PARADIGM SIZE, MORPHOLOGICAL TYPOLOGY, AND UNIVERSAL ECONOMY

FRANS PLANK

I

As every typologist knows, some languages (called 'synthetic') have inflexional categories, some of them having more, some less, while other languages (called 'analytic' or 'isolating') lack inflexional categories.

Languages which have inflexional categories need not have exactly the same set; for example, some languages have inflexional Switch-Reference, others, maybe most, do not. Cross-linguistic variation is not entirely random in this respect, though: there are many notional categories (e.g. colour, kinds of disease, speed and perhaps manner of locomotion) which are never grammaticized as inflexional categories; other categories acquire inflexional status relatively rarely (e.g. Switch-Reference); some categories recur with more than chance frequency whenever languages have inflexional categories at all, including Number, Case, Gender/Class, Person, Possession, Definiteness and perhaps further deictic categories, Gradation, Tense, Aspect, Mood, and Diathesis. Incidentally, I do not regard the boundary between inflexional and derivational categories as categorical and always well-defined (cf. PLANK 1981, ch. 2), hence would not be surprised if generalizations were found which to some extent cover both inflexional and derivational categories. Evidently, inflexional languages can be classified according to the presence or absence of particular inflexional categories (e.g into Aspectual and Aspect-less, Casual and Case-less languages); but such classifications do not seem very popular among typologists, perhaps owing to the relative lack of cross-linguistic variation concerning the major inflexional categories, and also the apparent lack of correlates that would invest such classifications with systemic significance.

0165–4004/86/0020–29 $ 2,–

Languages which share particular inflexional categories may nevertheless differ in their categorial infrastructure, viz. in the number and kind of terms which realize the respective categories, and in this sense define the structure of inflexional paradigms. For example, languages sharing the inflexional category of Number may differ in that Number is either realized by two, three, four, or even more terms. Such variation in categorial infrastructure obviously has not gone unnoticed. Trivially, languages can be compared as to whether or not they manifest, for example, a Dual as one term of their category of Number. Less trivially, categorial infrastructures have been found to be to some extent amenable to universal implicational generalizations. For example, if Number is realized by Dual, this inflexional category is invaribly realized by Singular and Plural as well, whereas the presence of Singular and Plural terms implies nothing about the Dual.

The purely quantitative aspect of infrastructural variation, on the other hand, has attracted much less attention among typologists and others keen on cross-linguistic generalizations. Some inflexional categories (such as Switch-Reference or Definiteness) in fact appear to have a relatively uniform, and invariably relatively small, number of terms in all their language-particular instantiations, hence need not rouse the typologist, who thrives on differences, at least to get started. Term numbers may vary considerably across languages, however, with other inflexional categories, including Case, Gender/Class, Tense, Aspect, Mood, and Diathesis. And it would be unjust to accuse typologists of ignorance concerning such quantitative variation. To the extent that term numbers, hence paradigm sizes, correlate with the degree to which languages are synthetic or analytic, such variation indeed was bound to appear on the typological record. The relevant correlation is based on considerations of economy, and, not implausibly, states that morphological categories should tend to have the fewer terms and the smaller paradigms, the more alternative, especially non-morphological, means are available in a language to express the relevant distinctions: For example, there should be the fewer Cases, the richer are the analytic means (such as adpositions, relational nominals, serial verbs, word order) or other morphological means (such as agreement or derivation) exploited in a language to encode grammatical relations.

The recognition of such interrelationships is about all typology can take credit for as far as the question is concerned why certain

inflexional categories vary across languages in the number of terms realizing them. Even so, the preferred policy still seems to be to relegate such variation to the domain of the language-specific. Thus, DRESSLER (1985, §5), comparing the make-up of nominal and pronominal inflexional paradigms of Turkish, Hungarian, Latin, Russian, German, Italian, French, and English, carefully extracts features which tend to be type-specific or even type-constitutive (e.g. the presence or absence of Gender, inflexional classes, non-affixal exponents, cumulative exponents), but includes the widely variable number of Cases among the language-specific variables. Similarly, CARSTAIRS (1984a, 9—12), when dividing facts about Latin inflexional morphology into 'accidental facts about Latin' and 'facts possibly reflecting some general linguistic principle', tentatively classes the absence of a Dual, i.e. the fact that Number in Latin is a two-term rather than a three-term category, among the accidental facts, amenable at best to a diachronic explanation.

But also worth recalling here is at least one attempt, by BRØNDAL (1935), to specify an upper limit on the number of terms any inflexional category may possibly have. Taking into account only the 'elementary' categories of Mood, Aspect, Tense, Number, and Person, while ignoring allegedly more complex categories (such as Gender, Case and Diathesis) as well as 'croisements de deux categories' (such as between Aspect and Tense), BRØNDAL suggests that six is the maximal number of terms per elementary paradigmatic category — a number he arrives at by delimiting the possible kinds of paradigmatic oppositions. But even granting that there is such a universal upper limit, nothing follows from BRØNDAL's hypothesis about the actual cross-linguistic variation in the number of terms realizing particular categories — nothing that could explain, that is, why a given category may have the maximal six terms in some languages, but only five, four, three, or two terms (or only one, a possibility also admitted by BRØNDAL) in others. The same criticism may be levelled against HJELMSLEV (1935/1937), who suggests upper limits for a complex category ignored by BRØNDAL, viz. case: 216 terms for three-dimensional case systems (which is the 'maximum théorique', the 'maximum empirique' being 52 terms, in HJELMSLEV's experience), 36 terms for two-dimensional ones, and six terms for one-dimensional ones. Admitted, HJELMSLEV relates his maxima to another factor, the 'dimensions' of case systems

— but whether to utilize three, two, or only one dimension in turn appears to be an arbitrary language-particular decision.

The aim of this paper is to suggest that morphological typology may provide a clue to the question of paradigm sizes, beyond that afforded by the analytic/synthetic distinction, but likewise invoking economy. The relevant typological distinction is the time-honoured one between flective and agglutinative languages, which has come in for a great deal of harsh and maybe rash criticism due to its supposed lack of significant correlations with other parameters of cross-linguistic variation (e.g. by COMRIE 1981, 48f.). More accurately, it is a distinction of types of exponence of particular inflexional categories rather than of language types per se which I suggest bears upon the issue of cross-linguistically variable term numbers. Strictly speaking the ultimate points of comparative reference even are the individual terms of particular inflexional categories themselves. But, insofar as the types of exponence can be expected not to vary randomly from one term to the next of one inflexional category, and from one inflexional category to the next within one language, it may indeed turn out appropriate to refer, more comprehensively, to language types, defined in terms of the prevailing type of inflexional exponence.

The crucial distinction for present purposes is whether or not the exponents of an inflexional category, or of its individual terms, syntagmatically co-express other inflexional categories as well. Some technical terminology, partly familiar, will be helpful here. *Cumulative* exponents simultaneously express at least two co-occurring inflexional categories without being formally segmentable into two or more parts, while *separatist* exponents express only one inflexional category of a word form. Exponents can be further distinguished as *sensitive* and *insensitive*. Sensitive exponents are alternative expressions of the terms of an inflexional category, the choice among which is conditioned by lexical or morphological, rather than phonological (or maybe also semantic), properties of co-occurring morphemes; exponents are insensitive when lacking such alternants. Both cumulative and separatist exponents may be sensitive or insensitive. In particular, sensitivity does not perforce render separatist exponents cumulative. Suppose, for example, the exponents of a Case term are sensitive to Number as expressed elsewhere in the same word forms: Should this not be taken for an instance of extended exponence, with Number terms being syntag-

matically expressed twice, by (separatist) Number exponents and cumulative CaseNumber exponents? This conclusion would seem inappropriate if the Number exponents also appear independently of those particular Number-sensitive Case exponents, hence qualify as the *principal* exponents of Number (cf. CARSTAIRS 1984a, 32).

Sensitive cumulative and insensitive separatist exponents are the hallmarks of flective and agglutinative languages respectively, but are not the only ingredients of this typological distinction conceived of systemically (see in particular SKALIČKA 1979). What should be noted here is that the distinctions between cumulative and separatist and sensitive and insensitive exponents are not ones of 'techniques of construction' (in the sense of SAPIR 1921), even though they tend to correlate with the distinction of 'agglutinative' vs. 'fusional' and 'symbolic' techniques. Insensitive cumulative exponents seem to be regarded as still characteristic of flective languages, even if not prototypically so. Sensitive separatist exponents are often considered a deviation from the agglutinative norm, rather than a non-prototypical manifestation of flective morphology. This subordination of the sensitive/insensitive distinction to the cumulative/separatist distinction does not strike me as most felicitous, as will become clear, rather incidentally, from the following, where sensitive separatist exponents will be found to side with cumulative rather than insensitive separatist exponents. Note also that inflexional classes, where the choice among alternative exponents of inflexional categories is determined by individual nominal or verbal stems, are one manifestation of sensitivity, and declension and conjugation classes are more characteristic of flective than of agglutinative languages.

What I am suggesting, then, is this correlation between types of exponence and the number of terms of inflexional categories:

(TENT) Cumulative and sensitive separatist exponents limit the number of terms that may potentially realize inflexional categories more severely than insensitive separatist exponents.

It follows from TENT that the limits on term numbers should be the narrower, the more categories and terms exponents cumulate or are sensitive to. It does not follow from TENT that any type of

exponence is less suitable than any other to express categories of relatively few terms. Thus, a situation like that encountered in the extinct Iranian language Sogdian (cf. SIMS-WILLIAMS 1982), where 'light-stem' nouns (i. e. a phonologically defined class) had a six-term Case inflexion expressed by exponents cumulating Case with Number, whereas 'heavy' stems had only a two-term Case inflexion expressed by insensitive separatist exponents following the separatist Plural exponent, is not against the spirit of TENT. Specifying no absolute limit on the number of terms still conveniently expressible by cumulative and sensitive separatist exponents, TENT merely predicts that insensitive separatist exponents should prevail over cumulative and sensitive separatist ones when the inflexional terms to be expressed are really numerous.

Despite the somewhat vague quantitative specification of the threshold where it is supposed to become predictive, TENT is evidently an empirical hypothesis rather than a logical necessity: logically, nothing is to be said against very rich (say, 53-term) inflexional categories expressed by cumulative or sensitive separatist exponents. We are thus faced with two tasks: firstly, to establish the empirical validity, or at least plausibility, of TENT; secondly, to explain why things should correlate as suggested by TENT (if they do) rather than differently (with cumulative and sensitive separatist exponents admitting of richer categorial realizations) or not at all. As to the first task, what can be offered here is little more than a preliminary and global look at some encouraging evidence, preparing the ground for more thorough cross-linguistic investigations. Although TENT is meant to pertain to all inflexional categories with a variable and potentially reasonably high number of terms, particular attention will be paid here to Case, as that category whose numerical realization presumably varies most drastically — between two terms and, allegedly, 53 (attributed to one variety of the East Caucasian language Tabasaran). The second, explanatory task would seem to require less effort in principle: upon a little reflection it is not difficult to see why insensitive separatist exponents should admit of more richly structured categories than their cumulative and sensitive separatist counterparts.

II

To compare the number of Case terms across languages, we need to know which forms are Case forms in any language and into how many terms any paradigm of Case forms is to be divided — basic questions on which grammatical descriptions often enough fail to agree. The major difficulties consist in the reliable distinction of inflexional Case-marked word forms from adpositional phrases (a vexed question above all in languages with agglutinative morphology) and also from derivational word forms, and in the principled distinction of Case terms when forms are syncretic, with formal distinctions failing to correspond neatly to functional distinctions for all Case-marked items alike (which pattern seems more widespread with cumulative than with separatist exponents). It is impossible here to go to any lengths justifying the decisions taken whenever Case identifications and counts are controversial for any particular language. Nor is this really necessary: Occasional disagreements about identifying and counting Cases, unless they are on a massive scale or confined to languages of a particular morphological type, should not invalidate our general conclusion about TENT, which after all is partly statistical. At any rate, for present purposes I pragmatically tried to settle on numerical compromise solutions whenever language-particular descriptions I drew on disagreed about Case numbers and it was not immediately obvious which analysis was more adequate.

I also assume without argument that Case indeed is a unitary and elementary category (pace BRØNDAL). The category of Case may, thus, subsume terms marking semantically relatively transparent ('local' and other adverbial) as well as semantically more opaque ('grammatical') relations, and even relations outside syntactic clause structures proper (as do Vocatives). Exponents of Case terms are regarded as separatist rather than cumulative even if the relations encoded are notionally complex (e.g. have pragmatic, semantic, and syntactic ingredients, as does the subject relation in particular).

Now, one piece of suggestive evidence for or against TENT could be derived, it seems, from the average number of Case terms in sufficiently large samples of languages with different types of exponence for Case. What TENT leads one to expect is that the average should be higher with insensitive separatist than with cumulative

and sensitive separatist exponents, if perhaps not dramatically so because TENT does not rule out insensitive separatist expression of few-term categories.

But there are a few preliminaries that need to be attended to before embarking on this calculation. It would be ideal if there were only a single number of Case terms per language. Actual languages do not always live up to this ideal, though: the number of Case terms may differ depending on classes of lexemes and on the presence of other inflexional categories. When term numbers diverge widely depending on class memberships of nominal lexemes (as in Sogdian), it is advisable to count the different paradigms separately rather than to settle on an average value for the language as a whole. This policy has not been adopted here if such divergences involve nominal and suppletive pronominal Case marking, where nominal paradigms alone are taken into account. In all other cases the calculation of the cross-linguistic average is based on the highest number of Case terms productively manifested in the languages considered.

It would further be ideal if all Case terms in any language were expressed by the same type of exponence. In actual fact, however, Case and other paradigms occasionally turn out to be mixed in this respect. For example, in Basque, the three-term category Number and the eleven- or twelve-term category Case are usually expressed by (insensitive) separatist exponents, but the PluralAbsolutive exponent *-ak* is clearly cumulative. Similarly in Estonian, where two-term Number and 14-term Case are regularly expressed by separatist exponents in eleven Cases, but by cumulative exponents in the Nominative, Genitive, and Partitive. In Old English, all cumulative NumberCase exponents are sensitive to stem classes, except the PluralDative exponent *-um*, which is insensitive. As these examples indicate, one type of exponence usually prevails in any paradigm; and if the preponderance is considerable, it seems legitimate for the gross measure of the average term number per language to discount clear minority types of exponents. Thus, Basque and Estonian have been taken as exemplifying the (insensitive) separatist type of exponence. Tocharian A has not been included in the sample for which the average Case number has been calculated because the ratio of terms with cumulative (three 'primary' Cases) and separatist exponents (five 'secondary' Cases) is exceptionally well balanced. Strictly speaking, to examine the

validity of TENT, different types of exponents co-occurring in mixed paradigms ought to be considered separately.

The most common paradigmatically prevalent or exclusive exponent types are (mostly sensitive) cumulative exponents, the categories cumulated with Case being Number (most frequently), Possession, and/or Definiteness, and mostly insensitive separatist exponents. Languages with predominantly sensitive separatist Case exponents seem much rarer; and for lack of further specimens I have provisionally classed the only possibly pertinent language considered, Lapp (where Case exponents generally seem Number-sensitive), with the insensitive separatist set.

To turn from these preliminaries to the language samples, let me refer to Table 1 below for the full lists. The sample of languages with cumulative Case exponents includes 16 languages, all Indo-European with the exception of (West Greenlandic) Eskimo. The sample of languages with separatist Case exponents includes 31 languages, including Indo-European ones such as Sogdian (heavy-stem paradigm), Ossetic, and Bengali. Both samples are not random but designed to include languages exemplifying the maximal and the minimal or the approximate middle number of Case terms attested in the relevant language families or subfamilies. The Indo-European preponderance among the cumulative languages is obviously inevitable. Who is intent on redressing the genetic balance so far as possible, might include Arabic in the cumulative set, which arguably has a cumulatively expressed two-term Case category under certain circumstances (with external Plural), but otherwise (with internal Plural) three separatist Case terms.

With cumulative exponents, then, Case is realized by 5.6 terms on average, and with overwhelmingly separatist exponents by 8.6 terms, or, if the 48-term Case category of Tabasaran is excluded from the sample (for reasons given below), by 7.3 terms. Having experimented with different language samples, my impression is that 5.6 really is a maximal average value for cumulative exponents, while 7.3 rather strikes me as a minimal average for separatist exponents. These average term numbers, even if the difference is not necessarily huge, tend to support TENT — granting that average values indeed bear on TENT.

Average values do not reveal how the total has been arrived at: quantities of about uniform size may add up to the same total as quantities of extremely different sizes. Thus, the above average

value of 5.6 for languages with cumulatively expressed Case terms could equally well have resulted from adding up 16 languages with 5—6 Case terms each or from adding up four languages with 15, five languages with three, and seven languages with two Case terms. In the latter case, however, if the average value of 8.6 (or 7.3) for languages with separatist Case term exponents should turn out to have resulted from adding up 31 (or 30) languages with 8—9 (or 7—8) Case terms each, TENT would actually be disconfirmed: the number of terms maximally realizing Case, given this breakdown of the average values, would be higher with cumulative (15) than with separatist exponents (9/8). A closer look at the languages in our samples is imperative, therefore, to safeguard against rash interpretations of the evidence from average values.

To this effect I have arranged the languages (or, in the case of Sogdian, paradigms) considered in five groups depending on whether their number of Case terms is Minimal (2 terms), Small (3 terms), Medium (4—6 terms), Large (7—10 terms), or Extra Large (more than 10 terms). These five size classes are not as arbitrary as it may seem, but are roughly justifiable by cross-linguistic regularities in the elaboration of Case paradigms drawing on at least three different but potentially overlapping domains (grammatical, local, other adverbial). But for present purposes we need not dwell on the implicational generalizations underlying these class divisions. What counts is how languages (or paradigms) with cumulative and separatist exponents for Case distribute among these classes, and this is represented in Table 1.

Unlike the average values, Table 1 reveals immediately where the real difference lies between the two samples. No noticeable difference emerges with regard to the two classes at the lower end of the spectrum: both cumulative and separatist exponents are found to express few-term Case categories, but with either type of exponent such languages (or paradigms) are the clear minority. The majority with both types of exponents indiscriminately are Case categories with between four and ten terms, the ratio of representatives of the Large to those of the Medium class being slightly higher with separatist than with cumulative exponents. But where separatist exponents most conspicuously outrank cumulative ones is in the Extra Large class: no single language with cumulative exponents affords more than ten terms of Case, a luxury that is by no means the exception with separatist exponents. And this kind of

TABLE 1

Distribution of Case exponent types among classes of paradigm sizes

	Minimal (2)	Small (3)	Medium (4—6)	Large (7—10)	Extra Large (11 +)
Cumulative	Manx English	Hindi	Greek Albanian Latin*** Old Irish Old English Panjabi Sogdian (light stems)	Sanskrit Romany Lithuanian Russian** Polish Eskimo (W. Greenlandic)	
Separatist	Ubykh Swahili Yaqui Sogdian (heavy stems)	Gondi* Ostyak*	Gilyak Vogul Manchu Turkish Georgian Bengali Tarascan	Chukchi Nenets Lapp Ossetic Hurrian Kannada Dyirbal Yidiny Fore Quechua	Burushaski Basque Evenki Hungarian Finnish Estonian Tabasaran Udi

Notes: * Language could possibly belong to Minimal class
** Language could possibly belong to Medium class
*** Language could possibly belong to Large class

pattern is exactly what is predicted by TENT. Faced with Table 1, and presupposing (plausibly, I assume) that the empty cell of the matrix is no accident due to an inadequate language sample, one is even tempted to reformulate TENT more accurately: at least as far as Case is concerned, cumulative exponents (and sensitive separatist exponents: note that Lapp exemplifies the Large class) appear to limit the number of terms that may potentially realize this inflexional category to no more than ten terms, while this limit may be exceeded with insensitive separatist exponents of Case terms.

I have not systematically examined the implication of TENT that term numbers should be the lower, the more categories and terms exponents cumulate or are sensitive to. The indications, however, are that there are no significant additional constraints of this kind. Observe, for example, that many languages make it to the Large class in spite of exponents cumulating Case (8-term in Sanskrit, 8- and formerly 10-term in Lithuanian, 8- or 9-term in Eskimo) with 3-term Number (Sanskrit, Lithuanian, varieties of Eskimo) and also with many-term Possession (Eskimo), and in spite

of sensitivity of Case exponents to several stem-class and Gender distinctions. Or consider Old English, which has five Cases in adjectival paradigms as opposed to only four of nouns, although with adjectives Case exponents cumulate an additional two-term category, Definiteness (or whatever notion underlies the distinction of strong and weak declension). On the whole, it does not seem possible, thus to predict in such terms the class membership of languages or paradigms utilizing cumulative and sensitive exponents for Case. Cumulation of Case with one or two other inflexional categories of three or more terms and highly differentiated sensitivity certainly does not predispose languages or paradigms to membership in the Minimal or Small classes — rather on the contrary.

III

Why, then, should cumulative and sensitive separatist exponents differ from insensitive separatist ones in limiting the number of the terms of Case or other inflexional categories more severely, to about ten in the case of Case? I shall suggest that they do so simply for reasons of formal economy, but first want to dispose of an alternative diachronic explanation. Given that the vast majority of languages with cumulative (and sensitive) Case exponents historically derive from a language, Proto-Indo-European, with no more than eight Case terms, does this ancestry not suffice to account for the quantitative limitations observed? Not quite, because it is in principle possible for Case paradigms to expand diachronically, a kind of development exemplified by the acquisition of one or two further Case terms in Ossetic and Lithuanian, in Ossetic concomitantly with the exchange of cumulative for separatist Case exponents. And as other, mostly non-Indo-European languages provide additional illustration of more drastic increases in Case terms, usually drawing on postpositions (cf. the survey in Kahr 1976), a purely historical explanation for the limited size of Case paradigms evidently proves insufficient.

The economic explanation focuses on the number of exponents required to express the terms of co-occurring inflexional categories. This number is radically different with insensitive separatist exponents on the one hand and cumulative and sensitive separatist exponents on the other. With insensitive separatist exponence, the

number of exponents required to express distinctively the entire inflexional paradigm results from adding all terms of all inflexional categories. With insensitive cumulative exponence, the number of exponents required results from multiplying all terms of all categories expressed cumulatively by one another. With sensitive exponence, this number similarly results from multiplying all sensitive terms by the number of distinctions they are sensitive to.

Here is a simple example to illustrate this arithmetic. Imagine an inflexional paradigm involving 4-term Case and 2-term Number, yielding a combined paradigm of eight members. Utilizing insensitive separatist exponents throughout, one gets along with six of them in expressing distinctively all eight members of the paradigm:

(INSENSSEP)	Noun-N1-C1	Noun-N2-C1
	Noun-N1-C2	Noun-N2-C2
	Noun-N1-C3	Noun-N2-C3
	Noun-N1-C4	Noun-N2-C4

Utilizing insensitive NumberCase-cumulative exponents throughout, distinguishing all eight members of the same inflexional paradigm requires eight exponents:

(INSENSCUM)	Noun-N1C1	Noun-N2C1
	Noun-N1C2	Noun-N2C2
	Noun-N1C3	Noun-N2C3
	Noun-N1C4	Noun-N2C4

What the paradigm INSENSCUM can do without is the rule required in INSENSSEP to specify the relative position of the co-occurring separatist exponents. Utilizing separatist Case exponents which are sensitive to the Number distinction between N1 (Singular) and N2 (Plural), one even ends up with ten different exponents:

(SENSSEP)	Noun-N1-C1n1	Noun-N2-C1n2
	Noun-N1-C2n1	Noun-N2-C2n2
	Noun-N1-C3n1	Noun-N2-C3n2
	Noun-N1-C4n1	Noun-N2-C4n2

Taking into account the familiar typological correlation of separatist exponents and paradigms where one member is unmarked (and maybe, strictly speaking, even extra-paradigmatic) and systemat-

ically devoid of any overt exponent, the number of exponents required for INSENSSEP even reduces to four (viz. six minus two extra overt exponents for N1 (Singular) and C1 (Nominative or Absolutive)); and analogously for SENSSEP. Taking into account the familiar typological correlation of cumulative exponents with the fusional and symbolic techniques of construction, even more formal idiosyncrasies should need to be memorized in INSENSCUM, and of course in SENSCUM paradigms, as the number of inflecting lexemes grows.

The differently severe quantitative limitations TENT imposes on terms realizing inflexional categories, thus, are presumably not to be explained by reference to a magical number 10 marking the threshold to the exclusive domain of insensitive separatist exponents. Rather, the explanation, based on the arithmetical fact that it takes more cumulative or sensitive separatist than insensitive separatist exponents to express distinctively any multicategorial inflexional paradigm where at least one category has more than two terms, ought to invoke general considerations of linguistic economy, viz. constraints on the profusion of grammatical, as opposed to lexical, forms. If we accept that the fund of grammatical forms of each language *in toto*, and of forms catering for particular grammatical categories, should not be unnecessarily large, and at any rate incomparably smaller than the lexical resources, it follows that insensitive separatist exponents are the most economic solution conceivable, the more differentiated categories and paradigms get. Indeed, analytic markers of grammatical, local, and adverbial relations such as adpositions, relational nominals, or serial verbs ideally have the properties of being insensitive and separatist; and therefore come in handy whenever a many-term Case category, cumulatively expressed, threatens to offend against the requirements of formal economy. With such analytic devices as well as with insensitive separatist morphological exponents grammatical paradigms may grow relatively extensive before the sheer number of forms to be memorized becomes uneconomically large.

The number of distinctions which the grammatical apparatus for relation-coding may potentially be required to take care of, if difficult to specify precisely, clearly is substantial enough, in any language, to warrant economy considerations. Assuming that adpositions and other non-morphological (at least semi-)grammatical devices preferably encode local and other adverbial relations and

that their number, in any language, is limited to about 30 (BRØNDAL 1950 suggests that languages may have between nine and 27 adpositions, which on the whole seems a realistic estimate, regardless of BRØNDAL's own theoretical rationalization), and assuming furthermore that the total of non-local and non-adverbial relations preferably encoded by cases and other morphological devices will never exceed about six, the maximum number of relations requiring grammatical encoding would seem to be about 35. (BLAKE's 1930 count of about 65 potentially distinct relations strikes me as somewhat exaggerated, motivated by 'the laws of thought' rather than 'general grammatical principles'.) To encode all of these relations cumulatively and/or sensitively would call for a supply of forms that is not exactly negligible.

Notice, at any rate, that we are now seeking to explain the differential limitations TENT imposes on the number of terms realizing inflexional categories by a general constraint against an unnecessarily rich fund of grammatical forms. The limit suggested by TENT turned out to be fairly specific, disallowing more than ten Case terms when Case exponents are cumulative or sensitive separatist. Our appeal to formal economy, on the other hand, so far has not produced anything like a vaguely quantifiable limit. The specific suggestion of TENT thus remains unaccounted for, even though we can plausibly point to formal economy as providing the rationale why there should be some quantitative limit at all. The rest of this paper is devoted to exploring how this situation can possibly be remedied.

Consider Basque, which with its eleven to twelve Case terms barely attains membership in the Extra Large group of Table 1. Since its nominal inflexion further includes a 3-term category of Number (Indeterminate, Singular, Plural), Basque ideally could get along with as few as 14 inflexional exponents (or only 12, if Indeterminate Number and Absolutive Case remain without one) if exponents were consistently of the insensitive separatist type. In reality Basque indulges in the luxury of a few more inflexional forms, viz. some adaptive variants and the cumulative PluralAbsolutive suffix *-ak*, disregarding dialectal support of some Case distinctions by change of accent. If Basque were to utilize consistently cumulative exponents, it would need at least 36 to take care of an inflexional paradigm of the same size. May we conclude, then, that 36 different exponents of nominal inflexion are an amount that is beyond the limit dictated by formal economy requirements?

It seems that this can hardly be true, in view of the common assumption that Extra Large Case categories alone may be realized by as many as 48 (or, in one dialect, 53) terms in Tabasaran, or slightly fewer in other East Caucasian languages, requiring at least the same amount of separatist exponents to express the Case paradigm irrespective of further inflexional categories such as Number. But this view of the giant Case paradigms of East Caucasian languages is somewhat misleading, and quantitatively fairly exaggerated. Case paradigms there in fact consist of three or four subsystems the terms of two or three of which productively enter into syntagmatic combinations. There is usually a Medium-size, rarely a Large-size, array of grammatical Cases; and the numerous local Cases then are constructed by combining 3- to 7-term 'series' markers expressing dimensionality (e.g. 'within', 'below', 'upon', 'behind') with 3- to 6-term markers of directionality (expressing e.g. location, direction to or from), with another marker added at the end to express what HJELMSLEV (1935, 138ff.) calls 'cohérence/incohérence'. All this maximally adds up to a supercategory of about twenty simple terms, the numerous local 'Cases' resulting from term combinations. Assuming more or less consistently insensitive separatist exponence for these terms, and making provisions for (2-term) Number and maybe (2- to 6-term) Class inflexion, the amount of exponents required for fully distinctive noun inflexion turns out to be not much larger than about 25. And this quantity seems a realistic assessment of the number of noun-inflexional exponents actually utilized in this language family. It does not exceed the limit tentatively suggested in connection with Basque, whose number of Case terms at first sight seemed far less impressive.

It is probably Hungarian which provides the real yard-stick for determining the amount of inflexional exponents tolerable in languages favouring insensitive separatist exponence. Hungarian has variously been analysed as manifesting 16, 18, 21, 23, or 27 Cases, and its noun inflexion further includes 2-term Number (with zero exponence for Singular) and 6-term Personal Possession. Allowing for certain deviations from pure insensitive separatist exponence, the number of exponents required to express the entire inflexional paradigm again does not seem to exceed about 30. Consistently (insensitive) cumulative expression of the three inflexional categories could only be accomplished with a fund of about 240 exponents!

If we accordingly set the limit on the number of permissible (noun-)inflexional exponents at about 30, does this enable us to account for the limitation of Case terms to about 10 if their expression is cumulative and maybe sensitive? Hardly, it would seem, because a closer look at cumulatively expressed inflectional paradigms tends to reveal a much larger amount of members to begin with. Consider Latin, where six to seven Cases are cumulated with two Numbers, the exponents of these twelve to 14 paradigm members being sensitive to, roughly, a 5-way distinction of stem classes, which yields a total of 60 to 70 paradigmatic distinctions. Or Lithuanian, topping the Large group of Table 1: here eight to ten Case terms are cumulated with three Numbers and inflexional exponents are again sensitive at least to a 5-way stem-class distinction, yielding 120 to 150 paradigmatic distinctions. In the nominal inflexion of Eskimo (West Greenlandic) 8-term Case is cumulated with 2-term (dialectally 3-term) Number and 8-term Personal Possession (Possession marking being optional), yielding 16 (or, with Dual, 24) distinctions within the Unpossessed paradigm plus 128 (or, with Dual, 288) distinctions for the Possessed paradigm. Clearly, if these languages were found to provide as many different exponents as there are paradigmatic distinctions, i.e. up to a little more than 300, there would be no way of setting a limit on exponent number sufficiently low to account for the exclusively insensitive separatist exponence when Case terms number more than about ten. Recall that Hungarian would need only about 240 different exponents to express its entire noun-inflexional paradigm (insensitive-)cumulatively.

These languages with cumulative morphology, however, do not have at their disposal such an abundance of nominal exponents at all. For one thing, they need not use cumulative exponents throughout. Thus, in Eskimo Possession is expressed cumulatively with Absolutive and Relative Cases only, but essentially separately from the remaining six oblique Cases. And other inflexional forms too may well yield to further morphological segmentation than is often assumed in descriptive grammars of Eskimo and for purposes of the present survey (cf. HAMMERICH 1951).

Even more importantly, these languages tend to cut down on exponents by employing them multiply, viz. to express several paradigmatic distinctions rather than only a single one. For example, in Latin *-m* may express SingularAccusative in all five declensions (perhaps aided by different stem formatives), and SingularNomina-

tive and SingularVocative of 2nd declension Neuters. In Eskimo, *-(i)t* may express PluralAbsolutiveUnpossessed, PluralRelativeUnpossessed, and 2nd Singular Possessor/Singular Possession AbsolutivePossessed. Such syncretisms are a familiar feature of cumulative inflexional paradigms, but are far less familiar with separatist exponents. This unequal distribution is best motivated in terms of economy (as argued in particular by CARSTAIRS 1984b, §7): with separatist exponence in a Case-Number paradigm, for example, an additional exponent is needed if a Case distinction that is formally recognized in the Singular were to be syncretized in the Plural; with cumulative exponence, on the other hand, one exponent fewer is required if two Cases are allowed to coincide formally. Exploiting the potentials of syncretism, languages with richly structured, cumulatively expressed inflexional paradigms usually manage to get along with limited funds of grammatical exponents. In fact, no more than approximately 30 distinct exponents appear to be utilized in the nominal inflexion of Latin, Lithuanian, and Eskimo alike. In the adjectival paradigm of Old English, where inflexions cumulate 4- or 5-term Case, 2-term Number, 3-term Gender, and 2-term Definiteness, yielding 50 paradigmatic distinctions, the number of exponents is even limited to ten or eleven — about the same number required if all inflexional categories were expressed by separatist exponents.

In conclusion, I thus tentatively suggest that the number of exponents potentially available for nominal inflexion in any language is limited to about 30, another (so far) magical number. With insensitive separatist exponents more paradigmatic distinctions are expressible than with cumulative or sensitive separatist exponents. The more inflexional categories are cumulated and the more distinctions exponents are sensitive to, the smaller will be the overall inflexional paradigms constructible with 30 exponents. Cumulation of 10-term Case, the limit suggested by TENT, with another 3-term category (say, Number) would suffice to exhaust the maximally available fund of exponents — unless paradigmatic distinctions are syncretized. Note that the 10-term limit TENT imposes on Case terms is, strictly speaking, still unexplained: it is conceivable, for example, for a 15-term Case category to be cumulated with 2-term Number, requiring exactly 30 non-syncretic exponents. However, on account of the typological correlation of cumulation with sensitivity, the principled limitation on Case terms is bound to be lower:

10-term Case cumulated with 2-term Number and sensitive to a 2-way distinction of stem classes would already require 40 exponents—again unless paradigmatic distinctions fall victim to syncretism. And with this proviso there is the rub of our explanation. Granted that an arbitrarily large proportion of paradigmatic distinctions may be syncretized, the requirement of formal economy is almost rendered vacuous, allowing no reasonable limitations of the number of terms realizing inflexional categories. A 20-term Case and 3-term Number paradigm, after all, is comfortably expressible with 30 cumulative exponents if half of the paradigmatic distinctions are syncretized, i.e. every exponent expresses two NumberCase combinations. The present survey indirectly suggests that syncretism is universally subject to certain limitations, for otherwise it would be difficult to understand why Case cross-linguistically appears to be restricted to ten terms or less with cumulative and sensitive exponents, even though extensive exploitation of syncretism would easily allow Case to be realized by more than ten terms. But the laws of syncretism are not the subject of this paper.

Address of the author: *Frans Plank*
Fachgruppe Sprachwissenschaft
Universität Konstanz
Postfach 5560
D-7750 Konstanz 1 (West Germany)

REFERENCES

BLAKE, F. R. 1930.
"A semantic analysis of case", in: *Curme volume of linguistic studies*, ed. by J. T. HATFIELD, W. LEOPOLD & A. J. F. ZIEGLSCHMID, 34—49. Baltimore, Waverly Press (Language Monographs VII).
BRØNDAL, V. 1935.
"Structure et variabilité des systèmes morphologiques", reprinted in his *Essais de linguistique générale*, 15—24. Copenhague, Munksgaard, 1943.
BRØNDAL, V. 1950.
Théorie des prépositions. Introduction à une sémantique rationelle. Copenhague, Munksgaard (original Danish edition 1940).
CARSTAIRS, A. 1984a.
Constraints on allomorphy in inflexion. Indiana University Linguistics Club. (Abridged version of PhD thesis, London University, 1981.)
CARSTAIRS, A. 1984b.
"Outlines of a constraint on syncretism", *Folia Linguistica*, 18, 73—85.
COMRIE, B. 1981.
Language universals and linguistic typology. Syntax and morphology. Oxford, Blackwell.

DRESSLER, W. U. 1985.
"Zur Bedeutung der Sprachtypologie in der Natürlichen Morphologie", in: *Energeia und Ergon. FS E. Coseriu* (ed. J. ALBRECHT et al.) Tübingen, Narr.
HAMMERICH, L. L. 1951.
"The cases of Eskimo", *International Journal of American Linguistics*, 17, 18—22.
HJELMSLEV, L. 1935/1937.
La catégorie des cas. Étude de grammaire générale (Two Parts). Aarhus, Universitetsforlaget (Acta Jutlandica VII, 1/2).
KAHR, J. C. 1976.
"The renewal of case morphology: Sources and constraints", *Working Papers on Language Universals*, 20, 107—151.
PLANK, F. 1981.
Morphologische (Ir-)Regularitäten Aspekte der Wortstrukturtheorie. Tübingen, Narr.
SAPIR, E. 1921.
Language. An introduction to the study of speech. New York, Harcourt, Brace & World.
SIMS-WILLIAMS, N. 1982.
"The double system of nominal inflexion in Sogdian", *Transactions of the Philological Society*, 1982, 67—76.
SKALIČKA, V. 1979.
Typologische Studien. Braunschweig, Vieweg.