb forms are replaced by bound verb stems, here *niósł* by *niosł-*; (partial) prosodic univerbation with verb stems (not shown), unchanged word-boundary sandhi is now word-internal (not shown).

2.1.3.3. *Degrammation*. Auxiliary > personal pronoun. Russian, like Polish, regrammatizes present perfect as general past and loses imperfect and aorist. With these past tenses gone, the present-tense auxiliary is degrammatized for tense. Its forms retain their content as person and number markers and become allomorphs of personal pronouns. As such they are used as nominatives with present as well as past-tense verbs. They are later lost (2.3.3).

E.g., MidR ¹*kupilŭ* ²*esmi na torgu, a togo ži* ²*esmi* ³*ne znaju, u kogo* ¹*kupilŭ* “²I ¹bought [it] at the market, but ²I ³not know the one, from whom [I] ¹bought [it]” (²*esmi* is the former AUX.PRES.1SG); *a ženy ne* ¹*vidělŭ* ²*jesi* ³*budešŭ vŭ sně* “and ²you ³have (fut.) not ¹dreamt of [your] wife (²*esi* is the former AUX.PRES.2SG)” (Zaliznjak 2004: 179).

Content: tense, person, number > person, number. Content syntax: downgrading from VP to subject NP. Expression: unchanged. Morphosyntax: unchanged free forms.

2.1.3.4. *Degrammation*. Case form > stem allomorph. Danish. Here, before the loss of case inflection, several dozen fixed prepositional phrases were lexicalized with a case form (dat., gen.) governed by the preposition. The noun forms are no longer segmentable since no content can be ascribed the erstwhile desinences; the forms are (stem) allomorphs. (In the examples, | marks the former desinence boundary.)

E.g., Da. *ad år|e* “in the future, lit.: in a year”, *af gård|e* “off the farm”, *fra bord|e* “off ship-board”, *i sind|e* “in mind”, *på fod|e* “onto (poss.) feet”, *til lands* “on land”, *til bord|s* “to the table”; Skautrup (1944: 267).

Content: case meaning > “Ø”. Content syntax: case assignment > fixed form in lexicalized phrases. Expression: unchanged. Morphosyntax: boundary loss.

2.2 Content-syntactic changes

2.2.1. *Semantactic upgrading*, change from dependent to head or enlargement of scope, is typically entailed by grammation and regrammation. The Gzn literature is full of examples. Some examples in the preceding sections: 2.1.1.1 (lexical verb > modal), 2.1.2.1 (derivation > inflection), 2.1.2.3 (aspect > tense), 2.1.2.4 (participle > finite verb).

2.2.2. *Semantactic downgrading* typically accompanies regrammation or degrammation. Some examples in the preceding text: 2.1.2.2 (case inflection > DP derivation), 2.1.3.1 (modal > lexical verb), 2.1.3.3, 2.1.3.4 (auxiliary > participant marker).

2.2.3. *Semantactic permutation.* Several examples were seen above. In the double regrammation in 2.1.2.4, as the retrospective past is regrammatized as a non-resultative past, the pluperfect participle is upgraded to finite verb, and its auxiliary is downgraded to procedural marker. Other examples: 2.1.3.2 (auxiliary > verbal person-and-number marker), 2.1.3.3 (auxiliary > subject pronoun allomorph).

2.3 Expression changes

2.3.1. *Expression reduction.* What the textbooks call phonetic erosion is the creation and subsequent generalization of variants in which phonological features or segments are omitted (ellipsis; 3.1.1). The most frequent type, syntagmatic reduction (here, stem reduction) involves grammatical elements; note that there is no Gzn in 2.3.1.1, 2.3.1.2, where all the elements are grammatical to begin with. In 2.3.1.3, a vocative is grammatized as politeness marker with subsequent pragmatically conditioned reduction in specific environments; outside these the expression may remain unreduced (layering). In paradigmatic reduction (here desinence reduction), distinctions in a paradigm are reduced to a functional minimum (syncretism), typically more in marked categories than in corresponding unmarked categories; contrast F and M in 2.3.1.4.

2.3.1.1. *Syntagmatic reduction.* English. What the tradition insists on calling ‘contractions’ are examples of clitic shortening. E.g., /aɪ=m/ “I am”, /ʃi=z/ “she is; she has”, /wi=r/ “we are”, /wi=v/ “we have”, /wi=d/ “we had; we would”, /wi=l/ “we will”, /ar=nt/ “are not”, /kæn=t/ “cannot”.

Content: unchanged. Content syntax: unchanged. Morphosyntax: word > clitic. Expression: Reduction from syllable to rhyme.

2.3.1.2. *Syntagmatic reduction.* Polish. OPol auxiliary “be”, present. (i) {*jeśm*.1SG, *jeś*.2SG, *jest*.3SG, *jesmy*.1PL, *jeście*.2PL, *sa*.3PL} > (ii) {=*śm*.1SG, =*ś*.2SG, =∅.3SG, =*sm*.1PL, =*ście*.2PL, =*sa*.3PL} > (iii) {=*m*.1SG, =*ś*.2SG, =∅.3SG, =*śmy*.1PL, =*ście*.2PL, =∅.3PL}. Copula and existential “be” in (i) has been renewed as (iv) {*jest=em*, *jest=eś*, *jest*, *jest=eśmy*, *jest=eście*, *sa*} > (v) {*jest-e-m*, *jest-e-ś*, *jest*, *jest-e-śmy*, *jest-e-ście*, *sa*} (layering); Andersen (1987).

Content: present tense, person, number > person, number (degrammation; 2.1.3.2). Morphosyntax: word > clitic > affix. Expression: reduction by one syllable, leveling.

2.3.1.3. *Pragmatically conditioned reduction.* Russian. OR *gospodar’* “master” > *gosudar’* “sovereign; sire”, MidR *sudar’* > Pre-1917 =s “Sir”, e.g., *Gotóvy-s* “Yes, they are ready, Sir”; *Ták-s* “Right, Sir”.

Content: lexical > grammatical (politeness). Content syntax: lexeme > discourse marker. Morphosyntax: word > enclitic. Expression: reduction, tetrasyllabic > single onset.

2.3.1.4. *Paradigmatic reduction*. Russian. Adjective endings, M.SG and F.SG of *takoj* “that kind of, such a”. Loss of phonological features and segments in the bolded forms: loss of palatalization in desinence-final labials and three non-phonetic vowel changes in the feminine paradigm resulting in increasing syncretism—5 > 4 > 3 distinct endings.

Table 4. Paradigmatic reduction in Russian

	MidR				1700s		1900	
	MASC	FEM	MASC	FEM	MASC	FEM	MASC	FEM
NOM	<i>tak-ój</i>	<i>tak-ájja</i>	<i>tak-ój</i>	<i>tak-ájja</i>	<i>tak-ój</i>	<i>tak-ájja</i>	<i>tak-ój</i>	<i>tak-ájja</i>
ACC	=N/G	<i>tak-úju</i>	=N/G	<i>tak-úju</i>	=N/G	<i>tak-úju</i>	=N/G	<i>tak-úju</i>
GEN	<i>tak-óvo</i>	<i>tak-ójě</i>	<i>tak-óvo</i>	<i>tak-ójji</i>	<i>tak-óvo</i>	<i>tak-ój</i>	<i>tak-óvo</i>	<i>tak-ój</i>
LOC	<i>tak-óm'</i>	<i>tak-ójji</i>	<i>tak-óm</i>	<i>tak-ójji</i>	<i>tak-óm</i>	<i>tak-ój</i>	<i>tak-óm</i>	<i>tak-ój</i>
DAT	<i>tak-ómu</i>	<i>tak-ójji</i>	<i>tak-ómu</i>	<i>tak-ójji</i>	<i>tak-ómu</i>	<i>tak-ój</i>	<i>tak-ómu</i>	<i>tak-ój</i>
INST	<i>tak-ím'</i>	<i>tak-ójju</i>	<i>tak-ím</i>	<i>tak-ójju</i>	<i>tak-ím</i>	<i>tak-ójju</i>	<i>tak-ím</i>	<i>tak-ój</i>

2.3.2. *Expression elaboration*. This is a common type of expression change in languages with productive morphology. It is traditionally labeled analogical change. It is more interesting to note that elaboration changes may be system motivated, type motivated, or motivated by universal preferences.

2.3.2.1. *Expression elaboration*. Shorter > longer allomorphs. Russian. MidR $-i \sim -m'i \sim -am'i$ > $-am'i$ INST.PL. Shorter desinences are more often replaced by longer desinences than vice versa (Mańczak 1963); but both the narrow and the wider context here call for a different interpretation (4.3.1, 4.3.2): individual segments in this and other desinences are ascribed content in a development towards agglutination; in $-a-m'-i$, $-a-$ marks nouns, $-m'-$ marks marginal case (here, DAT or INST), terminal $-i \sim -a$ marks INST; contrast pronominal $-e-m'-i$, adjectival $-i-m'-i$, numeral “two” $-u-m'-a$ (Jakobson 1958).

Content: distributed among submorphemes. Content syntax: unchanged. Expression: allomorphy sharply reduced. Morphosyntax: unchanged.

2.3.2.2. *Expression elaboration*. Zero > overt allomorphs. Russian. Through sound change, Russian develops a $-\emptyset$ allomorph for the GEN.PL in several noun declensions. Over time the relationships among the GEN.PL allomorphs change, as the overt allomorphs are extended and become phonologically conditioned, whereas the $-\emptyset$ allomorph is curtailed. OR $-\check{i} \sim -ov\check{i} \sim -\check{i}\check{j}$ > MidR $-\emptyset \sim -ov \sim -ej$ > ModR $-\ov/-ej \sim -\emptyset$. The system context here is a preference for overt endings to zero endings for marked categories (4.3.4).

Content: unchanged. Content syntax: unchanged. Expressions: unchanged. Morphosyntax: allomorphs redistributed in favor of overt variants, especially in paradigms with $-\emptyset$ allomorph in NOM.SG.

2.3.3. *Expression substitution*. All expression changes involve the replacement of one variant by another in some or in all environments; expression reduction (2.3.1) and elaboration (2.3.2) are special cases in which a shorter, respectively, a longer allomorph is generalized. On the word level, the variants are allomorphs, as in the examples above. On the phrase level, the expression substitution may involve clitics or words. The partial displacement of auxiliary “have” by “be” in Italian dialects (Cennamo 2006) is an example; the fact that

some verbs in one sense take “have”, in another, “be” creates the impression that “have” and “be” differ in meaning; in fact they are indexes of the lexical–semantic categories across which they are distributed (2.4.3). In language-internal change, the generalized variant has arisen through neologism (ellipsis) or (analogical) extension; see the examples under 2.3.1, 2.3.2; also the redistribution of Perfect desinences in Haug (2006). But new variants may be introduced as well through dialect or language contact (3.2).

2.3.4. *Expression doubling*. Expression doubling is presumably the typical source of concord and agreement. The following examples may be idiosyncratic and may not shed much light on how regular concord and agreement arise. Both 2.3.4.1 and 2.3.4.2 may have their origins in intrusions (3.2.2). See also Harris 2006.

2.3.4.1. *Expression doubling*. Prepositional concord. Russian. Some Old and Middle Russian dialects regularly iterate prepositions in prepositional phrases with (i) conjoined head nouns, (ii) conjoined modifiers, or (iii) appositive nouns.

E.g., (i) OR *ot Jakimŭ i ot Šimŭjuna* “from J. and (from) S.”; (ii) *po starym po knjažim gramotam* “in accordance with old, (with) official documents”; (iii) *k Juriju k Ončiforovu* “to Jurij (to) Ončiforov [family name]”; *k bratu k Ygnatu* “to [my] brother (to) Ignatius”; *k sinu k svojemu k Isaku* “to his (to) son (to) Izak” (Zaliznjak 2004: 164).

Content: unchanged. Content syntax: unchanged. Expression: prepositions iterated. Morphosyntax: doubling accompanies ‘expanded’ NP heads or modifiers.

2.3.4.2. *Expression doubling*. Auxiliary. Gallipoli Serbian. In this dialect, the future auxiliary, which is a Wackernagel clitic, becomes affixed to the infinitive of its complement when the infinitive is in first position, but continues to occur sentence-second when the infinitive occurs further to the right.

E.g. (with standard forms in parenthesis), *ja=ćem ti kupi-ć-em* (st. *kupi-ti*.INF) “I will buy [it] for you”, *ti=ć-eš se zadiša-ć-eš* (st. *zadixa-ti*.INF) “you will get out of breath”, *On=će na-će* (st. *na-ći*.INF) “he will find [it]”, *mi=ć-emo ga savrva-ć-emo* (st. *savlada-ti*.INF) “we will fix it/him” (Ivić 1994: 253).

Content: unchanged. Content syntax: unchanged. Expression: auxiliary doubled. Morphosyntax: auxiliary both sentence clitic and verb affix

2.3.4.3. *Expression doubling*. Inflected complementizers. German. In North Bavarian dialects, participant markers have come to be affixed both to verb and complementizer. The development is thought to presuppose (i) morphosyntactic emancipation—a stage with participant markers movable from verb to complementizer, (ii) Wackernagel placement, and (iii) a subsequent generalization in both positions, as in Serbian (2.3.4.2).

E.g., (North Bavaria) *wen=ts ēęts baęen wā-đz* “if you (2PL) were farmers”, *wal=ts ēęts baęen ha-ts* “because you (2PL) are farmers”, (Egerland) *wen=s dā wil-st* “if you (2SG) will”, *wail=n sā wol-n* “because they want to”, *das=n sā mis-n* “that they have to” (Žirmunskij 1956: 484).

Content: unchanged. Content syntax: unchanged. Expression: doubled. Morphosyntax: participant markers both complementizer enclitic and verb affix.

2.4. Morphosyntactic change

The Gzn Schema focuses attention on changes in bonding, which can be described as changes in boundary strength or freedom of placement (2.4.1–2.4.2). But besides these, morphosyntactic changes include changes in index relations, similarly relevant to all levels of structure (2.4.3) and changes in element order, relevant to affixes within the morphological word, and to clitics and words within the phrase and the clause (2.4.4).

2.4.1. *Bond strengthening (integration)*. In the usual course of events, as far as it can be determined, affixes develop from clitics, clitics from full words, and some full words go back to phrases. These facts are traditionally (over)stated as phrase > word > clitic > affix.

Several examples have been cited in the preceding sections. But two things are worth emphasizing. First, each of the stages in the phrase > affix ‘cline’ may be quite complex. The change from clitic to affix, for instance, may require an extension of bound stems to host the new affix (2.1.3.2) and univerbation along several parameters, morphophonemic, prosodic, segmental (Andersen 1987). Secondly, a single step on the ‘cline’ may take a long time. In the Polish clitic > affix change in 2.1.3.2, for instance, bound stems became affix hosts by the 1500s, but prosodic univerbation, which began in the 1500s has only recently come to completion—five hundred years later. The development has consisted of numerous morphosyntactic innovations, spreading the affixes from environment to environment, each composed of myriad little modifications of traditional usage. The changes in 2.4.1.1, 2.4.1.2 have been no less complex.

2.4.1.1. *Integration*. Word > proclitic. Greek. MGk. *thélo (i)na [ʔáfo* “I.want that I.write” > *tha.FUT [ʔáfo* “I’ll write” (Joseph 2003, Heine 2003).

Content: volition > future. Content syntax: modal verb > tense marker. Morphosyntax: auxiliary + complementizer > proclitic. Expression: four syllables > one.

2.4.1.2. *Integration*. Word > inflected enclitic > suffix. Serbian. Old *hoć-u pisa-ti* “I.want to.write” > *pisati=hc-u.FUT.1SG* > *pisa=ć-u.FUT.1SG* (Andersen MS.c).

Content: volition > future. Content syntax: modal verb > tense marker. Morphosyntax: auxiliary > inflected Wackernagel enclitic. Expression: infinitive host truncated, auxiliary stem reduced, one syllable > single consonant.

2.4.2. *Bond weakening (emancipation)*. Although the development of greater morphosyntactic cohesion is more frequent, morphosyntactic emancipation occurs both as an internally motivated change and as a result of contact.

2.4.2.1. *Emancipation*. Affix > clitic. Russian. When in EMidR the IMPV.2SG was apocoped, the 2PL imperative desinence *-t'e* became enclitic; table 5. Its new morphosyntactic status is shown by the imperative's word-final, not word-internal sequential constraints, e.g., *postáv'=t'e* with [ft']—devoicing, but palatalization retained before consonant; word-internal constraints would call for [ft']. The enclitic is subsequently extended to the 1PL hortative to specify plural addressee or politeness, e.g., *id'-ó-m* “let's (both) go”, *id'-ó-m=t'e* “let's (all) go”; as well as to other directives, e.g. *ná.SG, ná=t'e.PL* “here you are”, *nú.SG, nú=t'e.PL* “well then”, *pólno.SG, pólno=t'e.PL* “enough!”, *bonžúr=t'e.PL* (jocular greeting, Chekhov). The larger context for this change is a typological trend towards agglutination (2.3.2.1; 4.3.2).

Content: unchanged. Content syntax: extended from imperative to other directives (mands). Expression: unchanged. Morphosyntax: verb desinence > enclitic.

Table 5. Affix > clitic in Russian

	ORuss.		MidRuss.		ModRuss.	
2-3sg	<i>dar'-í</i>	<i>stáv'-i</i>	<i>podar'-í</i>	<i>postáv'-Ø</i>	<i>podar'-í</i>	<i>postáv'-Ø</i>
2pl	<i>dar'-i-t'e</i>	<i>stáv'-i-t'e</i>	<i>podar'-i-t'e</i>	<i>postáv'-i-t'e</i>	<i>podar'-i=t'e</i>	<i>postáv'=t'e</i>
1pl	<i>dar'-i-m</i>	<i>stáv'-i-m</i>	<i>podar'-i-m</i>	<i>postáv'-i-m</i>	<i>podar'-i-m</i>	<i>postáv'-i-m</i>
1pl+2pl					<i>podar'-i-m=t'e</i>	<i>postáv'-i-m=t'e</i>

2.4.2.2. *Emancipation*. Affix > word. American English. *-ade* “product made from or with X” (e.g., *lemonade, orangeade, gingerade, limeade*) > *ade* “refreshing beverage”; cf. *Kool-Ade* (renamed *Kool-Aid* 1929), *Gatorade* (developed for the *University of Florida Gators* football team, 1965). Note that the content change in *-ade* implies a prior content reanalysis of the first constituent from “ingredient” to “characteristic”. This indicates that the content of *ade* is a result of reanalysis (3.1.4), and the noun is not a conversion, such as *ism*; conversion is a type of neologism (3.1.1).

Content: derivational > lexical; and see preceding paragraph. Content syntax: now combinable with constituents that do not denote ingredient. Expression: unchanged. Morphosyntax: suffix > head of compound > free word.

2.4.2.3. *Emancipation*. Word > compound. English. Based on contact with the orthography, old univertated compounds are resolved. E.g., *forehead* [fɔːrd] > [fɔːrhɛd], *coxswain* [kɒksɪn] > [kɒksweɪn].

Content: unchanged. Content syntax: synthetic to (semi)analytic. Expression: spelling pronunciation (contact adoption; 3.1.3). Morphosyntax: word > compound.

2.4.2.4. *Emancipation*. Compound > phrase. English. Based on contact with prescriptive grammar, the adverb suffix *-ly* is extended to what appear to be unsuffixed adverbs. E.g., *full-fledged, full-grown* > *fully-fledged, fully-grown*.

Content: unchanged. Content syntax: unchanged. Expression: overt adverb marking (intrusion due to transference; 3.1.2). Morphosyntax: compound > phrase.

2.4.2.5. *Emancipation*. In addition to the examples in this section, consider the emancipation (desinence > phrase enclitic) entailed by the regrammation of the English *-s* (2.1.2.2, 2.4.3; Askedal 2006), and several other examples discussed by Askedal (2006): Japanese *-ga* (enclitic > word) entailed by its regrammation and upgrading to complementizer; Northern Irish *-mid* > *muid* “we” (desinence > word), subsequently the basis for a new emphatic *muid-í*; Northern Saame *taga* “without” (enclitic > word); Gk. *ksana* “again” (prefix > adverb).

2.4.3. *Grammatical indexes*. An important kind of grammatical change is the development of grammatical indexes. At the word level, these are cases of allomorphy, viz. grammatically conditioned alternants. At other structural levels, there is phrase-internal (concord), clause-internal (agreement), cross-clausal (switch reference), text-grammatical (endophoric elements), and speech-act indexing (exophoric elements, discourse markers). On each level, indexing serves to signal coherence and create textuality.

Languages with rich morphology can have correspondingly rich systems of grammatical indexing. Several examples are analysed and further categorized in Andersen (1980). Gaeta (2004) analyses a number of examples of grammatical indexing, on the word, the phrase, and the clause level; they are all results of phonological change that have become grammatically conditioned, that is, grammatical indexes developed from below.

In this volume, Ottósson (2006) refers to word-level indexes as “quasi-exponence” and “morphosyntactic sensitivity”. Maiden (2006) discusses systems of word-level indexes, but does not recognize them as having any sign value. Schøsler (2006) describes the major typological shift in case-role indexing from the Latin system of morphological cases to the “variety of organisational devices ... lexical, morphological, analytical, and topological” in modern French.

Since grammatical indexes have content, it seems reasonable to speak of their development as *index grammar*. The examples that follow are intended to be suggestive and therefore limited to simple word-internal indexes. To conserve space, I omit examples of index regrammation and degrammation.

2.4.3.1. *Index grammar from above*. In Old Ukrainian, nouns, adjectives, and pronouns had a DAT-INST.DU ending *-V-m-a* (where *-V-* stands for a variety of desinence-initial vowels) and INST.PL desinences of the general form *-(V)-(m)-y* (e.g., *-amy*, *-my*, *-y*). When the dual number was lost, the DAT-INST.DU, now a DAT-INST-PL allomorph, survived as *-oma* in the numeral “two”. Subsequently, *-(o)ma* replaced the inherited INST.PL endings of the numerals “three” (*tr-ymy* > *tr'-omá*) and “four” (*čotyr-my* > *čotyr-má*), and later it was extended to all cardinal numerals (*pjat'-omá* “five”, *st-omá* “hundred”) and quantifiers (*bahat'-má* “many”, *ki'k-omá* “some”); Andersen (MS.e).

Content: (DAT-INST.DU allomorph > INST.PL allomorph.) Content syntax: unchanged, concord marker. Expression: no change. Morphosyntax: idiosyncratic lexical > lexical-grammatical category index.

2.4.3.2. *Index grammar from below*. German. In early Germanic, front vowels develop from fronted allophones of back vowels in syllables at one time preceding /i/ or /j/ (Janda 2003; Gaeta 2004). The resulting alternations of certain back vowels with corresponding front vowels, called Umlaut, are subsequently grammatized with a diversity of indexical functions. Umlaut indexes derived causatives, certain person forms in conjugation, comparatives, diminutives, and, most famously, it is an index of “plural” number of nouns. Thus in Modern German, in some nouns the replacement of a back vowel by a corresponding front vowel

accompanies an overt plural ending: *Mann–Männ-er* (M) “man”, *Gast–Gäst-e* (M) “guest”, *Land–Länd-er* (NT) “country”, *Frucht–Frücht-e* (F) “fruit”; in others, a zero ending: *Vogel–Vögel* (M) “bird”, *Vater–Väter* (M) “father”, *Mutter–Mütter* (F) “mother”. Umlauting has historically been extended to many nouns, but continues to be lexically limited. Nouns with stem vowels that are not umlautable lack this grammatical index.

Morphosyntax: Phonological conditioning > grammatical indexes.

In English, umlaut indexing of “plural”, once comparable to that described for German, has been curtailed, and it no longer exists as a regular segmental alternation. In the handful of Modern English nouns that show reflexes of umlaut—*mouse–mice*, *goose–geese*, etc.—the best analysis, it seems, is to recognize that the plural allomorph as a whole (e.g., *geese*) symbolically represents the same lexical content as the ‘singular allomorph’ (*goose*) and additionally serves as an index of the zero allomorph of the plural. The ‘singular’ allomorph is an unmarked, general-purpose form, now being extended to the plural of neologisms (*mouses*, *gooses*). Nouns such as *sheep*, *craft* differ from these by having no singular–plural allomorphy and thus no grammatical index for the zero plural allomorph.

2.4.4. *Changes in element order*. Element order is a simple means of indexing meaningful elements. It is grammaticized on the clause level in topological schemas, e.g., S–IO–DO in Gm. *Hat Braun Roth Weiss empfohlen?* “Has Braun recommended Weiss to Roth?”; on the grammation of SVO in French, see Schøsler (2006). Clitics commonly index the margins of phrases or clauses. Thus the regrammation of the English ‘genitive’ *-s* (2.1.2.2) was followed by two morphosyntactic changes: it was emancipated to clitic status (2.4.2.5) and came to be enclitic to the determiner phrase it marks.

2.4.4.1. *Split infinitives*. English. Attested since the 1300s, they have been problematicized from a number of points of view (van Gelderen 2004, Askedal 2006). It appears that from being proclitic to the infinitive form, the infinitive marker *to* has become proclitic to the infinitive clause, including any pre-verbal adverb(ial), in most environments; the change may still be in progress. There is no apparent change in its proclitic status (pace Askedal 2006), and no apparent change in its scope (i.e. content syntax; Fischer 1999). Since the change from preposition-like to complementizer-like placement does not reflect any upgrading, it calls for a different explanation. The similarity between the movement of *to=* to the left margin of the infinitive clause and that of the DP marker *=s* to the right margin of the determiner phrase leaps to the eye. Both changes seem motivated by the universal preference for placing clitics at the margins of the relevant constituents.

2.4.4.2. *Proclitic ‘climbing’*. French. In some periods and varieties of French, the (prefix >) proclitic *re=* “again” is preposed larger constituents than the verb whose signified activity is repeated: Older Fr. *Tu=me=devrois re=dire* > *Tu=me=re=devrois dire* “You should tell me again”; ModFr. *Revoilà le chien qui hurle* “There is that dog howling again” (Maupassant). Similar changes in other Romance languages, e.g. It. dial. *Chi ar c’era?* “Who was that again?” (i.e. *chi ar=c’era*; st. *Chi c’era di nuovo?*); Askedal (2006).

No change in Content, Expression, or Content syntax (specifically, no scope increase). Morphosyntax: emancipation (prefix > proclitic); preverbal > prephrasal placement.

2.4.4.3. *Enclitic > proclitic*. Estonian. According to one interpretation, after its regrammation as interrogative enclitic, OEst. *es.Q* was (i) decliticized and placed in Wackernagel’s position,

and subsequently (ii) became a “free particle” in “first position” (Askedal 2006). It appears, however, there was no “decliticization”. First, =*es* was a sentence-initial Wackernagel enclitic, i.e. it was placed in relation to the first orthotonic word (or phrase) of the sentence and, since it was an enclitic, followed it. Later, it became a sentence-initial proclitic. To understand the change enclitic > proclitic, it may be useful to view *es* in its areal context. Formerly, =*es*.Q conformed to the northern area in Eastern Europe with Wackernagel enclisis (Fi. =*ko*.Q, Russ. =*li*.Q). Now, both Est. dial. *es*= and st. *kas*= conform to the southern area with sentence-initial proclisis (La. *vai*=.Q, Li. *ar*=.Q, Pol. *czy*=.Q, Br. *c’i*=.Q, Ukr. *čy*=.Q).

No change in Content, Content syntax, or Expression. Morphosyntax: sentence enclitic > sentence proclitic. Possibly a contact change (3.3–3.4).

2.5. Conclusion

The preceding survey has exemplified categories of change defined by the constituent elements of linguistic signs, which are subject to change—content, expression, content syntax, and morphosyntax—and some subcategories that are logically possible or happen to be well attested. If we look carefully at actual chronological developments in language histories we are likely to discover other types of change which—like those reviewed above—will raise questions about their origins; see, for instance, Andersen (MS.bc).

3. Types of innovation

Each of the types of change that have been exemplified in the preceding pages is, technically speaking, a diachronic correspondence, that is, a correspondence between two states separated in time, a ‘before’ and an ‘after’ which have been interpreted as a change event. This interpretation may be an oversimplification in the sense that a correspondence may be the result of a sequence of changes. But even when it is reasonable to consider a correspondence the reflection of a single, temporally well-defined event in a tradition of speaking, we have to recognize that a change is a complex event that progresses over time only thanks to myriad minor adjustments in the competences and usage of the speakers. These innovations, as the adjustments are traditionally called, are the ‘atoms’ of which a change is composed.

The historical linguist’s task, accordingly, is to resolve every observed diachronic correspondence into the change or changes that brought it about, and each change, again, into the innovations from which it resulted.

The normal course of events through which a new expression originates, gains currency, and becomes established as part of a tradition of speaking is the following: one or more speakers (i) make a (primary) innovation and (ii)

actualize it in usage; other speakers (iii) adopt the new expression and (iv) actualize it in their usage; if the new expression is used widely and long enough, new cohorts of speakers (v) will acquire it as an integral part of their competence, and (vi) actualize it in their usage; the new expression becomes generalized in the community through repeated cycles of (iii)–(vi). One can speak of such a series of overlapping kinds of innovation as a ‘change scenario’ and of the constituent innovation types (i)–(vi) as ‘subchanges’.

3.1. The four basic types

There are four basic types of innovation. Each of these should be understood as a (covert) modification of competence that is (overtly) actualized in speech (Andersen 2001c; MS.bc).

3.1.1. *Neologism*. The creation of new signs. (i) New combinations of content and expression (coinage); e.g., *grammation*. (ii) New expressions for extant content (remedial innovations, jocular deformations, new collocations, etc.); e.g., *nuked* “nude, naked”; “have” + NP_{DO} + past participle, eventual source of retrospective tenses (Detges 2000). (iii) New, derived realizations of extant expressions (clippings; ellipsis); e.g., the omission of a lexical verb in 2.1.3.1 or phonological material in 2.3.1.

A neologism is always an alternative to an existing expression (or to an *ad hoc* circumlocution) and hence implies the creation of a set (paradigm) of alternative expressions, variants. The paradigm may subsequently be reduced as the variants are ascribed markedness values in reanalysis and the unmarked variant is extended, the marked variant curtailed (expression reduction, 2.3.1).

3.1.2. *Extension*. The application of extant linguistic means in new usage. E.g., the use of a retrospective present (present perfect) to refer to past situations without necessary relevance to the present, eventually replacing simple past tense(s) in 2.1.2.3; the use of the subjectless passive for definite objects (Eythórsson 2006). The extended expression typically has less semantic depth or intension than its source (bleaching); the two contents, full and bleached, coexist in the case of layering; but there is no diminution of intension when an unmarked term is extended. Extension may imply a curtailment in the use of other linguistic means; e.g., the extension of the present perfect entailed a curtailment in the use of the Old Russian imperfect and aorist (2.1.2.3); French documents a similar relation between its *passé composé* and *passé simple*.

3.1.3. *Adoption*. The acceptance of newly encountered usage for passive or active use.

3.1.4. *Reanalysis*. The fresh analysis of received usage in the course of new speakers' grammar formation. From the linguist's point of view, there are several kinds of reanalysis, which reflect that the learner decides issues of analysis of two kinds, substantive (which concern the content, expression, content syntax, or morphosyntax of elements of grammar), and valuative (the ascription of markedness values). More in 4.2.

3.2 The four contact types

In contact situations, including bilingual and multilingual situations, the four types of innovation may give rise to specific kinds of change. The results of these are traditionally called 'borrowings', with little or no attention paid to their different origins in innovations.

3.2.1. *Borrowings*. Borrowings (in the narrow sense) are kinds of bilingual neologism, in which the content and expression of a sign, or just an expression is introduced from another language. The term *calque* refers to borrowed composite signs whose expressions have been translated element by element.

3.2.2. *Intrusions*. Intrusions are new elements that enter the usage of a matrix language L_1 due to transference from a contact language L_2 , or which enter the usage of a matrix language L_2 due to interference from a contact language L_1 ; transference and interference are kinds of extension from one language to another.

What is recognized as substratum influence is the result of interference intrusions; superstratum influence results from transference intrusions. Both kinds of intrusion play an essential role in the formation of *koinés* and language alliances. Where a written standard serves as a superstratum, typical transference intrusions are spelling pronunciations and the elements of prescriptive grammar (2.4.2.3, 2.4.2.4).

3.2.3. *Contact adoption*. This is the acceptance of alternative linguistic means from a contact dialect; it plays an essential role in areal diffusion and in the development of transitional dialects.

3.2.4. *Bilingual reanalysis*. This is the acquisition of composite competence in conditions of social bilingualism or bidialectism; an important aspect of this is

the ascription of values to functionally equivalent L_1 and L_2 elements, which then determine speakers' preference in usage. It plays an essential role both in areal diffusion and in the formation of language alliances.

3.3 Conclusion

Among the different kinds of innovation, it is surely most difficult to understand what enables speakers to make up new expressions (neologism) or apply their repertory of expressions in new ways (extension), or what decides them to adopt and use expressions they encounter in the usage of others. These kinds of innovation presuppose a creativity that has a grammatical competence as its point of departure. Innovations that arise in reanalysis are simpler and therefore more likely to afford us direct insight into the premisses that guide learners in the formation of a grammar. They will be the topic of the next chapter.

4. Reanalysis as a source of innovation

One of the most remarkable facts about linguistic change is its determinate direction. Changes that we can observe in real time—for instance, as they are attested in the textual record—typically progress consistently in a single direction, sometimes over long periods of time. Recently some newcomers to the field have speculated that the progress of linguistic change is analogous to the statistical drift in evolution (Croft 2002, Roberts and Roussou 2003). But linguistic drift is different (Andersen MS.a). Not only does linguistic drift have direction, as Sapir observed (1921: 155). It has structure: innovations are typically actualized across the social, stylistic, and grammatical categories of a language in an orderly fashion that reflects the hierarchical organization of those categories (see below).

Sapir realized that the sustained, determinate direction of drift can be understood only by supposing that “the drift of a language is constituted by the unconscious selection on the part of the speakers of those individual variations that are cumulative in a special direction”. To account for the directedness of drift, he called for an investigation of “the intuitive bases of speech” and, specifically, a study of linguistic patterning “and the ‘weights’ and psychic [i.e. cognitive] relations of the single elements in linguistic patterns” (p. 183)—a project entirely germane to the topic of this volume.

Sapir's great insight was, first, that it is the ‘weights’ of synchronic variants that determine the direction of change, and secondly, that these

weights must be intuitively assigned by speakers and, in turn, consistently guide the speakers' selections of certain variants over others in speech.

Most modern readers have no difficulty agreeing with Sapir that linguistic elements are weighted or, as we say now, participate in markedness (m-ness) relations. But to bring his insight to fruition we need to answer two questions: How do those m-ness values translate into change? And, What are the criteria by which m-ness values are assigned.

4.1 M-ness in actualization

In earlier work I have put forward a theory of m-ness that makes it possible to clarify how synchronic m-ness values translate into change. Here I will mention very briefly just two elements of this theory, the nature of m-ness and m-ness agreement.

4.1.1. *The nature of m-ness.* M-ness is in essence an inclusive relation. Inclusion is the most primitive cognitive relation (e.g., y^M is a kind of x^U), and it seems that this relation is assigned to related elements prior to any analysis of their logical relation (inclusive or exclusive, contrary or contradictory, etc.). Hence even relations that are logically exclusive have m-ness values; e.g., $long^U : short^M$ (contraries), $man^U : woman^M$ (contradictories), $east^U : west^M$ (converses) (Lyons 1977; Andersen 2001b).

All the syntactic and paradigmatic (taxonomic) relations in language are inclusive. Hence they are ideally represented by tree diagrams, in which any lower node is included under a higher node, and in any pair of branches, one will be marked, the other unmarked, reflecting their different usage potential (see below).

To understand the process of change, it is particularly important to appreciate the dynamic character m-ness imparts to variation. In a paradigm of covariants, one will be unmarked, the other(s), marked; if there are several marked covariants, they may form a hierarchy if the variation is subject to hierarchical constraints (Wolfram and Fasold 1974). Since the m-ness relation is inclusive, the usage potential of any marked term is included in that of the unmarked counterpart even though in normal usage they may be in complementary distribution. This difference in potential can be exploited in usage: the unmarked term may be used as proxy for a marked term, just as a hypernym can stand in for its hyponym(s). If such a 'skewed' usage arises and is acquired by a new cohort of learners, it may become the baseline for additional skewing in those speakers' usage, and so on through successive cohorts of speakers: as the variation is acquired by successive cohorts, an age

gradient develops in which the unmarked term becomes extended more and more and the marked term(s) become more and more curtailed (3.1.2) until the unmarked term is eventually generalized. This is how change results from “the unconscious selection on the part of the speakers of those individual variations that are cumulative in a special direction” (Sapir 1921: 155).

4.1.2. *M-ness agreement*. The principle of m-ness agreement is manifested in change in a number of ways. Typically in internally motivated change, an innovation is extended to unmarked contexts earlier than to corresponding marked contexts, and to some categories earlier than others in conformity with a hierarchy of grammatical categories (Timberlake 1977; Andersen 1990; Andersen 2001abc and the other contributions in Andersen 2001d; Cennamo 2006; Ottósson 2006). M-ness agreement will not be discussed in the following exposition, but it is worth mentioning because it is strong evidence that m-ness is intrinsic to all (underlying) linguistic relations and is not to be identified with such superficial observables as relative frequency or complexity.

4.2 M-ness in reanalysis

We assume that in the process of grammar formation, learners decide analytic issues of two kinds, substantive (which concern the content, expression, content syntax, and morphosyntax of elements of grammar), and valuative (the ascription of m-ness value to every individual element of grammar in relation to other elements).

Innovations in reanalysis show that the first kinds of issues are decided on the basis of such elementary procedures as segmentation of expressions, analysis of contents, and assignment of content elements to elements of expression; the formation of sentence, phrase, and morphosyntactic patterns for the expression of content-syntactic relations; the organization of expressions into hierarchical paradigms on the basis of their content; et cetera. It is possible to formulate more or less extensive sets of learners’ ‘operating procedures’ (Slobin 1985, Faarlund 2006).

The learner’s analysis moves forward on two levels simultaneously resulting in the formation of a base grammar (a system of productive grammatical rules) and a system of usage rules that capture current appropriateness norms (Andersen 2001c). The usage rules are based on (couched in terms of) the base grammar and enable the learner to produce, first of all, established usage where this cannot be derived by the productive rules (e.g., irregularities, idioms, archaisms) and, secondly, variable usage

appropriate to the learner's changing age-group, gender, social-class membership, and ethnicity as well as relevant register and stylistic demands.

Since m-ness values are implicit in all syntactic and paradigmatic relations, they are found throughout the base grammar and usage rules, including the complex of variable rules.

4.3 Iconicity in reanalysis

Here I will return to a few examples presented in sec. 2 that illustrate some of the possible grounds for m-ness ascription. Specifically, the examples show the role of iconicity in the creation and selection of variants.

4.3.1. *System-specific iconicity: segment count as a diagram of grammatical number.* In Old Russian, several Common Slavic declensions merge in one, the masculine First declension (e.g., ModR *stól* "table", *m'ód* "honey", *góst* "guest"). In Middle Russian, several case endings in this declension have two or three allomorphs differing in number of segments; see table 6. In modern Russian, this allomorphy has been sharply reduced. The outcome is an orderly distribution of desinences and allomorphs such that for each pair of singular and plural case endings, the opposition singular^U vs. plural^M ("unspecified number" vs. "more than one") is diagrammed by a difference between n segments^U and $n + 1$ segments^M. (A few details that are not shown in table 6 will be taken up in 4.3.4.)

Table 6. Segment count as a diagram of grammatical number. First-declension M nouns

	Middle Russian		Modern Russian	
	SG	PL	SG	PL
NOM	∅	1 ~ 2 ~ 3	∅	1
ACC	∅	1 ~ 2 ~	∅	1
GEN	1	∅ ~ 2 ~ 2	1	2
LOC	1	2	1	2
DAT	1 ~ 3	2	1	2
INST	2	1 ~ 2 ~ 3	2	3

In interpreting this development it is not enough to know that there is a tendency in languages for a diagrammatic relation between grammatical number and desinence length (Jakobson 1965). The actual outcome here is a much tighter relationship between content and expression and calls for a more specific explanation. It is not enough either to know that across language histories, shorter allomorphs are more often replaced by longer allomorphs than vice versa (Mańczak 1963). This generalization fits the facts: here shorter allomorphs were generalized in three cases (DAT.SG, NOM-ACC.PL), but longer, only in two (GEN.PL, INST.PL); but Mańczak's "more often" is no better than Jakobson's tendency.

On the assumption that in each instance, the allomorph that is generalized is the one that is evaluated as unmarked, one might simply suppose that shorter allomorphs were unmarked in the singular, but in the plural, allomorphs were evaluated as unmarked not according to their own length, but in relation to singular desinences: those plural allomorphs were unmarked that were just one segment longer than the corresponding (unmarked) singular allomorph. This is what is suggested in the first paragraph of this section.

One cannot exclude this possibility a priori. But it is just possible that the neat correspondences between grammatical number and segment count seen in table 6 are the result of several, more general iconicity principles acting, or applied, independently.

One of these could be a universal principle reflected in Mańczak's statistical generalization: other things equal, of several allomorphs the one with the fewest segments is unmarked. This is in fact precisely the principle that is behind Zipf's Law. This famous 'law', which links frequency with word shortening, conflates three innovations: (i) the innovation of phonetically shortened variants (ellipsis; 3.1.1), (ii) the ascription of m-ness values to longer^M and shorter^U variants (valuative reanalysis) and (iii) the subsequent curtailment of the longer^M and extension of the shorter^U variants (actualization) (Andersen 2001c).

Let us provisionally adopt the m-ness principle in (ii) (the 'Zipf principle'), which will account for three of the simplifications in table 6, and turn to the cases in which longer allomorphs were generalized.

4.3.2. *Type-motivated iconicity: the preference for agglutination*, 1. Ex. 2.3.2.1 exemplified expression elaboration (from shorter to longer allomorphs) with the Russian INST.PL of nouns, in which the lexically conditioned allomorphy MidR $-i \sim -m'i \sim -am'i$ yielded to $-am'i$.

The background for this change was an overabundant allomorphy in Middle Russian plural paradigms, especially in the LOC-DAT-INST.PL. The LOC.PL and DAT.PL endings ($-V-x$ and $-V-m$) had a variety of initial vowels— i , \hat{e} , o , \hat{o} , a , with complex, cross-cutting lexical and phonological conditioning—followed by invariant $-x$, respectively $-m$. In the INST.PL the regular endings were First decl. $-i$ (written $-y$ and $-i$) (lexical exceptions with $-m'i$), Second decl. $-am'i$, Third decl. $-m'i$, that is, $-(a)-(m')-i$; but adjectives, pronouns, and some numerals had other desinence-initial vowels, so the general shape was $-(V)-(m')-i$.

The complex allomorphy in these endings formed the basis for reanalyses. First of all, $-x$ was posited as LOC.PL marker, $-m-$ as peripheral-case (DAT-INST) marker, and $-i$ as INST.PL marker. It should be mentioned that the consonant $-m-$ ($\sim -m'$) also occurred, and occurs, in LOC-DAT-INST.SG and DAT-INST.PL endings of pronouns and adjectives. Secondly, desinence initial vowels were generalized according to nominal part of speech: $-a-$ for nouns, $-\hat{e}-$ for pronominal adjectives, $-i-$ (written $-y$ and $-i$) for adjectives, $-u-$ for "two", $-o-$ for "three, four").

We can now see why the longest of the three INST.PL allomorphs would be unmarked and hence came to be generalized. Recall from above that the shortest, and highly frequent, INST.PL allomorph $-i$ was actually included in the longer allomorphs $-m'i$ and $-am'i$; but it alone among all the plural peripheral-case allomorphs lacked a case-specific consonant; by comparison with the others, it was opaque. The allomorph $-m'i$ began with a consonant, or, by comparison with the others, it lacked a desinence-initial vowel. The $-am'i$ ending included $-i$ as its final segment, had the case-specific consonant $-m'$, and had an initial vowel. It was the only one of the INST.PL allomorphs with a shape ($-V-m'-i$) that paralleled LOC.PL ($-V-x$) and DAT.PL ($-V-m$).

The initial vowel of $-a-m'-i$ occurred in the Second declension LOC.PL and DAT.PL endings ($-a-x$, $-a-m$); of all the relevant desinence-initial vowels, $-a-$ alone did not alternate with another vowel. Furthermore, $-a-$ was the inherited desinence-initial vowel in the personal pronouns (viz. $n-a-s$, $n-a-m$, $n-a-m'-i$ "us", $v-a-s$, $v-a-m$, $v-a-m'-i$ "you", LOC-DAT-INST.PL). Although formerly meaningless, the $-a-$ became meaningful by the contrast of these substantial case desinences with those of other nominal parts of speech.

In terms of its internal structure, its paradigm-internal relations, and its relation to other paradigms, then, $-a-m'i$ was unmarked.

The actualization of these m-ness valuations, in which marked allomorphs and submorphemic elements were gradually curtailed and eliminated, stretched over more than

four hundred years. Through innumerable innovations in usage rules, usage gradually came to generalize a base system of unmarked allomorphs with (i) desinence consonants and non-initial vowels representing case features or cases, (ii) desinence-initial vowels indexing (2.4.3) part of speech, and, as seen in table 6, (iii) segment count as a (diagrammatic) representation of grammatical number in the INST.PL as in the LOC-DAT.PL.

In the process, an inflective or fusional system of nominal morphology was transformed into an (incipiently) agglutinative one (Jakobson 1958). There is an old belief that agglutinative systems develop into fusional ones, and not vice versa; but see Igartua (MS). In the history of Russian (and the other modern Slavic languages), however, not only is the opposite development widely attested, but it originated through a well-motivated resolution of the rather complex allomorphy that had accumulated in the medieval systems. Of the two types of morphosyntactic mapping, agglutination is evidently unmarked in relation to fusion. This would make every transparent concatenation of submorphemic elements in the reanalysed base grammar unmarked in relation to the corresponding received, fused allomorph captured in the usage rules, $-a-x^U : -\acute{e}x^M, -ix^M, -ox^M$ LOC.PL; $-a-m^U : -\acute{o}m^M, -om^M$ DAT.PL; $-a-m'-i^U : -m'i^M, -i^M$ INST.PL. This transparency must have played a role in the adoption, transmission, and gradual diffusion of the innovated desinences.

4.3.3. *Type-motivated iconicity: the preference for agglutination*, 2. Ex. 2.4.2.1 described the morphosyntactic emancipation of the Middle Russian imperative desinence $=t'e.2PL$, which becomes an enclitic some time before the 1700s. Its extension to the 1PL hortative and some other directive forms leaves no doubt that it has become emancipated from the imperative. The phonological constraints that apply to the stem-stressed IMPV.2PL (2.4.2.1) show that $=t'e.2PL$ follows a word boundary: it is attached to IMPV.2SG as an enclitic. It is interesting that the same phonological constraints apply before the reflexive morpheme. This was earlier a second-position enclitic, but has become fixed to verbal forms, phonologically not as a suffix, but an enclitic.

Since clitics develop into affixes more often than vice versa, the emancipation of the Russian IMPV.2PL morpheme is interesting. It is certainly relevant to its interpretation that the inflectional systems of both verbs and nouns in this language have become agglutinative in character (Jakobson 1957). The transformation of nominal and verbal inflection occurred from the late 1200s on. Against this background, the emancipation of the IMPV.2PL morpheme can be seen as part of the emergence of this new, clearly concatenative morphosyntactic type.

4.3.4. *Universals of iconicity*. Since the 1100s–1200s, Russian has gradually replaced its $-\emptyset$ GEN.PL allomorph with overt allomorphs, $-ov/-ej$, the latter having become phonologically conditioned allomorphs (2.3.2.2).

It is a fact that across language histories, zero endings are more often replaced by overt endings than vice versa (Mańczak 1963). From this one can infer that language learners consider zero allomorphs marked more often than not. Now, Russian (and other Slavic languages) have zero allomorphs also in the NOM.SG (First-declension masculines, Third declension), yet here there has been no tendency to replace zero with overt allomorphs. This suggests a possible correlation between unmarked and marked categories on one hand and zero and overt marking on the other, which is well attested across languages and may represent a universal (Greenberg 1969). In the NOM.SG forms then, the zero allomorph would be unmarked, part of the base system, and stable, whereas in the GEN.PL the zero allomorph would be marked, assigned by usage rules, and liable to curtailment and eventual replacement by unmarked overt allomorphs.

There is an additional detail regarding this replacement by overt GEN.PL allomorphs: it has progressed further in masculine First declension nouns than in neuters and Second-declension nouns. Since masculine First-declension nouns have a zero ending in NOM.SG one

might wonder whether the ‘Greenberg correlation’ is sensitive to the presence of other zero endings in the paradigm. This may be so. But in Russian, the fact that the introduction of overt allomorphs has progressed further in the First declension (non-feminine nouns) than in the Second declension (mostly feminine nouns), and within the First declension, further in masculines than in neuters may simply manifest the principle of m-ness agreement (4.1.2).

4.3.5. Summary. The last few examples have illustrated how the language-specific diagram of grammatical number by desinences of contrasting length in table 6 may have arisen thanks to several markedness-principles. The ‘Zipf principle’ accounts for three of the cases in which the shortest allomorph was generalized (DAT.SG, NOM-ACC.PL; 4.3.2). The ‘Greenberg correlation’ of m-ness in content categories with zero vs. overt expressions accounts for one (GEN.PL; 4.3.4). The interpretation of the remaining, fifth allomorph generalization (INST.PL, 4.3.3) brings several principles into play in the resolution of the relevant allomorphs into submorphemic units (*-m’-*, *-i*) and the grammation of part-of-speech indexes (*-a-*): transparency vs. opacity, unity vs. diversity of expression.

In all instances it appears coincidental that the application of the m-ness principles gave rise to a diagram in which n vs. $n + 1$ segments represented the relation “unspecified number” vs. “more than one”.

5. Conclusion

In the last few pages I have tried to illustrate in some detail how the values that are ascribed in the process of grammar formation may be reflected in innovations that in the long run lead to change.

Traditionally many linguists have been satisfied with easier explanations. Some discover analogical relations in grammatical changes and come to believe that all change is analogical. Others observe how new forms may be more economical than the forms they replace and become convinced that all change leads to simplification. Recent investigations of Gzn have not changed these traditional attitudes much.

The larger aim of this paper has been to advocate that we change our focus from Gzn to grammatical change in general and shift our attention from diachronic correspondences to the types of innovation in which changes originate, especially to innovative reanalysis. It is at this level of detail, where new variants are created, that the most productive questions can be asked about any given innovation: whether it conformed to existing rules of the language, whether it was motivated by surface ambiguities in received usage, whether it conformed to typological properties of the language, or whether it embodied some general principle of language manifested as a learner's strategy.

One strong argument in favor of asking such questions is that each of these questions and others like them opens a door to further investigation.

REFERENCES

- Andersen, Henning. 1969. "The peripheral plural desinences in East Slavic". *International Journal of Slavic Linguistics and Poetics* 12: 19–32.
- Andersen, Henning. 1980. "Morphological change: towards a typology", *Historical Morphology*, ed. Jacek Fisiak, 1–50 [Trends in Linguistics. Studies and Monographs, no. 17]. The Hague: Mouton. Reprinted in *Morphology. Critical Concepts in Linguistics*, vol. 6. *Morphology: Its place in the wider context* ed. by Francis Katamba. 397–443. London–New York: Routledge, 2004.
- Andersen, Henning. 1987. "From auxiliary to desinence". *Historical Development of Auxiliaries* ed. by Martin Harris and Paolo Ramat, 21–52. Berlin: Mouton de Gruyter.
- Andersen, Henning. 1990. "The structure of drift". In *Historical Linguistics 1987. Papers from the 8th International Conference on Historical Linguistics* ed. by Henning Andersen and Konrad Koerner, 1–20. Amsterdam: John Benjamins.
- Andersen, Henning. 2001a. "Introduction". In Andersen 2001d: 1–20.
- Andersen, Henning. 2001b. "Markedness and the theory of linguistic change". In Andersen 2001d: 19–57.
- Andersen, Henning. 2001c. "Actualization and the (uni)directionality of change". In Andersen 2001d: 225–248.
- Andersen, Henning. 2001d. *Actualization* ed. by Henning Andersen [Current Issues in Linguistic Theory, 219]. Amsterdam–Philadelphia: John Benjamins.
- Andersen, Henning. MS.a. "Synchrony, diachrony, and evolution". In *Competing Models of Linguistic Change* ed. by Ole Nedergaard Thomsen. Amsterdam–Philadelphia: John Benjamins. To appear.
- Andersen, Henning. MS.b. "Grammation, regrammation, and degrammation: tense loss in Russian". To appear in *Diachronica*.
- Andersen, Henning. MS.c. "Slavic periphrastic futures. Divergence and convergence". In *Change in verbal systems. Issues in explanation* ed. by Kerstin Eksell and Thora Vinther. Bern: Peter Lang. To appear.
- Andersen, Henning. MS.d. "The last days of the Russian future perfect". In *Festschrift for NN*. To appear.
- Andersen, Henning. MS.e. "Possibility, impossibility and contingency". In *Festschrift for NN*. To appear.
- Askedal, John Ole. 2006. "'Degrammaticalization' versus typology. Reflections on a strained relationship". This volume.
- Avanesov, Ruben I. and Varvara G. Orlova. 1965. *Russkaja dialektologija*. (Second edition.) Moskow: Nauka.
- Beths, Frank. 1999. "The history of *dare* and the status of unidirectionality". *Linguistics* 37: 1069–1110.
- Browne, Wayles. 1993. "Serbo-Croat". In Comrie and Corbett 1993: 306–387.
- Burridge, Kate. 1995. "From modal auxiliary to lexical verb: the curious case of Pennsylvania Dutch *wotte*". In *Historical Linguistics 1995, 2. Germanic Linguistics* ed. by Richard M. Hogg and Linda van Bergen,

- 19–33 [Current Issues in Linguistic Theory, 162]. Amsterdam–Philadelphia: John Benjamins.
- Cennamo, Michela. 2006. “The rise and development of analytic perfects in Italo-Romance”. In this volume.
- Comrie, Bernard and Greville G. Corbett. 1993. *The Slavonic Languages*. London–New York: Routledge.
- Croft, William. 2000. *Explaining Language Change. An Evolutionary Approach* [Longman Linguistics Library]. Harlow, Essex: Longman.
- Eythórrsson, Thórhallur. 2006. “The new passive in Icelandic: merge over move”. In this volume.
- Faarlund, Jan Terje. 2006. “A mentalist interpretation of grammaticalization theory”. In this volume.
- Fischer, Olga S. 1999. “On the role played by iconicity in grammaticalization processes”. In *Form Miming Meaning: Iconicity in language and literature* ed. by Max Nänny and Olga S. Fischer, 345–374 [Studies in Language Companion Series, 53]. Amsterdam–Philadelphia: John Benjamins.
- Gaeta, Livio. 2004. “Exploring grammaticalization from below”. In *What makes Grammaticalization? A look from its Fringes and its Components* ed. by Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer, 45–76. Berlin–New York: Mouton de Gruyter.
- Gelderen, Elly van. 2004. *Grammaticalization as Economy* [Linguistik Aktuell/Linguistics Today, 71]. Amsterdam–Philadelphia: John Benjamins.
- Gelderen, Elly van. 2006. “Linguistic cycles and economy: some challenges”. In this volume.
- Greenberg, Joseph H. 1969. “Some methods of dynamic comparison in linguistics”. In *Substance and Structure of Language* ed. by Jaan Puhvel, 137–203. Berkeley–Los Angeles: University of California Press.
- Greenberg, Joseph H. 1991. “The last stages of grammatical elements: contractive and expansive desemanticization”. In *Approaches to Grammaticalization, vol. 2. Focus on Theoretical and Methodological Issues* ed. by Elizabeth C. Traugott and Bernd Heine, 301–314 [Typological Studies in Language, 19.1]. Amsterdam: John Benjamins.
- Harris, Alice. 2006. “Explaining exuberant agreement”. In this volume.
- Heine, Bernd. 2003. “Grammaticalization”. In Joseph and Janda 2003: 575–601.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: University Press.
- Heltoft, Lars, Jens Nørgård-Sørensen and Lene Schøsler. 2005. “Grammatikalisering som strukturforandring”. In *Grammatikalisering og struktur* ed. by Lars Heltoft, Jens Nørgård-Sørensen and Lene Schøsler, 9–30. Copenhagen: Museum Tusulanum, Københavns Universitet.
- Heltoft, Lars. 2001. “Reanalysing structure—the Danish definite article, its predecessors and development”. *Acta Linguistica Hafniensia* 33.67–90.
- Hodge, Carleton. 1970. “The linguistic cycle”. *Language Sciences* 13: 1–7.

- Hopper, P. J. 1991. "On some principles of grammaticalization". In *Approaches to Grammaticalization. Volume 1* ed. by Elizabeth C. Traugott and Bernd Heine, 17–35. Amsterdam–Philadelphia: John Benjamins.
- Humboldt, Wilhelm von. 1822 [1884]. "Ueber das Entstehen der grammatischen Formen und ihren Einfluss auf die Ideenentwicklung". In *Die sprachphilosophischen Werke Wilhelm's von Humboldt* ed. by H. Steinthal, 67–101. Berlin: F. Dümmler.
- Igartua, Ivan. MS. "Ciclos invertidos. En torno a la direccionalidad del cambio tipológico. To appear
- Ivić, Pavle. 1994. *Celokupna dela*, vol. 1. *O govoru Galipoljskih Srba* ed. by Dragoljub Petrović. Sremski Karlovci–Novi Sad: Izdavačka Knjižarnica Zorana Stojanovića.
- Jakobsen, Henrik Galberg and Peter Stray Jørgensen. 1999. *Håndbog i nudansk*. (Third edition.) Copenhagen: Politiken.
- Jakobson, Roman. 1931. "Principes de phonologie historique". Reprinted in his *Selected Writings*, 1. 1962: 202–220 and in Jakobson 1990: 184–201.
- Jakobson, Roman. 1957 [1971]. "Shifters, verbal categories and the Russian verb", Cambridge, Mass.: Department of Slavic Languages and Literatures, Harvard University. Reprinted in Jakobson 1971: 130–147 and in Jakobson 1990: 386–394.
- Jakobson, Roman. 1958 [1971]. "Morfologičeskie nabljudenija nad slavjanskim sklonenijem (Sostav russkix padežnyx form)". In *American Contributions to the Fourth International Congress of Slavists*, 127–156. 's-Gravenhage: Mouton. Reprinted in Jakobson 1971: 154–183.
- Jakobson, Roman. 1971. *Selected Writings*, vol. 2. *Word and Language*. The Hague: Mouton,
- Jakobson, Roman. 1990. *On Language* ed. by Linda R. Waugh and Monique Monville-Burston. Cambridge, Mass.–London, England: Harvard University Press.
- Janda, Richard D. 2003. "'Phonologization' as the start of dephoneticization – Or, On sound change and its aftermath: of extension, generalization, lexicalization, and morphologization". In Joseph and Janda 2003: 401–422.
- Joseph, Brian D. 2003. "Morphologization from syntax". In Joseph and Janda 2003: 472–492.
- Joseph, Brian D. and Richard D. Janda (eds.). 2003. *The Handbook of Historical Linguistics*. Malden, Mass.–Oxford–Melbourne–Berlin: Blackwell.
- Karskij, E. F. 1911 [1956]. *Belorusy. Jazyk belorusskogo naroda*, vols. 2–3. Moscow: Izdatel'stvo Akademii Nauk.
- Kuryłowicz, Jerzy. 1965. "The evolution of grammatical categories". *Diogenes* 51. Reprinted in his *Esquisses linguistiques*, vol. 2, 38–54. Munich: Wilhelm Fink Verlag.
- Langacker, Ronald W. 1977. "Syntactic reanalysis". In Charles N. Li (ed.), *Mechanisms of Syntactic Change*, 59–139. Austin: University of Texas Press.

- Lyons, John. 1977. *Semantics*, vol. 2. Cambridge: University Press.
- Maiden, Martin. 2006. "Lexical non-sense and morphological sense: on the real importance of 'folk etymology' for historical linguists". In this volume.
- Mańczak, Witold. 1963. "Tendances générales du développement morphologique". *Lingua* 12: 19–38.
- Meillet, Antoine. 1912. "L'évolution des formes grammaticales". *Scientia (Rivista di scienza)* 12, no. 26: 6–24. Reprinted in his *Linguistique historique et linguistique générale*, 130–148. Paris: Honoré Champion, 1965.
- Ordbog over det danske Sprog*, vols. 1–28. Copenhagen: Gyldendal, 1918–1956.
- Ottósson, Kjartan. 2006. "The diffusion of systemic changes through the inflectional system: Evidence from Nordic languages and German". In this volume
- Roberts, Ian and Anna Roussou. 2003. *Syntactic Change*. Cambridge: University Press.
- Sapir, Edward. 1921 [1949]. *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace & World.
- Schlegel, August W. 1818 [1971]: *Observations sur la Langue et la Littérature Provençales*. Neudruck der ersten Ausgabe Paris 1818 herausgegeben mit einem Vorwort von Gunter Narr: "August Wilhelm Schlegel – ein Wegbereiter der Romanischen Philologie" [Tübinger Beiträge zur Linguistik, 7]. Tübingen: Tübinger Beiträge zur Linguistik.
- Schøsler, Lene. 2006. "Argument marking from Latin to modern Romance languages. An illustration of 'combined grammaticalization processes'". In this volume.
- Shevelov, George Y. 1993. "Ukrainian". In Comrie and Corbett 1993: 947–998.
- Short, David. 1993. "Czech". In Comrie and Corbett 1993: 455–532.
- Skautrup, Peter. 1944. *Det danske Sprogs Historie*, Bind 1. *Fra Guldhornene til Jyske Lov*. Copenhagen: Gyldendalske Boghandel, Nordisk Forlag.
- Skautrup, Peter. 1968. *Det danske sprogs historie*, bind 4. *Fra J. P. Jacobsen til Johs. V. Jensen*. Copenhagen: Gyldendalske Boghandel, Nordisk Forlag.
- Slobin, Dan. 1985: "The child as icon-maker". In *Iconicity in Syntax. Proceedings of a Symposium on Iconicity in Syntax, Stanford, June 24–6, 1983* ed. by John Haiman [Typological Studies in Language, 6]. Amsterdam: John Benjamins.
- Tauli, Valter. 1958 *The structural tendencies of language*. [Annales Academiae Scientiarum Fennicae, series B, tomus 115, 1]. Helsinki: Suomalainen Tiedeakatemia.
- Timberlake, Alan. 1977. "Reanalysis and actualization in syntactic change". In *Mechanisms of Syntactic Change* ed. by Charles N. Li, 141–177, Austin: University of Texas Press.

- Traugott, Elizabeth C. 2003. "Constructions in grammaticalization". In Joseph and Janda 2003: 624–647.
- Whitney, William Dwight. 1867 [1870]. *Language and the Study of Language. Twelve Lectures on the Principles of Linguistic Science*. (Third edition.) London: N. Trübner.
- Wolfram, Walt and Ralph W. Fasold. 1974. *The Study of Social Dialects in American English*. Englewood Cliffs, N.J.: Prentice-Hall.
- Zaliznjak, Anatolij A. 2004. *Drevnenovgorodskij dialekt*. (Vtoroe izdanie, pererabotannoe s učetom materiala naxodok 1995–2003 gg.) [Rossijskaja Akademija Nauk. Institut Slavjanovedenija. Jazyk. Semiotika. Kul'tura]. Moscow: Jazyki slavjanskoj kul'tury.
- Žirmunskij, V. M. 1956. *Nemeckaja dialektologija*. Moscow: Izdatel'stvo Akademii nauk SSSR.