

Copyright 2006
February 2006 (this version)
Please contact the author before citing.

Explaining Exuberant Agreement

Alice C. Harris

SUNY Stony Brook
Centre for Advanced Study, Oslo

Explaining Exuberant Agreement¹

Alice C. Harris

Multiple (“exuberant”) agreement with a single argument is rare among languages of the world but is found in several languages of the Nakh-Daghestanian family. The origins of such a system in verbs are investigated in one of the languages of the family, Tsova-Tush. The paper explains both why exuberant agreement is typologically rare and why it occurs at all. It is argued that the typological rarity of this structure is explained by simple probability, ranging not over types but over events and conditions. The rarity of the structure is the result of the fact that so many events and conditions are required to bring it about.

1. Introduction

Several languages of the Nakh-Daghestanian (Northeast Caucasian) family, including Archi, Khinalug, Tabasaran, and Lak, have developed multiple (“exuberant”) agreement marking with a single argument. In this paper, the origins of such systems are investigated in verbs in one of the languages of the family, Tsova-Tush, not closely related to any of those named above.

- (1) (a) v-ux-v-erc'-v-ie² Tsova-Tush
v-back-V-return-V-AOR³
'he turned him back'

¹The research reported here was supported by the National Science Foundation under Grant Number BCS-0215523 and by the Centre for Advanced Study in Oslo in 2005. I am grateful to participants at the symposium “Linguistic Theory and Grammatical Change” and at its reprise for their helpful comments, especially to Dag Haug and Keren Rice, who provided extensive written comments. I thank also those in the audience at the University of Utrecht and at the Max Planck Institute, Leipzig, where this was presented as part of the conference “Rara et Rarissima”.

²The last morpheme in both examples is actually *-ie'*; Dešeriev did not write nasalization.

³The following abbreviations are used in glosses: ABS absolutive, ALL allative, AOR aorist, CAUS causative light verb, CON contact case, DAT dative, DIR directonal, ERG ergative, EV epenthetic vowel, EX exclusive first person plural, GEN genitive, IMPF imperfect, INST instrumental, INTR intransitive, INTRLV intransitive light verb, LOC locative, OBL oblique, ORD ordinal, PAST immediate past, PL plural, PRES present, RPT reported, SG singular, SUBJV subjunctive, SUPERESS superessive, TRLV transitive light verb, UNACC unaccusative (Georgian only); I, II numbers of classes (Chamalal only). In the text, CM is class marker, N-D Nakh-Dagestanian (North East Caucasian).

- (b) **y-ux-y-erc'-y-ie** (Dešeriev 1967: 239)
 Y-back-Y-return-Y-AOR
 'he turned her back'

My goal in this paper is to explain both why exuberant agreement is typologically rare and why it occurs at all. The explanation has two parts. First, it is hypothesized that multiple agreement of this kind is ordinarily eliminated, but that in languages of this family the presence of agreement is one of the factors that distinguishes one verb from another. Second, it is suggested that the typological rarity of this structure is not explained as the result of its inability to function, the difficulty of its acquisition, the difficulty of processing it, our innate endowment, or by any universal rule specifically outlawing it. Rather, it is suggested here, this structure results from a complex sequence of quite ordinary diachronic events.

The results reported here are relevant to theories that claim naturalness as a basis of morphological change, to theories of grammar optimization, and to the Elsewhere Condition, identity avoidance, and the representation of extended exponence in morphological theory.

In section 2, I introduce exuberant agreement and Tsova-Tush. Section 3 proposes the specific origin of exuberant agreement in this language. I discuss in section 4 the explanation of this phenomenon and in §5 the wider implications of my proposal for synchronic and diachronic linguistics.

2. Exuberant Agreement in Languages of the Nakh-Dagestanian Family

Tsova-Tush is a severely endangered member of the Nakh or Vainakh subgroup of the Nakh-Dagestanian (henceforth N-D) language family. The other members of this subgroup are Chechen and Ingush, which have somewhat similar phenomena, as do some more distantly related languages. This section makes specific proposals about the origins of forms, such as those in (1), with exuberant agreement.

2.1. Exuberant Agreement

Extended exponence (or multiple exponence) is a term used by Matthews (1972), and by many others since, to refer to the realization of a feature by more than one morpheme within a word. In exuberant exponence, illustrated in (1), more than one instance of the **same** morpheme realizes a morphological feature within a single word. Clearly exuberant agreement, then, is a subtype of extended exponence. The Old Georgian example in (2) illustrates extended exponence but not exuberant agreement.

- (2) da ušjulo x-i-kmn-n-es (1 Esdras 1:46-48) Old Georgian
 and iniquity 3-UNACC-do-PL-3PL
 'and iniquities were committed'

Here both of the suffixes indicate plurality of the nominal *ušjulo* 'iniquity', whose plurality is not in fact marked on the noun itself. The first suffix, *-(e)n*, marks plurality of subjects of unaccusative verbs and direct objects of transitive verbs (Harris 1985: 209-230). The second suffix, *-es*, marks the plurality of third person subjects. Since both realize the feature [+plural], this is one example of extended exponence. The prefix *x-* refers to the same nominal, but marks

person and grammatical relation, not number. Because *x-* and *-es* mark person of the same argument, they constitute a second example of extended exponence. Neither of these qualifies as exuberant in the sense defined above, since no morpheme in the Georgian example is repeated.

Although I have defined exuberant agreement as more than one occurrence of the same agreement morpheme, in this paper I also pay attention to the occurrence of more than two realizations of the same feature (by same or different morphemes), since both this and exuberant exponence are of particular interest to linguistic theory. Thus Yip (1998: 220), having categorized cases of identity avoidance, writes

“The first type [i.e. a language in which the same morpheme appears two or more times in a given word – ACH] is rare, perhaps non-existent, but it is not clear that the morphophonology underlies this: in most cases it seems likely that syntactic and morpho-syntactic principles will achieve this end without identity avoidance being involved at all.”

On the other hand, Noyer (1992) proposes a principle of Feature Discharge to regulate potential cooccurrence of exponents. Once a given feature has been realized by means of one (occurrence of an) affix, that feature has been “discharged” and cannot be realized again, even in a different position (“discontinuous bleeding”). This is modified to permit a set of morphosyntactic features to have a “primary” and a “secondary” exponent, where the latter is an exponent of more than one feature. Thus Yip assumes that exuberant exponence will not occur, while Noyer proposes a principle that would prevent realization of a single feature by more than two morphemes.⁴

2.2. Tsova-Tush

Tsova-Tush has the largest number of gender-classes (henceforth “classes”) of any N-D language – eight (Č’relašvili 1967). Holisky and Gagua (1994: 162-163), however, define only five classes, on the grounds that the remaining three are very small. The class markers (CM) are listed in Table 1, with an example of a noun in each class.

⁴The theoretical relevance of the data presented here is dependent upon these CMs occurring in a single word. The paper is preliminary, in that I have not yet been able to undertake the fieldwork that would provide specific evidence regarding wordhood. In fact some sources write the preverb as a separate word, while others write it with the verb. This is not deemed here to be of great importance, since the preverb is only one of six sources of CMs in Tsova-Tush. Arguments for wordhood are presented briefly in the last paragraphs of §§3.1, 3.2, 3.4 and 3.5.

SINGULAR	PLURAL	
v	b	<i>mar</i> ‘husband’
y	d	<i>pst’u</i> ‘wife’
y	y	<i>cark’</i> ‘tooth’
d	d	<i>bader</i> ‘child’
d	y	<i>bat’r</i> ‘lip’
b	b	<i>borag</i> ‘knit slipper’
b	d	<i>ča</i> ‘bear’
b	y	<i>bak</i> ‘fist’

Table 1. Major grammatical classes in Tsova-Tush
(after Holisky and Gagua 1994: 162)

As illustrated by the examples, class markers do not appear on most nouns.

In Caucasian linguistics it is conventional to refer to classes by number, but there are too many in Tsova-Tush for the reader to easily remember the corresponding class marker of each. Therefore, in glosses of Tsova-Tush, I give the singular class marker with nouns, because it does not generally show up on the noun itself. If the noun is in the plural in the example, I give the singular CM followed by the plural, e.g. ‘ear(**d/y**)-PL.ABS’, a noun of the fifth class in Table 1, conditioning the CM *d*- in the singular and *y*- in the plural. In glossing the verb, I simply repeat the marker, instead of providing a class number.

In Tsova-Tush, as in the proto-language, some verbs agree in class and number. The agreement trigger is the absolutive nominal, that is, the subject of an intransitive verb or the direct object of a transitive verb. In the proto-language, verbs that lacked class agreement had no agreement at all. Only agreeing verbs are illustrated here.

- (3) mux **d-oj-eⁿ** (Dict 491a)⁵
fly(**d**).ABS D-flee-AOR
‘The fly flew away.’
- (4) o-qu-s yoħ **y-ex-ö** (Dict 237b)
that.one-OBL-ERG woman(**y**).ABS Y-ask-PRES
‘He (distal) is asking (proposing to) the woman.’

Some adjectives also agree with their heads, as in (5).

- (5) vir-e-go-ħ **y-ax:eⁿ** lark’-i **y-a** (Dict 61b)
donkey(**d**)-OBL-ALL-LOC Y-long ear(**d/y**)-PL.ABS Y-be
‘The donkey has long ears.’

⁵Most of the examples in this paper are from Kadagiže and Kadagiže (1984) and are cited as “Dict”, with a page number. Superscript <ⁿ> represents nasalization of the vowel in word-final position (see Holisky and Gagua 1994: 160). Breve, <~>, over a vowel marks reduction (see Holisky and Gagua 1994: 153).

remains, only the CM prefix and various tense, mood, voice suffixes.⁸ It forms transitives.

- (8) aqsb-a-x gagn-i c'eg-y-o (Dict 759b)
 Easter(d)-OBL-CON egg(y)-PL.ABS red-Y.TRLV-PRES
 'At Easter they paint (lit. make) eggs red.'

This verb can also be used independently, as in (9).

- (9) vux-k' d-in-o-lo-s! (Holisky & Gagua 1994: 182)
 what.ABS-ever D-AOR-RPT-SUBJV-1SG.ERG
 'Whatever have I done?'

If one of the light verbs that take a prefixal CM incorporates a verb stem that also takes a CM, we have the beginnings of exuberant agreement.

- (10) darʃaⁿ it':-lye-čö k'lasi y-ot'-y-iyēⁿ (Dict 494a)
 Darejan(y).ABS ten-ORD-in grade Y-go.over-Y.TRLV-AOR
 'They took Darejan over into the tenth grade.'
- (11) k'ab y-ox-y-iyēⁿ (Dict 498a)
 dress(y).ABS Y-rip-Y.TRLV-AOR
 '[She] ripped the dress.'
- (12) k'nat ču-reⁿ kalak-ĩ v-eh-v-al-iⁿ (Dict 241b)
 boy(v).ABS in-from city-DIR V-steal-V-INTRLV-AOR
 'The boy stole away from home to Tbilisi.'

Thus, by adding light verbs to existing verb stems, together with their prefixal CMs, we get – in some cases – two CMs.

In (10-12) it is easy to see that the lexical roots in in complex verb forms (-ot'- 'go over', -ox- 'rip', and -eĥ- 'steal') have no tense-aspect-mood suffix, unlike the verbs in all other examples. In Tsova-Tush, verbs (apart from 'be', which is irregular in this regard) must have either a special ablaut marking of the present tense, a suffixal marker of tense-aspect-mood, or a suffixal formant of a non-finite form, such as a masdar or participle. The lexical root in the complex verb form has none of these, and thus it cannot be an independent verb form and a word distinct from the light verb. This confirms that complex verbs formed with light verbs are single words.

3.2. Second Innovation: Grammaticalization of Pronouns

More recently, Tsova-Tush, like a few of its sister languages, has independently innovated person-number-case agreement. Personal pronouns listed in (13) have become suffixes on the finite verb form, indicating agreement, usually with the subject, sometimes with an object (Holisky 1994: 144). Subjects of transitives occur in the ergative form of the pronoun,

⁸This verb has no root in the other Nakh languages, as noted also by Imnajšvili 1976 and by Pareulize 2005.

while subjects of intransitives are in the absolutive form.

(13)		Absolutive	Ergative	
	1SG	so	as	
	2SG	ho	aħ, aho	
	1EX	txo	atx, atxǒ	
	2PL	šu	aiš, aišŭ	(Data from Holisky and Gagua 1994: 173)

The first person inclusive, *vai*, occurs as a separate word immediately following the verb, not as a suffix (see Holisky and Gagua 1994: 177-178). Third person pronouns and nouns do not condition person-number-case agreement in the verb (see also Gagua 1952).

Suffixal agreement is illustrated in (14), where the older class marking is also present.

(14)	pst'u-yaš-e-n	t'atb-u- ⁿ	t'ard	d -eyš-n- as (Dict 234a)
	wife-sister(y)-OBL-DAT	silver-OBL-GEN	ring(d).ABS	D-promise-AOR-1SG.ERG
	'I promised a silver ring to my wife's sister.'			

In (14) the CM prefix *d-* agrees in class with the direct object, 'silver ring', while the final suffix, *-as* agrees in person, number, and case with the subject 'I', which does not, however, show up overtly in this particular example. In an intransitive, the two kinds of agreement are conditioned by the same nominal, as in (15).

(15)	so	osi	v -a-ra- sǒ (Dict 24a)
	I.ABS	there	V-be-IMPF-1SG.ABS
	'I (male) was there.'		

Both agreement affixes are conditioned by *so* 'I'; *v-* indicates grammatical class and number, while *-sǒ* indicates person, number, and case. The agreement suffixes are not part of the exuberant agreement, as I have defined it, but they are nevertheless interesting in this context. When the inherited verb takes a prefixal CM, a light verb with a CM occurs, and suffixal agreement is present, we have three layers of agreement, as in (16).

(16)	čxindur	d -ol- d -in- as (Dict 489a)
	sock(d).ABS	D-start-D-AOR-1SG.ERG
	'I started [e.g. to knit] a sock.'	

Holisky and Gagua (1994) refer to the agreement markers in (13) as suffixes, and I assume this analysis here. It is supported by the fact that before this suffix, the aorist suffix *-(i)n* does not condition nasalization of the preceding vowel and undergo word-final deletion (see footnote 4 above), as illustrated in (14) and (16). (This can be contrasted with (12), for example, where these processes do apply.) Their analysis can also be supported by the fact that a vowel does not undergo word-final reduction when followed by these suffixes, as shown in (15).

3.3. Third Innovation: Grammaticalization of Preverbs

- (21) teg-**b**-a-**b**-ol-**b**-ie (Dešeriev 1967:241)
 order-B-EV-B-begin-B-PAST
 ‘S/he just began to work.’

In (21), *teg-ar* is an intransitive verb meaning ‘be put in order’. From it, a transitive verb is derived regularly: *teg-d-ar* ‘do; put in order’. In the example, this is combined with another regular verb *d-ol-ar* ‘begin (intr)’, from which a transitive is derived regularly: *d-ol-d-ar*.

The argument for wordhood given in §3.1 applies also to compounds. The first element in compounds has neither a tense-aspect-mood suffix nor a suffix forming a non-finite form, and thus it cannot be an independent verb form and a word distinct from the second element in the compound. This confirms that compound verbs are single words. In (19) we see additional evidence that the compound is a single word. There is only one marker of the first person subject. First person singular conditions a person agreement suffix, as described in §3.2; and if the compound in (19) were really two words, each should have an agreement suffix.

3.5. A Further Source of Exuberant Agreement: Periphrastic Reported Tenses

Č’relašvili (1984) and Holisky and Gagua (1994: 180-181) describe a set of reported tenses (evidentials). These reported tenses are formed, at least historically, with an aorist reported form of the verb ‘be’. The copula alone is illustrated in (22), and the reported tenses are illustrated in (23-24).

- (22) st’ak’ v-a (Dict 24a)
 man(v).ABS v-is
 ‘He is a man.’
- (23) tet’-o-**d**-a-n-o (Holisky and Gagua 1994: 181)
 cut-PRES-D-be-AOR-RPT
 ‘she was apparently cutting it’
- (24) xet’:-(o)-**d**-a-n-o (Č’relašvili 1984: 108)
 read-PRES-D-be-AOR-RPT
 ‘apparently he was reading it, has read it’

Although these tenses are not frequently used, any one of them would increase the number of CMs in a given word.

While these are described as periphrastic, both (sets of) authors write them as a single word. This contrasts with the set of auxiliaries described in Holisky (1994), which she writes as independent words, some of which “show evidence of becoming suffixes” (1994: 157). There is additional evidence that the copula *d-a-* has become an affix and is no longer an independent word. Here the CM indicates agreement with the absolutive argument of the (main) verb, indicating that the copula is not an independent auxiliary. If ‘be’ were an independent word, it would agree instead with its subject.¹²

3.6. Summary

¹²This argument is used in Handel (2003: 137-138) to characterize the perfect and pluperfect of Ingush, which are cognate to the periphrastic reported tenses of Tsova-Tush.

Although three is the highest number of CMs I have found in a single example, there is the potential for as many as six CMs in Tsova-Tush: (i) before the preverb, as in (18), (ii) before the first conjunct in a compound verb, as in (19), (iii) as a transitive or intransitive marker of the first conjunct, as in (8) and (12), (iv) before the second conjunct in a compound verb, as in (20), (v) as a transitive or intransitive marker of the second conjunct, as in (21), (vi) before the formant of a periphrastic reported tense, as in (23). This number does not include the CMs in the auxiliaries, some of which also take a CM and are described as on the way to becoming suffixes (see Holisky 1994, especially p. 157).

4. Explanation

There are two parts to the explanation of why structures like (18) occur in Tsova-Tush and not, for example, in Norwegian. One part of the explanation involves grammaticalization of the inflection, and the other, the reason the structure is rare.

4.1. Loss or Retention of Trapped Inflection?

The first part of the explanation must confront the reality that when two units fuse, trapping inflectional morphology between them, either of two things can happen.

Harris and Faarlund (in press) provide both attested and reconstructed examples of inflection that is lost when the host to which it is attached and another word collide in grammaticalization. The authors argue that this process is not phonological, but morphological. Although phonological changes may also take place, of course, loss of inflection under these circumstances does not depend on phonological change. Faarlund and Harris's strongest examples, attested changes, include the loss of case in Mainland Scandinavian when trapped between a noun and a definite article, the loss of a person marker in Georgian when trapped between a main verb and an auxiliary, the loss of a verbal suffix in Mainland Scandinavian when trapped between a verb and a reflexive clitic, and other losses. Loss of trapped morphology preserves the iconicity of agreement (one morphological representation for one referential entity), and it preserves the universally preferred order of morphemes, with agreement on the outside.

We have seen the other possibility in Tsova-Tush examples above. It is clear that when a light verb incorporated a verb stem, the morphology trapped between them was not lost. When a preverb fused with a verb, trapping a CM, the CM was not lost. When two verbs were joined in a compound, trapping a CM, the result was the same, and the CM was retained. And when an auxiliary was grammaticalized as a formant of reported tenses, the CM was again retained. Retention of the trapped CM results in a counter-iconic structure; for example, there is no possible reading for (18) where each *y*- refers to a different entity. Retention of the trapped CM also results in a dispreferred morpheme order, with some agreement marking close to the root, instead of the preferred order, with agreement marking on the edges of the word. The development of both counter-iconic agreement and dispreferred morpheme order needs to be explained.

While multiple realizations of an agreement feature are unusual, they are not unique to N-D languages. They also occur, for example, in Skou (Donohue 2003), Kiranti (van Driem 1990, 1991, 1993, 1997), North Omoto (Hayward 1998), and Hocank (Helmbrecht and Lehmann, in press).

What determines what the outcome will be? At the time of the change which produced the first interior CM in Pre-Tsova-Tush, what led speakers to tolerate the trapped inflection, rather than to lose it, as in some other languages? I hypothesize that loss is the ordinary outcome, because it preserves iconicity and preferred word order. I further suggest that the explanation of the different outcome in Tsova-Tush lies in the fact that in this language, as in other N-D languages, only some verbs have (verb-initial) class agreement.¹³ This means that one lexeme can be distinguished from another by the presence vs. absence of the verb-initial CM. Examples are given in (25).

(25) Some verb pairs in Tsova-Tush:

ak'-ar 'burn (intr.stat.)	d-ak' -ar 'burn (intr. active)'
aʔ-ar 'say'	d-aʔ -ar 'give'
ebl-ar 'arrange, place; bail, carry liquid'	d-ebl -ar 'put, place; lay (egg); put down in'
ebc'-ar 'push, pull; weigh; milk; smoke'	d-ebc' -ar 'tie'
ot'-ar 'spread'	d-ot' -ar 'go, go over'
ot:-ar 'stand, stay'	d-ot:- ar 'pour into'
ol:-ar 'thread (e.g. needle); put on (e.g. clothing)'	d-ol:- ar 'put down (inside s.t.); lock'

In addition, some perfective/imperfective pairs differ in taking vs. not taking the initial CM, though these pairs may also differ in vowel quality. For example, **d-ek'**-ar 'fall' is perfective and contrasts with *ak'*-ar 'fall', which is imperfective. A similar pair is **d-at'**-ar 'run PERFECTIVE' vs. *it'*-ar 'run IMPERFECTIVE' (see Holisky 1994: 153 and Holisky and Gagua 1994: 179 for more).

The light verb that forms transitives has no actual root (see examples (8) and (9) above). If its CM had been lost, there would be no marker at all of transitivity. The light verb that forms intransitives, on the other hand, has the root *-al-*. While there is not currently a lexeme that contrasts with it in terms of the presence vs. absence of a CM, the examples in (25) show that this is a potential distinguishing feature of lexemes in Tsova-Tush. Therefore, I suggest that the loss of trapped morphemes is usual, that the trapped suffix on the core was lost as shown in Schema 1 (if indeed there was one), and that the CM prefix on the light verb was not lost because in Tsova-Tush this is a potential distinguishing feature of a lexeme.

Af1-	Core	-Af2	Af3-	Periphery	-Af4
v-	eħ	-AF2	v-	al	-i ⁿ



¹³Since agreement is inflection, it is usually assumed that it will occur with all members of the category verb. Sign languages, too, have agreement with only a subset of verbs; see Aronoff et al. (2005: 321-324) for a recent summary and references. In Tsova-Tush, unlike sign languages, the set of verbs taking agreement cannot be predicted on a semantic basis.

Af1-	Core	-Af2	Af3-	Periphery	-Af4
v-	eħ		v-	al	-i ⁿ

Schema 1. Fusion of *d-eħ-ar* ‘steal’ with *d-al-ar*.

Schema 1 uses the example of the verb in sentence (12), repeated here as (26).

- (26) k’nat ču-reⁿ kalak-ĩ v-eħ-v-al-iⁿ (Dict 241b)
 boy(v).ABS in-from city-DIR V-steal-V-INTRLV-AOR
 ‘The boy stole away from home to Tbilisi.’

While the Affix 2 position may have been empty from the beginning, in some related languages in which a verb form is incorporated by a light verb, an affix that nominalizes the verb is included. For example, in Udi, the infinite marker, *-es*, is included in this context.

Why was the verb-initial CM not lost when adverbs were grammaticalized as preverbs and when verb compounding was initiated? When a child acquires a language that already has an interior CM, she learns that this kind of interior inflection is **unmarked for this language**. For such a speaker, there would be no reason to lose the trapped inflection in the course of grammaticalization of a preverb.

I conclude that loss of trapped morphology is the expected outcome cross-linguistically. Tsova-Tush, in the first instance, namely grammaticalization of light verbs, presents an exception because class marking in N-D languages occurs with some, but not all, verbs. The presence vs. absence of a CM in these languages can potentially distinguish one lexeme from another and in this sense is part of the lexeme. In grammaticalization of the verbs *d-al-ar* INTRANSITIVE and *d-ar* TRANSITIVE as light verbs, the CM was not lost because it was part of the lexeme in this unusual sense. When the grammaticalization of adverbs as preverbs trapped the verb-initial CMs, and when the innovation of verb compounding trapped various CMs, interior CMs had been established as a language-specific unmarked feature. Children acquiring the language had no reason to lose these interior agreement markers.

4.2. Simple Probability

Let us turn to the issue of why structures such as those in (18) are rare. We might attribute the relative rarity of exuberant agreement to one or more of the following factors.

- our innate endowment discourages this structure (perhaps as part of a more general feature)
- this structure does not function well
- this structure cannot be acquired easily by children
- this structure is not easily processed.

All of these may, in fact, be true. True or not, each one leaves us wondering how such a structure can exist at all. Do speakers of Tsova-Tush not know that this is innately discouraged? Do they not need a language that functions well? Are their children able to acquire structures others would fail at? Are the speakers able to process speech better than the rest of us?

I have proposed in earlier papers (Harris 2005, in press) that the rarity of certain other phenomena are to be explained by probability, and I suggest that the development of this system too is a simple matter of probability. As we know, it is extremely common for an auxiliary to be added to a main verb, or for a light verb to incorporate a verbal or other base, as Tsova-Tush did in adding the light verbs *d-ar* ‘make, do’ and *d-al-ar*, the formant of intransitives. It is likewise extremely common for a subject pronoun to cliticize to a verb and become an affix, as *so* ‘I.ABS’, *as* ‘I.ERG’, and other pronouns did in Tsova-Tush. We know that many languages, including Latin, Greek, Germanic, Slavic, and Kartvelian languages, have turned adverbs into preverbs. Compound verbs may be somewhat less common than the others, but they are certainly not rare. One way in which Tsova-Tush differs from the other languages that have undergone these very common changes is that in Tsova-Tush all four have occurred. For this system to develop, it was, of course, essential that Tsova-Tush had agreement on verbs and on the verbs that became light verbs, then transitivity markers. I assume that it is entirely irrelevant that inherited agreement in Tsova-Tush happens to be class agreement, but in any case this in itself is not uncommon (the Bantu languages are probably most famous for it). It was also apparently essential that only some verbs in Tsova-Tush have agreement, and hence that the presence vs. absence of inherited class agreement has the potential of distinguishing one lexeme from another. This may be somewhat unusual, but the main thing that makes Tsova-Tush uncommon with respect to exuberant morphology is that it combines all these attributes and changes:

- fusion of light verbs as formants of transitive/intransitive structures
- fusion of a subject pronoun as an agreement affix
- fusion of an adverb as a locative preverb
- introduction of verb compounding
- fusion of an auxiliary as a marker of evidentiality
- inherited class agreement on some, but not all, basic verbs
- inherited class agreement on some light verbs
- inherited class agreement on an adverb/preverb
- agglutinative morphology elsewhere.

Few languages have exuberant agreement because it is simply statistically improbable that many languages would combine all of these conditions and changes.

But couldn’t exuberant exponence – two occurrences of the same morpheme – occur just through one of the five changes, without involving all five? This would seriously compromise the explanatory value of complexity and probability. It is true that one change, together with appropriate conditions, could give rise to two occurrences of the agreement marker. I have presented the full complexity of the Tsova-Tush case because its relevance to linguistic theory as a counterexample to Noyer’s (1992) proposal is dependent on the fact that it exemplifies more than two realizations of the same feature. Consider further that a language with two instances of an agreement marker is somewhat unusual and requires one change, with appropriate conditions; while a language with three instances is considerably more unusual and requires two changes, with appropriate conditions; and a language such as Tsova-Tush, which can potentially take many morphemes realizing a single feature is very rare and requires many changes, also with

certain conditions. Thus, the rarity of the phenomenon is in proportion to the number of change (and conditions) required to reach it, as would be predicted.

But it is lexical verbs that are shown in (25) to occur in minimal pairs, where the presence of the CM distinguishes one verb from another. Do light verbs also occur in minimal pairs, so that loss of the CM would confuse two light verbs? The light verb that forms transitives has no actual root (see examples (8) and (9) above); even as an independent verb, it consists only of the CM and tense-aspect-mood marking. If its CM had been lost, there would be no marker at all of transitivity. This is discussed in greater detail above in §4.1.

But perhaps there are other pathways to exuberant agreement? In another language exuberant agreement might develop somewhat differently. Nevertheless, I submit that no simple pathway could produce exuberant agreement. As discussed above, only multiple instances of fusion or compounding will introduce multiple copies of agreement. While agreement on some but not all basic verbs may not be essential, it is likely that some other characteristic distinguishes between languages that develop exuberant agreement and other languages.

Crucial to the explanation above is the fact that some verbs in Tsova-Tush have agreement, while others lack it; and this in itself is rare. How did this arise? Cercvaže (1970) has pointed out that throughout the N-D family, agreement markers occur in word-initial position only with verbs whose stems begin with vowel. Although there is no consensus on the exact form of the agreement markers in Proto-Nakh-Dagestanian, they are uniformly reconstructed as consonants, generally of approximately the form of those in Tsova-Tush, *w-, *y- (or *r-), *b-, *d- (see for example Alekseev 1985, 1988, and Schulze 1992). Some languages, such as Lak, have agreement markers before consonants in word-internal position, but not in word-initial, e.g. *d-a-r-čuna* ‘went’, where both *d-* and *r-* are markers of class agreement. Cercvaže observes further that throughout the N-D family there is a phonotactic constraint against a cluster of consonants in word-initial position. While we cannot be certain of developments with such time depth, it is likely that the use of agreement with only certain verbs is due to (i) this common phonotactic constraint, coupled with (ii) the fact that agreement markers were consonants and (iii) were prefixes, and (iv) the fact that some verbs were V-initial and others C-initial. These common conditions seem to explain the fact that some verbs agree in these languages, while others do not. And this, in turn, is a key to explaining the retention of exuberant agreement.

Notice that my explanation does not suffer from the drawback of the others described above. I have not claimed that any linguistic principle makes exuberant agreement uncommon, and therefore I have no need to explain why that principle is relaxed for Tsova-Tush.

I conclude that exuberant agreement is rare in part because the common changes and common conditions that lead to it are rare in combination – simply because of the low probability of any nine features being combined in one language.

5. Discussion: The Wider Implications of this Explanation

It is not claimed here that the low probability of combinations of historical events and other factors explains all rare phenomena. In Harris (2005) and Harris (in press) it is argued that the same considerations account for the existence of three distinct case marking systems in Georgian, a situation identified only in languages of the Kartvelian family, as far as I am aware, and for the existence of endoclitics situated inside root morphemes in Udi, the only language for

which such a system has been identified.

While the phenomena mentioned above are extremely rare, the same considerations, together with others, can account for the relative frequency of prefixes and suffixes, on the one hand, and the comparative rarity of infixes and circumfixes, on the other. We may assume that prefixes and suffixes develop in two changes, from independent words, via a clitic stage. It may be assumed that infixes develop from prefixes or suffixes through an additional step (cf. Yu 2003). Circumfixes, on the other hand, develop from a prefix (two steps) and a suffix (two additional steps) that become linked (a fifth step).

(27)	prefix	2 steps
	suffix	2 steps
	infix	3 steps
	circumfix	5 steps

Although (27) is perhaps over-simplified, it predicts that prefixes and suffixes will be equally widespread, infixes will be somewhat less common, and circumfixes will be considerably less common. Clearly one must take other factors into account also. One that is important is the fact that once an affix of any type is established, it sets a precedent in a particular language; it becomes part of system congruity, in the sense of Wurzel (1989), and we may then expect other tokens of this type to develop in this language. Another point of importance is that once an affix of any type is established in a language, its daughters inherit the affix, and its type becomes part of the system congruity of the family or subgroup. An affix type established early in a very large family may be expected to be found today in a large number of languages, while a type established equally early in a very small family may be expected to be found in few. This also contributes to the large or small number of languages having affixes of any given type.

While this approach may not explain the rarity of all phenomena, it is based on a sound footing, simple probability. In contrast, other explanations proposed to date – innateness, functionality, ease of processing, ease of acquisition – are not so firmly based. (i) There is no direct evidence, only indirect, to indicate what information our innate endowment provides. For the most part, evidence that a specific structure functions poorly, or is difficult to acquire, or is difficult to process is also lacking.¹⁴ (ii) The claim that a phenomenon is difficult to acquire, difficult to process, dysfunctional, or innate may be based on the fact that it is rare. Yet if this is offered as an explanation of its rarity, the reasoning is circular. (iii) Tsova-Tush has not been documented over a long period of time, but both Georgian and Udi have been, and it is known that the rare phenomena mentioned above from these languages have lasted for well over a millenium. It is not clear why such phenomena would demonstrate such longevity if they are contrary to our innate endowment, dysfunctional, difficult to acquire, or difficult to process. (iv) Finally, as mentioned above, it is not clear why rare phenomena should occur at all if they are to be explained in one of these four ways. On the other hand, the approach advocated here faces

¹⁴ But there is evidence concerning acquisition of the rare three coexisting systems of case marking in Georgian, and the evidence indicates that it is not difficult to acquire (Imedadze and Tuite 1988).

none of these problems. (i) Events are expected to be distributed according to probability. (ii) The reasoning here is not circular. (iii) It is not predicted that phenomena that require many steps or conditions to develop would disappear rapidly, and thus there is nothing to explain regarding the longevity of some rare phenomena. (iv) This approach explains why a rare phenomenon can occur.

6. Conclusions

In this paper I have shown that “exuberant” agreement – many instances of identical class-number agreement – may occur in the Tsova-Tush verb, together with a marker of person-number agreement. I have argued that this developed because the CM in Tsova-Tush, as in other N-D languages, is used with only some of the verbs of the language and thus helps to distinguish one verb from another. In the first instance, it was essential to maintain the CM on the grammaticalized light verbs as part of the identification of those verbs. Exuberant agreement is rare, it is argued here, because developing such a system requires many steps and conditions, and the probability of this combination of factors all falling together is low.

While I have not claimed that the probability of the occurrence of events and conditions explains the rarity of all rare phenomena, I have argued that this approach does explain many such phenomena, and that it should be taken into account as one important element in explaining the distribution of types generally. I have tried, in addition, to cast doubt on innateness, functionality, ease of acquisition, and ease of processing as explanations for the distribution of language types.

References

- Adams, Douglas Q. 1988. *Tocharian historical phonology and morphology*. New Haven: American Oriental Society.
- Alekseev, M.E. 1985. *Voprosy sravnitel'no-istoričeskoj grammatiki lezginiskix jazykov: Morfologija, sintaksis*. Moskva: Nauka.
- Alekseev, M.E. 1988. *Sravnitel'no-istoričeskaja morfologija avaro-andijskix jazykov*. Moskva: Nauka.
- Aronoff, Mark, Irit Meir, and Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81: 301-344.
- Bokarev, A.A. 1949. *Očerok grammatiki čamalinskogo azyka*. Moskva-Leningrad.
- Cercvaže, Ilia. 1965. *Andiuri ena* [The Andi language]. Tbilisi: Mecniereba.
- Cercvaže, Ilia. 1970. K'lasnišniani da uk'lasnišno zmnebi xunžurši. [Classed and classless verbs in Avar.] *IK'E* 17: 189-191.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Č'relašvili, K'onst'ant'ine. 1967. Gramt'ik'uli k'lasebis raodenobis šesaxeb C'ovatušur enaši. [On the number of grammatical classes in the Tsova-Tush language.] *Orioni: Ak'ak'i Šaniže*, ed. by G. Axvlediani, A. Šaramiže, I. Imnaišvili, A. Urušaže, G. Šalamberiže, Š. Žižiguri, K'. Č'umburiže, 332-335. Tbilisi: University.
- Č'relašvili, K'onst'ant'ine. 1984. C'ovaturšuri zmnis uylilebis sist'ema. [The conjugation system of the Tsova-Tush verb.] *Macne* 1984.2.106-112.
- Dešeriev, Ju.D. 1953. *Bacbijskij jazyk*. Moskva: Akademija.
- Dešeriev, Ju.D. 1967. Bacbijskij jazyk. *Jazyki narodov SSSR, IV: Iberijsko-kavkazskie jazyki*, ed. by K.V. Lomtadze (chief), E.A. Bokarev, Ju.D. Dešeriev, G.B. Murkelinskij, M.A. Kumaxov, S.M. Xajdakov, A.K. Šagirov, 228-246. Moskva: Nauka.
- Donohue, Mark. (2003). Agreement in the Skou language: A historical account. *Oceanic Linguistics* 42. 478--498.
- Driem, George van. (1990). An exploration of proto-Kiranti verbal morphology. *Acta Linguistica Hafniensia* 22. 27--48.
- Driem, George van. (1991). Bahing and the Proto-Kiranti verb. *Bulletin of the School of Oriental and African Studies* LIV (2). 336--356.
- Driem, George van. (1993). The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* LVI (2). 292--334.
- Driem, George van. (1997). A new analysis of the Limbu verb. In Bradley, David (ed.), *Tibeto-Burman languages of the Himalayas* (Papers in Southeast Asian Linguistics 14). Canberra: Pacific Linguistics. 157--173.
- Gagua, Rusudan. 1952. Zmnis cvla gramat'ik'uli k'las-k'at'egoriis mixedvit bacbur enaši. [Change in verb (form) according to the category of grammatical class in the Tsova-Tush language.] *Iberul-k'avkasiuri enatmecniereba* 4: 129-134
- Handel, Zev. 2003. Ingush inflectional verb morphology: A synchronic classification and historical analysis with comparison to Chechen. *Current trends in Caucasian, East European, and Inner Asian linguistics*, ed. by Dee Ann Holisky and Kevin Tuite, 123-

175. Amsterdam: Benjamins.
- Harris, Alice C. 1985. *Diachronic syntax: the Kartvelian case*. (Syntax and Semantics, 18.) New York: Academic Press.
- Harris, Alice C. 2002. *Endoclitics and the origins of Udi morphosyntax*. Oxford: Oxford University Press.
- Harris, Alice C. 2005. The challenge of typologically unusual structures. *Morphology and linguistic typology: On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4), Catania, 21-23 September 2003*, ed. by G. Booij, E. Guevara, A. Ralli, S. SgROI, S. Scalise, 277-284. 2005.
<http://morbo.lingue.unibo.it/mmm/mmm4-proceedings.php> (ISSN number 1826-7491).
- Harris, Alice C. In press. On the explanation of typologically unusual structures. *Language Universals and Language Change*, ed. by Jeff Good. Oxford: Oxford University Press.
- Harris, Alice C., and Jan Terje Faarlund. In press. Trapped morphology. *Journal of Linguistics*.
- Hayward, Richard. (1998). The Origins of the North Omoto verb agreement systems. *Journal of African Languages and Linguistics* **19**. 93--111.
- Helmbrecht, Johannes, and Christian Lehmann. In press. Hocank's challenge to morphological theory. *A World of Many Voices: Lessons from Documented Endangered Languages*, ed. by David Rood, David Harrison, and Arianne Dwyer.
- Holisky, Dee Ann. 1987. The Case of the Intransitive Subject in Tsova-Tush (Batsbi). *Lingua* **71**: 103-132.
- Holisky, Dee Ann. 1994. Notes on auxiliary verbs in Tsova-Tush (Batsbi). *Non-Slavic Languages of the USSR: Papers from the Fourth Conference*, ed. by Howard I. Aronson, 143-159. Columbus, OH: Slavica.
- Holisky, Dee Ann, and Rusudan Gagaa. 1994. Tsova-Tush (Batsbi). *The Indigenous Languages of the Caucasus, 4: North East Caucasian, Part 2*, ed. by Rieks Smeets, 147-212. Delmar, N.Y.: Caravan Press.
- Imedadze, Natela, and Kevin Tuite. 1988. The acquisition of Georgian. *The crosslinguistic study of language acquisition*, vol. 3, ed. by Dan Isaac Slobin, 39-109. Hillsdale, NJ: Erlbaum.
- Imnajšvili, D.S. 1976. Sistema obrazovanija vremen anomal'nogo sprjaženija v naxskix jazykax. *C'elic'deuli* 3. [[Incomplete]]
- Kabak, Bariş. 2006. An obstacle to the morphologization of postpositions. *Studies in Language* **30**: 33-68.
- Kadagiže, Davit, and Nik'o Kadagiže. 1984. *C'ova-tušur-kartul-rusuli leksik'oni*. [Tsova-Tush-Georgian-Russian dictionary.] Tbilisi: Mecniereba.
- Kibrik, A. E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka*. Tom I, *Leksika, Fonetika*; Tom II, *Taksonomičeskaja grammatika*; Tom III, *Dinamičeskaja grammatika*. Izdatel'stvo Moskovskogo Universiteta.
- Korhonen, Mikko. 1996 [1979]. Entwicklungstendenzen des finnisch-ugrischen Kasussystems. *Typological and historical studies in language by Mikko Korhonen*, ed. by Tapani Salminen 165-178. Helsinki: Suomalais-Ugrilainen Seura.
- Lomtaže, E. 1956. Čvenebiti nacvalsaxelebis agebuleba da bruneba didouri jgupis enebši. [The structure and declension of demonstrative pronouns in the Dido group of languages.]

- IK'E* 8: 269-305.
- Magometov, A.A. 1965. *Tabasaranskij jazyk*. Tbilisi: Mecniereba.
- Matthews, P.H. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.
- Nichols, Johanna. 2004. *Ingush-English and English-Ingush dictionary*. London: RoutledgeCurzon.
- Nichols, Johanna, and Arbi Vagapov. 2004. *Chechen-English and English-Chechen dictionary*. London: RoutledgeCurzon.
- Noyer, Robert Rolf. 1992. *Features, positions and affixes in autonomous morphological structure*. Ph.D. diss., MIT.
- Pareulize, Rost'om. 2005. Xmovanpužiani zmnebi čečnurši. *IK'E* 35: 272-276.
- Schulze, Wolfgang. 1992. Zur Entwicklungsdynamik morphosyntaktischer Subsysteme. *Caucasologie et mythologie comparée*, ed. by Catherine Paris, 335-362. Paris: Peeters.
- Talibov, B.B. 1961. Sistema grammatičeskix klassov v C'axurskom jazyke. *Voprosy izučeniya iberijsko-kavkazskix jazykov*, ed. by E.A. Bokarev, 214-226. Moskva: Akademija.
- Topuria, Varlam. 1937. Zogierti brunvis genezisisatvis megrul-č'anurši. [On the origin of certain cases in Mingrelo-Laz.] *Enimk'is Moambe* 1.179-82.
- Uuspõld, Ellen, and A. Valmet. 2001. *Populjarnaja grammatika èstonskogo jazyka*. Tallinn: Valgus.
- Wurzel, Wolfgang Ullrich. 1989. *Inflectional morphology and naturalness*. Dordrecht: Kluwer.
- Yip, Moira. 1998. Identity avoidance in phonology and morphology. *Morphology and its relation to phonology and syntax*, ed. by Steven G. Lapointe, Diane K. Brentari, and Patrick M. Farrell, 216-246. Stanford: CSLI.
- Yu, Alan C. 2003. *The morphology and phonology of infixation*. Ph.D. Dissertation. University of California, Berkeley.