

Third Cycle Agree Effects in Mordvin

Doreen Georgi*

Abstract

In this paper I claim that a probe which searches for a feature X can exceptionally be valued by a feature Y (where X and Y are variables over person and number) if i) X and Y form a natural class and if ii) massive probe impoverishment as suggested in Béjar (2003) applies. As a consequence of impoverishment, the probe, which I call *chameleon probe*, is underspecified for the exact feature it searches for and can be valued by any member of the natural class of features. I exemplify this phenomenon in the Uralic language Mordvin, in which a transitive verb can agree with its subject *and* object if the latter is definite. The *chameleon probe* approach allows the derivation of the unexpected distribution of person markers in the Mordvinian definite declension by an independently motivated change in Béjar's (2003) system, yielding third cycle Agree effects.

1. Introduction

In the Uralic language Mordvin, a transitive verb can potentially agree with both of its arguments in person and number. Whether agreement with an argument actually takes place or not depends on the grammatical function and feature values of that argument such that hierarchy effects emerge. Béjar (2003) develops an analysis which captures the distribution of markers across the paradigm with a single person and a single number probe. If, however, both arguments of a transitive verb are singular, double person marking arises, although the verb can usually cross-reference person only once. Béjar's analysis excludes such a pattern and cannot explain how it can be generated. I claim that this pattern can indeed be derived and is even expected to occur with

* For valuable comments I would like to thank the audiences of the *Workshop on Portmanteaus* at the DGfS annual meeting 2010 in Berlin, the *Workshop on Theoretical Morphology 5* in Wittenberg 2010, and *SICOGG 12* in Seoul 2010, as well as the participants of the syntax colloquium at the University of Leipzig, and especially Daniela Henze, Stefan Keine, Gereon Müller, Andrew Nevins, Martin Salzmann, Jochen Trommer, and Eva Zimmermann for discussion. Research was carried out in the project *The internal Structure of Person Portmanteaus* as part of DFG research group 742 *Grammar and Processing of Verbal Arguments*.

a minor and independently motivated change in Béjar's analysis. Moreover, the analysis reveals a yet unnoticed strategy of probes: An X-probe searching for a feature X on a goal can be valued by feature Y rather than remaining unchecked. For concreteness, I suggest that a number probe can exceptionally be saturated by person values. I call such a probe *Chameleon probe* because it adapts itself to the morphosyntactic environment.

The problematic double person marking pattern, the derivation of which is the main subject of this paper, can be observed if the complex verbal suffixes are subanalysed into strings of submarkers. The first part of the paper is thus devoted to the demonstration that such a segmentation is not only empirically motivated but also conceptually desired because it allows to derive transparadigmatic syncretism. As a by-product of it, we will see that there are no portmanteau morphemes (in a sense to be made precise) in Mordvin, contrary to claims in the literature.

The paper is structured as follows: In section 2, I introduce basic facts about Mordvin and in particular about its verbal inflection. In the subsequent section, I outline why the verbal suffixes should be subanalysed. Afterwards, I discuss the advantages and shortcomings of the generalizations about the segmented forms proposed by Béjar (2003) and I add a new generalization which leads to an improved analysis. In Section 5, I provide my theoretical background assumptions and show how *Chameleon probes* arise. I then go through the derivations in the definite conjugation in detail. Afterwards, I address predictions of the analysis with respect to morphological exponence. Finally, in section 7, it is shown that the system developed so far is also able to derive the distribution of person and number suffixes in the indefinite conjugation. Section 8 concludes.

2. Argument Encoding in Mordvin

Mordvin is a Uralic language spoken in the Mordva Republic in Russia. It subdivides into two dialects, Erza and Moksha, which differ mainly in their inventory of phonological rules and the lexicon (for a comparative description see Raun 1988). In what follows, I concentrate on Erza which is spoken by about 600, 000 people. But all the generalizations which the analysis is based on also hold for the Moksha dialect.

In Mordvin, verbal arguments are encoded by case (dependent-marking)

and by verbal suffixes (head-marking). Only the latter marking strategy is the subject of this paper. Like other Uralic languages (the Samoyedic and Ugric subbranch), Mordvin distinguishes two conjugation classes: the indefinite and the definite conjugation. In the former, the verb agrees solely with its subject (viz. the external argument) in person and number; in the latter it can also agree with its object (viz. the internal argument) in person and number.¹ All intransitive verbs fall of course in the indefinite conjugation, but transitive verbs can fall in either of the classes, depending on the following conditions:

(1) *Conjugation classes in Mordvin:*

- a. definite: transitive verbs in perfective aspect with a definite object
- b. indefinite: all other transitive verbs (and intransitives)

Mordvin verbal suffixes distinguish two numbers (singular, plural), three persons (1st, 2nd, 3rd), two tenses (past, non-past), and seven moods. Mood is of no relevance in what follows because it is expressed independently from the argument-encoding suffixes. The verbal template is shown in (2) (cf. Zaicz 1988: 198).

(2) *Verbal template:*

stem – mood suffix – tense suffix – person/number suffixes

The following tables show the surface forms of the verbal suffixes for the definite and indefinite conjugation in past and non-past tense (cf. Zaicz 1988: 199).² The superscript ‘j’ indicates palatalization of the preceding consonant.

¹Mordvin is the most complex of the Uralic languages in the sense that a verb can agree in number *and person* with its object in the definite conjugation. In all other Uralic languages with a definite conjugation either only number agreement obtains with the object (Khanty, Mansi, Nenets, Enets, Nganasan; cf. Kröhnert (this volume)) or the pure presence vs. absence of a definite object is marked without indicating any ϕ -features of the object (Selkup, Mator, Kamas).

²Compare also the data in Collinder (1957); Abondolo (1982); Raun (1988); Béjar (2003).

(3) *Definite non-past paradigm:*

Obj Subj	1sg	2sg	3sg	1pl	2pl	3pl
1sg	—	-tan	-sa	—	-tadiz ^j	-syn ^j
2sg	-samak	—	-sak	-samiz ^j	—	-sy ^t
3sg	-samam	-tanzat	-sy	-samiz ^j	-tadiz ^j	-syn ^j z ^j e
1pl	—	-tadiz ^j	-syn ^j ek	—	-tadiz ^j	-syn ^j ek
2pl	-samiz ^j	—	-synk	-samiz ^j	—	-synk
3pl	-samiz ^j	-tadiz ^j	-syz ^j	-samiz ^j	-tadiz ^j	-syz

(4) *Definite past paradigm:*

Obj Subj	1sg	2sg	3sg	1pl	2pl	3pl
1sg	—	-itin ^j	-ija	—	-id ^j iz ^j	-in ^j
2sg	-imik	—	-ik	-imiz ^j	—	-i ^j t
3sg	-imim	-in ^j z ^j it ^j	-iz ^j e	-imiz ^j	-id ^j iz ^j	-in ^j z ^j e
1pl	—	-idiz ^j	in ^j ek	—	-id ^j iz ^j	-in ^j ek
2pl	-imiz ^j	—	-in ^j k	-imiz ^j	—	-in ^j k
3pl	-imiz ^j	-idiz ^j	-iz ^j	-imiz ^j	-id ^j iz ^j	-iz _j

(5) *Indefinite non-past paradigm:*

1sg	-an
2sg	-at
3sg	-i
1pl	-tano
2pl	-tado
3pl	-it ^j

(6) *Indefinite past paradigm:*

1sg	-in ^j
2sg	-it ^j
3sg	-s ^j
1pl	-in ^j ek
2pl	-id ^j e
3pl	-s ^j t ^j

Some examples of the forms are given in (7).

- (7) *Mordvin* (Zaicz (1988: 197), Abondolo (1982: 14, 15)):³
- | | | |
|----|---|----------------------------|
| a. | sov-an
enter-1sg
I enter. | <i>indefinite non-past</i> |
| b. | sov-in ^j
enter-1sg
I entered. | <i>indefinite past</i> |
| c. | kund-an
catch-1sg
I catch. | <i>indefinite non-past</i> |
| d. | van-yn ^j
look-1sg
I looked. | <i>indefinite past</i> |
| e. | kunda-tan
catch-1sg>2sg
I catch you. | <i>definite non-past</i> |
| f. | van-yt ^j in ^j
look-1sg>2sg
I saw you. | <i>definite past</i> |
| g. | van-samam
look-3sg>1sg
S/he saw me. | <i>definite past</i> |

The initial consonants /s/ and /t/ in the definite non-past paradigm as well as the initial vowel /i/ in the past paradigms and /s^j/ in 3rd person past indefinite express tense and can thus be ignored for our discussion of argument encoding which only centers around ϕ -features of the arguments. The same holds for the vowel /a/ which immediately follows the tense marker in the non-past paradigm and the non-initial vowel /i/ in the past forms (these sounds are set in bold). Abondolo (1982) analyses these vowels as epenthetic elements which are inserted in order to avoid consonant clusters. The form of the vowel can be predicted: /a/ is inserted in the non-past and /i/ in the past. There are other analyses which take the vowel to form a morpheme together with (at least) the following consonant (Zaicz 1988; Raun 1988; Béjar 2003). Whatever the correct analysis is, the vowel does not contribute to the encoding of ϕ -features of the

³In the examples I only segment stem and suffix string. It is shown below that the suffixes can be further subanalysed into a sequence of suffixes.

arguments because it is invariable within a paradigm. Hence, I leave out this vowel in what follows.⁴

Before I continue, let me introduce some phonological processes that apply to underlying forms and which are important in the verbal inflection (cf. Abondolo 1982; Zaicz 1988; Raun 1988; Béjar 2003). The vowel [i] alternates with [y] which is a backed variant of /i/ that occurs after /s/. The plural morpheme *iz* is reduced to [i] when followed by phonologically non-zero morpheme. Furthermore, the vowel [i] / [y] triggers palatalization of the following coronal phonemes in the suffix string. Hiatus is not tolerated in Mordvin. The sequence /i-i/ is reduced to a single /i/ (e.g. when the past morpheme *i* is followed by the plural morpheme *iz*) and the sequence /i-a/ is realized with a glide as [ja]. Besides, the mid vowels [e] and [o] are allophones in non-word-initial syllables. These vowels assimilate in [±back] to the preceding vowel (vowel harmony, cf. also the alternation *nze ~ nzo, de ~ do*). Finally, voiceless stops are voiced if they appear before a vowel and after a voiced sound. Hence, the 2nd person object morpheme which shows up as the voiced dental stop /d/ is underlyingly a voiceless stop /t/.

3. Subanalysis of Verbal Suffixes

The verbal suffixes in Mordvin are often taken to be portmanteau morphemes in the sense that they are unsegmentable units which express features of subject and object simultaneously (Collinder 1957; Raun 1988; Aranovich 2007; Nevins 2010). Abondolo (1982) and Béjar (2003) whose bases her analysis on Abondolo's work are the only ones who propose a segmentation for (at least) some of the suffixes into separate person and number markers. He motivates this subanalysis with the observation that the postulated submarkers occur as discrete markers in the nominal paradigms of the language. For example, the marker /m/ expresses 1st person because it characterizes 1st person pronouns and it expresses a 1st person possessor of a noun. /m/ can also be found in the verbal paradigm: Whenever the object is 1st person in the definite non-past

⁴The vowel /i/ will nevertheless still show up in the definite past paradigm as part of the marker /iz/ in the combinations plural subject-3sg object and all forms with a 3pl object except 1st sg subject-3pl object. The reason is that in these cases the underlying sequence is /i/-/iz/(-/x/), the former being the past marker and the latter the plural marker (see the discussion below). A sequence of two /i/ is reduced to a single /i/ in Mordvin and hence there is only a single /i/ at the surface.

paradigm in (3), /m/ shows up, so it is probably the same marker expressing first person, in this case the person of the object. Such an approach can easily be pursued in the Uralic languages because exponents for nominal and verbal inflection are generally taken from the same set of markers. Verbal person markers, for example, are attached to nouns in order to express person features of the noun's possessor (= possessive declension in the Uralic languages). I follow the intuition of Abondolo (1982) and segment all verbal suffixes when the submarkers can also be found in other paradigms of Mordvin with the same meaning. Thus, I adhere to the *Syncretism Principle* (Müller 2004).

(8) *Syncretism Principle:*

Identity of form implies identity of function, in a domain Σ , and unless there is convincing evidence to the contrary.

In what follows, I go through other paradigms of Mordvin which can be found in Collinder (1957); Zaicz (1988) and Raun (1988) and I isolate the markers which are also found in the verbal paradigms.

The inventory of personal pronouns shows that /m/ expresses 1st person, /t/ 2nd person, /s/ 3rd person, and /i/ means plural (or more precisely /iz/ which is reduced to /i/). Note that we find the same markers in the definite paradigms when the object is 1st and 2nd person respectively.

(9) *Personal pronouns in Erza (Zaicz (1988: 198), Raun (1988: 104)):*

1sg	mon
2sg	ton
3sg	son
1pl	min'
2pl	tin'
3pl	sin'

The suffix indicating 2nd person singular in the imperative is /k/; /de/ means 2nd person plural, cf. (10). Both suffixes also show up in the verbal paradigm when the subject is 2nd person (/k/ in the definite conjugation with a 2nd person subject, /de - do/ in the indefinite conjugation; the alternation is due to progressive vowel harmony in [\pm back]). The plural of nouns is formed by -/t/ and can be found on nouns in the nominative (compare also the Proto-Uralic pluralizer /t/ (Zaicz 1988: 191)).

- (10) *Imperative (Zaicz 1988: 201):*
- a. vanok = look! (sg)
 - b. vanodo = Look! (Pl)
- (11) *Plural of nouns (Zaicz 1988: 191):*
- a. Nom sg.: kudo 'house'
 - b. Nom pl.: kudo-t 'houses'

In the possessive declension in which a noun is inflected for person and number of its possessor we encounter again /m/, /t/ and /k/. Their distribution supports what we have already seen so far with pronouns and imperative forms: /m/ marks 1st person and /t/ 2nd person. In addition, there is a second 1st person marker in the possessive paradigm that shows up with more than one possessor: /n/. Again this marker is found in the verbal paradigm when the subject is 1st person. What distinguishes the two 1st person markers /m, n/ and the 2nd person markers /t, k/? By looking at the definite paradigm, one can see that /m/ and /t/ express the person of the object whereas /n/ and /k/ realize subject person.

Furthermore, the markers /ze ~ zo/ and /nze/ express 3rd person as they do in some of the forms in the definite verbal paradigm with a 3rd person subject. The markers /nek ~ nok/ (1st person plural) and /ŋk/ (2nd person plural) are exactly those that we find in the definite verbal paradigms with the correspondingly specified subject acting on a 3rd person object. Finally, /st/ for a 3rd person plural possessor seems to be composed of the 3rd person marker /s/ we have already encountered in pronouns and the nominal plural marker /t/ which also occurs with 3rd person plural subjects of the indefinite conjugation.

- (12) *Possessive declension of the noun kudo 'house' (Raun 1988: 102):*
- | | |
|------------------------------------|----------|
| 1st sg, one possession | kudo-m |
| 2nd sg, one poss. | kudo-t |
| 3rd sg, one poss | kudo-zo |
| | |
| 1st sg, more than one poss. | kudo-n |
| 2nd sg, more than one poss. | kudo-t |
| 3rd sg, more than one poss. | kudo-nzo |
| | |
| 1st pl, one or more than one poss. | kudo-nok |
| 2nd pl, one or more than one poss. | kudi-nk |
| 3rd pl, one or more than one poss. | kudi-st |

The declension of personal pronouns is another place where person markers can be found. These pronouns decline like nouns in the non-oblique cases. The oblique case forms, however, are special: The pronominal stem which unambiguously expresses person is followed by the regular nominal case suffix and this suffix in turn precedes a suffix which redundantly expresses person for a second time. We can see in the following two examples that these final suffixes are drawn from the set of suffixes that are also used in verbal inflection and the possessive declension. In the examples in (13) /n/ is attached to 1st person pronouns and /nze/ to 3rd person pronouns, confirming their grammatical meanings exemplified above: 1st and 3rd person.

(13) *Inessive case of 1st/2nd sg personal pronouns (Zaicz 1988: 197):*

- a. mon^j-se-n;
 pron.1sg-iness-1sg
 in me
- b. son^j-se-nze
 pron.3sg-iness-3sg
 in him/her

(14) *Declension of a 3rd person pronoun:*

pronoun stem > case suffix > person suffix (Zaicz 1988: 197):

Nom	son
Gen/Acc	son-ze
Dat/All	son-e-nze
Iness	son-se-nze
Ela	son-ste-nze
Ill	son-ze-nze
Prol	son-ga-nzo
Abl	son-d ^j e-nze
Trans	son-kse-nze
Abess	son-t ^j eme-nze
Cfv	son-ška-nzo

A summary of the markers found in the nominal paradigms and the imperative forms is shown in (15).

(15) *Summary of submarkers:*

/m/	[1]	/t/	[2]
/s/	[3]	/i(z)/	[pl]
/k/	[2]	/de/	[2pl]
/t ₂ /	[pl]	/nze/	[3]
/nek/	[1pl]	/nk/	[2pl]

The crucial point is that markers which we find in nominal paradigms also show up as part of the alleged complex verbal inflection markers in (3) - (6) and that their grammatical meaning in the nominal paradigms is compatible with the morphosyntactic context in which they appear in the verbal paradigms. Hence, it is legitimate to postulate that the markers in the nominal and verbal paradigms are instances of the same abstract morpheme. I therefore subanalyse the verbal suffixes into a sequence of morphemes also found in noun inflection. In this way, we can capture transparadigmatic syncretisms across word classes. The result of segmentation is shown in (16) - (19) (remember that the tense marker and the epenthetic vowel are ignored; the paradigms show underlying forms, abstracting away from phonological processes).

(16) *Subanalysed definite non-past paradigm:*

Obj Subj	1sg	2sg	3sg	1pl	2pl	3pl
1sg	—	-t -n 2 1	-a ?	—	-t -iz 2 pl	-iz -n pl 1
2sg	-m -k 1 2	—	-k 2	-m -iz 1 pl	—	-iz -t pl 2
3sg	-m -m 1 1	-nze -t 3 2	-i 3	-m -iz 1 pl	-t -iz 2 pl	-iz -nze pl 3
1pl	—	-t -iz 2 pl	-iz -nek pl 1	—	-t -iz 2 pl	-iz -nek pl 1
2pl	-m -iz 1 pl	—	-iz -ŋk pl 2	-m -iz 1 pl	—	-iz ŋk pl 2
3pl	-m -iz 1 pl	-t -iz 2 pl	-iz -∅ pl 3	-m -iz 1 pl	-t -iz 2 pl	-iz -∅ pl 3

(17) *Subanalysed definite past paradigm:*

Obj Subj	1sg	2sg	3sg	1pl	2pl	3pl
1sg	—	-t -n 2 1	-a ?	—	-t -iz 2 pl	-iz -n pl 1
2sg	-m -k 1 2	—	-k 2	-m -iz 1 pl	—	-iz -t pl 2
3sg	-m -m 1 1	-nze -t 3 2	-ze 3	-m -iz 1 pl	-t -iz 2 pl	-iz -nze pl 3
1pl	—	-t -iz 2 pl	-iz -nek pl 1	—	-t -iz 2 pl	-iz -nek pl 1
2pl	-m -iz 1 pl	—	-iz -ŋk pl 2	-m -iz 1 pl	—	-iz -ŋk pl 2
3pl	-m -iz 1 pl	-t -iz 2 pl	-iz -∅ pl 3	-m -iz 1 pl	-t -iz 2 pl	-iz -∅ pl 3

(18) *Subanalysed indefinite non-past paradigm:*

1sg	-n 1
2sg	-t 2
3sg	-i 3
1pl	-t -no ? 1
2pl	-t -do ? 2
3pl	-i -t 3 pl

(19) *Subanalysed indefinite past paradigm:*

1sg	-n 1
2sg	-t 2
3sg	-∅ 3
1pl	-nek 1
2pl	-de 2
3pl	-t -∅ pl 3

As a result of subanalysis, each morpheme encodes either only subject or object features. Thus, there are no portmanteau morphemes in the sense that a single morpheme expresses features of both subject and object simultaneously. The

subanalysed paradigms are the basis for the following discussion of argument encoding in Mordvin.

Before I close this section, some comments on the glosses are in order at this point. The morpheme /a/ glossed as “?” in the definite paradigms is sometimes taken to be a 1st person morpheme (cf. Béjar 2003). Researchers claim that the expected 1st person subject marker /n/ has followed this /a/ in previous stages of the language (Zaicz 1988; Raun 1988) and hence /a/ has also been an epenthetic vowel. Its presence in modern Mordvin suggests that there is indeed a morpheme following /a/ which is however not spelled out. I adopt this view in what follows and treat /a/ as epenthetic, but it can also be taken to be an allomorph of the 1st person subject marker /n/. The marker /t/ in the indefinite non-past paradigm is also glossed as “?”. Zaicz (1988) and Raun (1988) are not sure about its origin or meaning. I will come back to this marker in my analysis in section 7 in which I propose that this /t/ means “plural”.

4. Generalizations and Challenges

Béjar (2003) also subanalyses the verbal suffixes in Mordvin on the basis of the segmentation proposed by Abondolo (1982) which largely corresponds to my subanalysis.⁵ She observes that there are hierarchy effects in that agreement can obtain with the subject or the object, depending on their grammatical function and feature values. She states the following generalizations on the marker distribution (ignoring the shaded cells for the moment):

(20) *Béjar (2003: 168) on marker distribution in Mordvin:*

a. PERSON AGREEMENT:

The verb agrees with the object if it is local person. Subject person is cross-referenced only if the object is 3rd person.

b. NUMBER AGREEMENT:

The verb agrees with the subject in number if it is plural. Plural of the object is cross-referenced only if the subject is singular.⁶

⁵She does, however, not motivate this segmentation at all and Abondolo (1982) did not provide a segmentation for all forms.

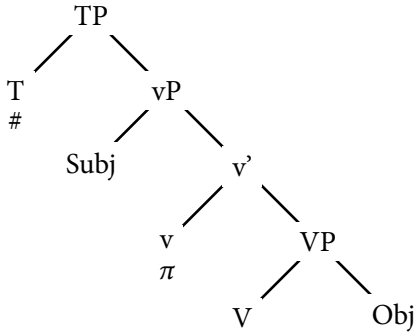
⁶The generalization on the distribution of number also holds in the possessive declension in (12): If the structurally higher argument, viz. the possessor, is plural, number of the lower argument, the possessum, is not indicated; the forms are ambiguous between one or more than

Thus, there is a preference for agreement with the object for person and a preference for agreement with the subject for number. The other argument can only be agreed with if the preferred argument does not provide a marked value (local person, plural). In general, the unmarked values of a feature (3rd person, singular) are not visible for the agreement system when agreement obtains with the preferred argument. But whereas the unmarked person value (3rd person) becomes a possible target for agreement when it obtains with the non-preferred coargument (the subject), the unmarked number value singular remains invisible even for agreement with the non-preferred object. Béjar (2003) argues that singular is complete absence of any number features in Mordvin and is thus always invisible for agreement, in contrast to 3rd person.⁷

The different behaviour of person and number agreement suggests that person and number probe are separated. With this assumption, Béjar (2003) implements the preferences of the probes by their different structural position: The person probe π is located on v , the number probe $\#$ is located on T , cf. (21). Each probe agrees with the closest accessible goal in its c -command domain. The closest goal for π is the object and for $\#$ it is the subject. A probe can expand its search domain to the coargument if the closest goal does not provide a marked value. Hence, $\#$ can then search for number on the object and π can search for person on the subject (in the latter case there is no c -command between v and the subject; this issue is addressed below).

one possessum. If, however, the possessor is singular, number of the possessum is indicated by the suffixes.

⁷Her arguments are that i) there is never an overt exponent for singular (but for 3rd person, cf. the paradigms above), and ii) that a singular argument is not an intervener for agreement between a higher number probe and a structurally lower argument (in contrast to other languages where an argument with an unmarked value does act as an intervener in such a scenario and thus must have some features). This means that the subject is not an intervener for Agree between the number probe on T and the object.

(21) *Probes in the structure:*

The analysis developed by Béjar (2003) is simple and elegant. However, when we look at the shaded cells in the definite paradigms above, we see that there are two person exponents in a suffix string, a fact which does not hold in any other cell. Béjar’s analysis cannot explain why double person agreement obtains. Such a pattern should even be impossible if there is only a single person probe in the structure. According to her analysis we expect there to be (maximally) a single person and a single number marker in a suffix string. Béjar (2003: 171) comments on this fact as follows: “Finally, what can be said about the irregularities in the shaded forms of (24) [\approx (16) and (17), D.G.]? Here, I would argue, the syntax of agreement has nothing to say about the actual form of these irregularities [...]”.

I take Béjar’s analysis of verbal agreement in Mordvin to be generally valid.⁸ In what follows I will show that there is a way to reconcile Béjar’s approach with the alleged irregularities. Thus, the marker distribution in the whole paradigm, including the double person marking pattern, can indeed be derived and is even expected to occur with a straightforward and independently motivated extension of Béjar’s analysis.

To this end, I first propose another straightforward generalization about the definite paradigms in Mordvin which will lead to a new analysis:

- (22) *Two-marker-generalization on the Mordvinian definite conjugation:*
 There are always two markers in a suffix string.
 (apart from the combination sg subject \rightarrow 3rd sg object which I will come back to)

⁸Different approaches which try to derive the syncretism patterns in Mordvin and which are not based on subanalysed suffixes can be found in Aranovich (2007) and Nevins (2010).

When looking at the definite paradigms in (16) and (17), one can see that these two markers are usually a person and a number marker. In the shaded cells, however, there are two person markers instead of a person and a number marker. But note that the generalization also holds for these cells. The absence of a number suffix is expected, because both arguments are singular in the context described by the shaded cells and the number probe # cannot “see” the value singular in Mordvin (recall that singular was argued to be absence of number features). What is exceptional is the occurrence of a second person marker because there is only a single person probe and once it is checked and valued it is inactive and cannot be valued again.

The generalization in (22) allows for an explanation of double person marking. Note that it is not a stipulation or a language-specific constraint; it is just an observation which follows from the fact that there are two probes in the structure, π and #, which have to be checked and valued if the derivation is to converge (cf. Chomsky 2000). The number probe searching for plural cannot be checked if both arguments are singular and could thus remain unchecked. Béjar (2003: 76f.) assumes that the number probe is checked by default agreement (= deletion of the probe without valuation) in this context in order to avoid the crash of the derivation, but this cannot explain the occurrence of a second person marker. Instead, I propose that the number probe is not deactivated by default agreement, but - descriptively spoken - it turns into a person probe and is valued by person features rather than remaining unvalued. I call such a probe a *Chameleon probe* because it accommodates itself to the syntactic context, being able to search for number or person. This account also guarantees that the derivation does not crash because of an unvalued “number” probe, without the need for a further stipulated concept like default agreement.

One could also propose another solution which is closer to Béjar’s analysis: Assume that i) the generalization in (22) is a filter which is active in Mordvin in the context of a transitive verb in perfective aspect with a definite object, and that ii) the number probe is indeed deactivated by default agreement. The valuation of the person probe leads to a single person exponent. However, the output would violate the *two-marker-generalization*. Hence, a repair strategy could apply in order to fulfill the generalization. One way to implement this is to say that the present person marker is simply copied in the phonological component. However, this cannot be true, because the two person markers in the shaded cells are not (always) identical. A different repair mechanism is the insertion of a second person probe. Such an approach is less preferable for two reasons:

First, the insertion of an additional probe violates *Inclusiveness* (Chomsky 1995) and secondly, it would have to stipulate the *two-marker-generalization* as a language-specific requirement on definite verbal paradigms which is a high ranked constraint and can thus trigger repair strategies. Hence, the approach would miss the fact that the generalization in (22) follows automatically from the number of probes in the structure. Because the alternative analyses have serious short-comings, I pursue the approach with a *Chameleon probe*.

To sum up, I suggest that double person marking obtains because the regular person probe and also the former number probe is valued by person in the shaded cells. The “false” valuation of # is restricted to contexts in which both arguments are singular, because if one of the arguments is plural, a plural exponent shows up (cf. (16) and (17)) and hence, # is then saturated by a number value.

5. Analysis

A number of questions arise for the *Chameleon probe* analysis. First of all, how can the “metamorphosis” of the probe be implemented and how can it be restricted to a context in which both arguments are singular? Besides, how can we express the fact that the probe under consideration cannot simply choose whether it agrees in person or number, but that it preferably searches for number values? Why is the person probe not a *Chameleon probe* which can exceptionally be valued by number? In this section I address these questions and I show that the change in the nature of the number probe and the required restrictions on this process do not have to be stipulated. They follow directly from the system that Béjar (2003) develops with only a minor change which is independently motivated. This means that no new mechanisms are applied which Béjar did not also require in her analysis of hierarchy effects - *Chameleon probes* are a direct consequence of her approach.

5.1. Béjar’s (2003) System

In order to see how *Chameleon probes* can come into existence, I have to outline the approach of Béjar (2003) in more detail. I repeat the assumptions which have already been introduced above:

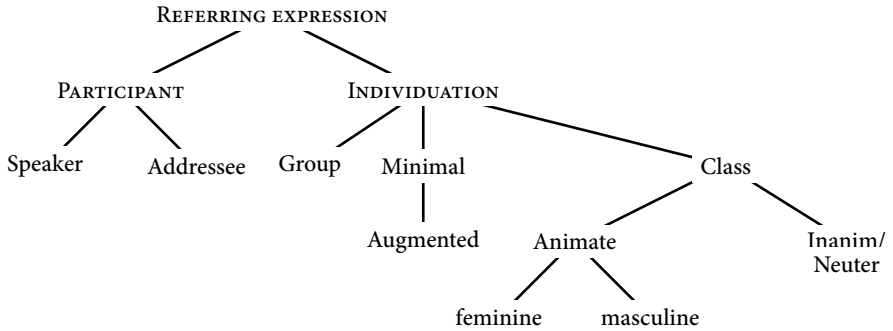
Derivation of hierarchy effects in Mordvin Béjar (2003):

- Person probe π and number probe $\#$ are separated.
- π is located on v , $\#$ is located on T .
- Probes have to be checked by a goal with matching features in their c -command domain (by Agree).
- Agree targets the closest accessible goal.
- Dynamic search domains: If the closest goal cannot value a probe, the probe expands its search domain to the next goal.

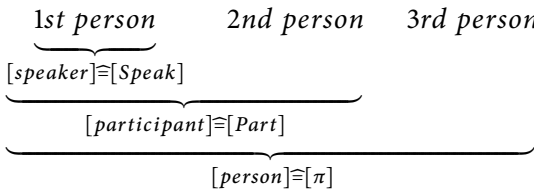
The last point requires some qualifications. Agree requires c -command: The number probe on T c -commands both subject and object, but the person probe on v does not c -command the subject. Béjar (2003) discusses several options to achieve c -command, what follows is compatible with any of these. For concreteness, I adopt the proposal that the probe is raised to T from where it c -commands the subject. This is a repair strategy which applies if the person probe does not find a suitable goal in its c -command domain. Rather than remaining unvalued which would cause a crash at the end of the derivation, it raises to T and starts probing again in its extended search domain in which a new potential goal (the subject) is present. But crucially, the number probe on T does not change its structural position in case it cannot be valued by the closest goal; thus $\#$ need not descend the tree if it wants to establish an Agree relation with the object, it just ignores the closest goal, because $\#$ c -commands both the subject and the object from the beginning and raising for repair is not necessary. Hence, $\#$ always stays on T , π starts out on v but can end up on T .

If valuation of a probe obtains with the closest goal, Béjar calls this *first-cycle Agree*; if it can only apply after expansion of the search domain to the coargument, she speaks of *second-cycle Agree*, a terminology which I adopt in what follows.

The most important point is the implementation of the fact that only marked values of a feature on a goal can check and value a probe on the first cycle. In order to achieve this, Béjar decomposes person and number values into bundles of more abstract privative features, based on the feature geometry developed by Harley and Ritter (2002) in (23).

(23) *Harley and Ritter (2002):*

Their claim is that there are implicational relations between nodes: A node on a level L_i implies the presence of the node on a higher level L_{i+n} . Hence, *Speaker* implies the presence of Participant etc. In a language like Mordvin with three person and two number values only the nodes REFERRING EXPRESSION, PARTICIPANT, *Speaker*, INDIVIDUATION, and *Group* are active. PARTICIPANT encodes speech-act participants (local persons), SPEAKER encodes the speaker; INDIVIDUATION encodes number in general, *Group* encodes plurality. Note that Béjar introduces an additional feature π in the geometry above PARTICIPANT. It means “person” in general, hence, if only π is active this encodes 3rd person. The feature bundles for the relevant values then look as follows:⁹

(24) *Bejar’s feature decomposition based on Harley and Ritter (2002):*

(25) 1st: $\begin{bmatrix} [\pi] \\ [Part] \\ [Speak] \end{bmatrix}$ 2nd: $\begin{bmatrix} [\pi] \\ [Part] \end{bmatrix}$ 3rd: $[[\pi]]$

⁹The feature bundles are also structured units and have to be read as abbreviations for a feature-geometric tree: A higher feature dominates the features below it.

$$(26) \quad \text{Sg: } [] \quad \text{Pl: } \left[\begin{array}{l} [\text{Indiv}] \\ [\text{Group}] \end{array} \right]$$

First person, for example, is encoded as $[\pi[\text{Part}[\text{Speak}]]]$ because it represents the speaker and the specification $[\text{Speak}]$ requires the presence of all dominating nodes (except $[\text{R}]$ which Béjar crucially does not take to be a part of the feature bundles - a fact I will come back to below); second person consists only of $[\pi[\text{Part}]]$, because it is a local person but not the speaker; etc. Remember that the value *singular* is complete absence of number features in Mordvin.

The two probes are complex, too, and can be represented in the same way, see (27). Note that the probes are specified for the value they require their closest goal to possess, namely local person (encoded by $[\text{Part}]$) and plural (encoded by $[\text{Group}]$). The absence of $[\text{Speak}]$ on the person probe allows to refer to the natural class of first and second person.

(27) *Person and number probe in Mordvin (Béjar 2003: 168):*

$$\pi: \left[\begin{array}{l} [\pi] \\ [\text{Part}] \end{array} \right] \quad \#: \left[\begin{array}{l} [\text{Indiv}] \\ [\text{Group}] \end{array} \right]$$

We can now return to the question why only marked values can satisfy a probe. Béjar suggests that a probe can only be checked by the features of a goal if the goal has a superset of features of the probe.¹⁰ The consequences of this requirement for first-cycle Agree are as follows:

(28) *First-cycle Agree:*

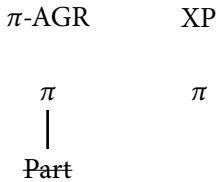
- a. Person agreement obtains only with a local person object.
- b. Number agreement obtains only with a plural subject.

¹⁰The set-theoretic concept of “subset” is applied if the elements of a set A are contained inside a set B; this means that A and B can (but do not have to) be identical. Note that Béjar’s analysis combines the checking and valuation approach to Agree. A probe is checked and thus becomes inactive (not able to search further) if it finds a goal which has a superset of features of the probe. Furthermore, the goal copies its features onto the head on which the probe is located (valuation). For example, if the person probe $[\pi[\text{Part}]]$ is valued by a 1st person object $[\pi[\text{Part}[\text{Speak}]]]$, the feature $[\text{Speak}]$ which the probe does not include is copied to v. The reason for this copying is that 1st and 2nd person are morphologically distinguished in Mordvin. If $[\text{Speak}]$ were not copied onto v, 1st and 2nd person objects could not be distinguished in a postsyntactic realizational morphological component.

First-cycle Agree with a 3rd person goal (being [π]) or a singular goal (being absence of features) is impossible because these feature bundles are not a superset of the probe's features. In this case, second-cycle Agree is enabled and the probes expand their search domain to the coargument: # looks for number on the object; π is raised to T and searches for person on the subject.

Now a complication arises. One can see in the paradigms that there can indeed be person agreement with a 3rd person subject (second-cycle Agree), although the feature matrix encoding 3rd person is not a superset of the probe's feature specification. Béjar (2003) observes this effect in a number of other languages. For this reason, she proposes a further repair mechanism, called *Probe impoverishment*: It applies if valuation of a probe fails at the first cycle and it deletes the most specific (viz. the bottommost) feature of the feature matrix in order to facilitate Agree at the second cycle.¹¹ In the case of the person probe this means that [Part] is deleted before the probe searches for a second time, cf. (29). Probe impoverishment applies simultaneously with raising to T in the case of π .

(29) *Probe impoverishment* (cf. Béjar 2003: 78):



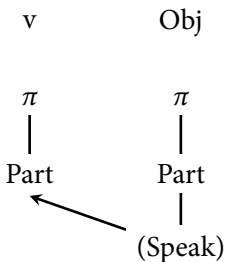
As a consequence, a 3rd person goal can check and value the person probe at the second cycle, having an identical feature set as the probe. Note that this does not help for the number probe: Even though [Group] is deleted, the probe consisting only of [Indiv] can still only be valued by a plural object at the second cycle, because singular is complete absence of features and therefore still not a subset of the probe. In order to avoid the crash of the derivation, Béjar proposes that default agreement obtains if the number probe can also not be valued at the second cycle because both arguments are singular. This system which is illustrated in the trees below derives the generalizations in (20) and proceeds in the following steps:

¹¹In fact, she proposes to delete all features of a probe except for the topmost feature, and not only the bottommost. But if a language has a probe that consist of two features like in Mordvin, the result is the same: Only the topmost feature of the probe feature bundle remains.

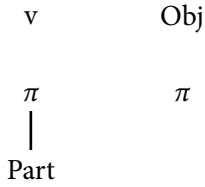
- The head containing the probe is merged.
- Immediately afterwards, the probe starts searching for a matching goal in its c-command domain.
- If it finds such a goal, it is checked and valued and becomes inactive (first cycle Agree). It does not probe further.
- If the probe does not find a matching goal, two repair mechanisms apply: i) The probe expands its search domain (which means raising to T in the case of π but not in the case of #) and ii) the lowest feature in the structures feature bundle is deleted.
- Afterwards, the probe starts searching for a matching goal in its c-command domain again.
- If the probe finds a suitable goal, it is checked and valued and becomes inactive (second cycle Agree). The person probe will be valued at the latest at this step of the derivation, the number probe does not have to be if the closest goal is singular.
- If the number probe does not find a matching goal at the second cycle, too, the two repair mechanisms apply again: i) The probe expands its search domain (which means restarting the search on the closest goal) and ii) the bottommost feature in the probe's structured feature bundle is deleted.
- The probe starts searching for a goal again and it will be valued at this stage of the derivation, because it is highly underspecified and will definitely find a matching person feature bundle on the closest goal. The probe is then checked and valued and becomes inactive.

ILLUSTRATION OF BEJAR'S SYSTEM (valuation indicated by an arrow):

(30) *Person agreement, 1st cycle valuation (local person object):*

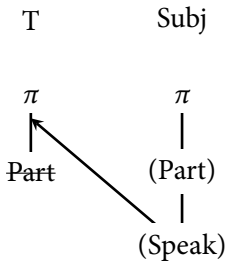


(31) a. *Person agreement, no 1st cycle valuation (3rd person object):*

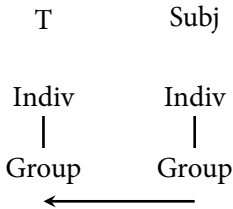


b. π raises to T, [Part] is deleted

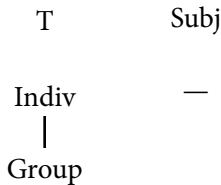
c. *Person agreement, 2nd cycle valuation (1st/2nd/3rd person subject):*



(32) *Number agreement, 1st cycle valuation (plural subject):*

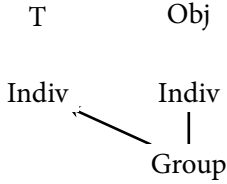


(33) a. *Number agreement, no 1st cycle valuation (singular subject):*



b. [Group] is deleted

- c. *Number agreement, 2nd cycle valuation (plural object):*



- (34) a. *Number agreement, no 2nd cycle valuation (singular object):*



- b. Bejar: default agreement; no number exponent shows up;
occurrence of a second person marker unexplained

5.2. Extended System: How Chameleon Probes Arise

In section 4 I proposed that the number probe is not saturated by default agreement if both arguments are singular, but that it is exceptionally valued by person in order to get checked, with the result of double person marking. But how can the number probe “see” person values? Remember Béjar’s condition that valuation is only possible if the goal has a superset of features of the probe. It is thus impossible in her decomposition of person and number features that a number probe is valued by person. My suggestion is that a scenario in which a person feature bundle is a superset of a “number” probe indeed arises if we follow the feature geometry by Harley and Ritter (2002) more closely. Note that Béjar (2003) by and large adopts their geometry, but she does not include the root node R(REFERRING EXPRESSION) in her feature bundles (cf. (23)). This root node is however active in Harley and Ritter’s system: Some combinations of features are represented by the presence of the node [R] alone (e.g. 3rd person pronouns are represented by [R] in a language which has no number distinctions for pronouns, e.g. Pirahã, cf. Harley and Ritter (2002: 501)). I thus include this node in the feature matrices such that the feature bundles look as in (35) and (36) with [R] on top. The probes in Mordvin which look for local person and plural, respectively, have the form in (37).

$$(35) \quad \text{1st: } \begin{bmatrix} [R] \\ [\pi] \\ [Part] \\ [Speak] \end{bmatrix} \quad \text{2nd: } \begin{bmatrix} [R] \\ [\pi] \\ [Part] \end{bmatrix} \quad \text{3rd: } \begin{bmatrix} [R] \\ [\pi] \end{bmatrix}$$

$$(36) \quad \text{Sg: } [] \quad \text{Pl: } \begin{bmatrix} [R] \\ [Indiv] \\ [Group] \end{bmatrix}$$

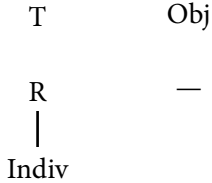
$$(37) \quad \pi: \begin{bmatrix} [R] \\ [\pi] \\ [Part] \end{bmatrix} \quad \#: \begin{bmatrix} [R] \\ [Indiv] \\ [Group] \end{bmatrix}$$

What is crucial for the following derivations is that person and number now form a natural class with respect to the node [R]. The remaining assumptions of Béjar (2003) are adopted without changes. For the valuation of the person probe the additional feature [R] has no consequences. The derivations are exactly as in (30) and (31), except for the presence of [R] on top of the goals and the probe. The person probe is checked at the latest at the second cycle (by the subject) because any person (1st/2nd/3rd) feature bundle is a superset of or identical to the probe's features at the second cycle.

For the number probe, nothing changes if one of the arguments is plural. The probe is then checked and valued by this argument. The derivations are as in (32) and (33-c) except that the probe and the goal include the node [R] now. If the subject is plural, valuation obtains at the first cycle, if it is singular, the probe is impoverished to [R [Indiv]] and it looks for plural on the object. Valuation obtains if the object is plural. But if the object is singular, too, (as is the subject) and # can still not be valued, we expect it to be impoverished for a second time in Béjar's system. If we apply probe impoverishment, the feature [Indiv] is deleted. As a result, the number probe consists only of the node [R], [Indiv] and [Group] having been deleted. Crucially, the probe is still active at this point, because the additional node [R] is present and the probe can start a new search. As the privative feature [R] is also a part of every person feature bundle, the probe can be checked and valued with any person value because the person feature bundle on a goal consists at least of [R [\pi]] and therefore fulfills the superset condition on valuation. Following the terminology of Béjar, this third attempt of the probe to find a goal can be called "third cycle Agree". It is illustrated in (38). Valuation of a number probe by person is possible because

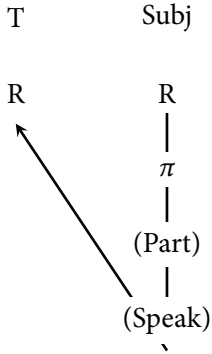
person and number form a natural class with respect to the privative feature [R] which has been suggested for independent reasons in Harley and Ritter (2002) anyway.

(38) a. *Number agreement, no 2nd cycle valuation (singular object):*



b. *second probe impoverishment: [Indiv] is deleted*

c. *Person agreement, 3rd cycle valuation of former “#”:*



The heavily impoverished probe # loses its feature [Indiv] which characterizes it as a number probe and it can then be valued by person. Note that there is no stipulated transformation of the number probe into a person probe, the effect results from the impoverishment which Béjar (2003) needed anyway to allow for the valuation of a person probe by a 3rd person goal at the second cycle (also in a number of other languages than Mordvin). Double person marking in the shaded cells arises because the regular person probe is valued by person and the highly impoverished “number” probe is valued by person features, too, because it does not find a plural argument.

The general goal of this analysis is not to propose a new type of probe, but to show that the probe impoverishment approach developed by Béjar (2003), coupled with the proposal that person and number form a natural class α Harley and Ritter (2002), predicts the existence of highly underspecified probes

which can in principle be valued by any feature that is part of α . As a result, we expect that an X-probe searching for feature X on a goal can be valued by feature Y (where X and Y are variables over features in α) in certain well-defined contexts, such that an exponent of Y shows up where an exponent of X is expected. The Mordvin data show that such a pattern is indeed attested. Thus, the indirectly predicted existence of *chameleon probes* can be seen as a desired result of Béjar's probe impoverishment approach, which qualifies as a strategy to avoid unchecked probes, which would lead to a crash of the derivation.

Finally, note that the term *chameleon probe* is just a descriptive term which clarifies the consequences of probe impoverishment on the surface: It looks as if the number probe turns into a person probe because we see a person exponent instead of the expected number exponent. But, of course, there is no real "metamorphosis" of the number probe which miraculously becomes a person probe. Rather, the *chameleon probe* is neither a person nor a number probe, it is heavily underspecified as a consequence of repeated impoverishment and can be valued either by person or number.

6. Predictions on Exponence

In this section I discuss several questions with respect to exponence that arise for the present analysis. These concern the properties of third cycle Agree, apparent exceptions to the *two-marker-generalization* in (22), and the choice of a person marker.

6.1. The Domain of Third Cycle Agree

In Béjar's analysis there are at most two agreement cycles: The first obtains with the closest accessible goal and the second with the coargument of the first goal. But what is the search domain for 3rd cycle Agree in Mordvin? Is it identical to the first or the second cycle search domain of the "number" probe? The answer does not fall out in any way from the system developed so far, it could be one or the other way so this turns out to be an empirical question. When we look at the exponents in the shaded cells in (16) and (17), we can see that the additional person marker (the rightmost marker) cross-references the person of the subject.¹² Hence, the *Chameleon probe* searches for a goal

¹²In general, the object marker is closer to the stem than the subject marker in Mordvin.

in the first cycle domain of the former “number” probe and does not search again in its extended second cycle domain. I conclude from this that the probe on T always targets the closest accessible goal as long as there are features by which it can potentially be valued.¹³ The only exception to this observation is the combination 3sg subject – 1sg object in the definite paradigms. Here the person of the object is expressed twice, although my analysis predicts that the outermost exponent should cross-reference the person of the subject which is the closest matching goal for the *Chameleon probe*. I do not have an explanation for why the subject is “invisible” in this context, but the important point for the present account is that there is double person marking, which should be impossible in Béjar’s system, but is expected under the extended system developed here. I will come back to this special form below.

6.2. The Choice of Person Exponents

A second question that arises is which exponent is chosen in which context. As can be seen in the definite paradigms, plural is always realized by the marker /iz/, regardless of whether the value is provided by subject or object. For person features, things are more complicated. We have already seen in section 3 that there are several exponents for 1st and 2nd person. Béjar (2003) proposes that the choice of an exponent for a person value is driven by the cycle at which it has been valued: For example, 1st person is expressed by /m/ when the object valued the person probe (first cycle Agree) and /n/ and /nek/ when the subject valued the person probe (second cycle Agree). The same holds for second person: /t/ expresses 2nd person object (1st cycle Agree), /k/ and /ŋk/ 2nd person subject (2nd cycle Agree); 3rd person morphology /nze/ and /Ø/ respectively, is necessarily second cycle morphology, because there is no valuation by 3rd person on the first cycle due to the condition on valuation.

(39) *Person exponents:*

	1st cycle	2nd cycle
1st	/m/	/n/, /nek/
2nd	/t/	/k/, /ŋk/
3rd	–	/nze/, /Ø/

¹³This fact could be derived in an optimality-theoretic approach by postulating a high-ranked locality constraint, viz. a constraint which demands valuation by the closest goal whenever possible.

One way proposed by Béjar to implement this difference in a post-syntactic realizational morphology (like Distributed Morphology, Halle and Marantz (1993; 1994)) is to use context restrictions. If the person probe is valued by the object (first cycle), it is located on *v*, if it is valued by the subject, it has been raised to *T*. This structural difference, the location of the probe, can be used as a context for vocabulary insertion. The different vocabulary items (VIs) can thus be represented as in (40) (exemplified for 1st person morphology):

- (40) *First person morphology*:¹⁴
- | | | |
|----|-----------------|-----------------------------|
| a. | /m/ ↔ [1] / [v] | <i>1st cycle morphology</i> |
| b. | /n/ ↔ [1] / [T] | <i>2nd cycle morphology</i> |

A list of all vocabulary items is given in (41). I will discuss the difference between the two second cycle morphemes of the person VIs (with the context *T*) below. According to the *Subset Principle* and *Specificity* the most specific matching vocabulary item is inserted into a terminal node.

- (41) *Vocabulary items*
- a. /m/ ↔ [1] / [v]
 - b. /t/ ↔ [2] / [v]
 - c. /n/ ↔ [1] / [T]
 - d. /k/ ↔ [2] / [T]
 - e. /nze/ ↔ [3] / [T]
 - f. /nek/ ↔ [1] / [T, pl]
 - g. /ŋk/ ↔ [2] / [T, pl]
 - h. /∅/ ↔ [3] / [T, pl]
 - i. /iz/ ↔ [pl]

The prediction is that the “number probe” which is valued by person when both arguments are singular is realized by second cycle person morphology. The reason is that the number probe # starts out on *T* and is not moved to another head, even if first cycle Agree does not obtain. The reason for this is that # *c*-commands both arguments of the verb and movement is not necessary

¹⁴For reasons of simplification and clarity I use the numbers 1, 2, and 3 in the VI entries as abbreviations for the complex features bundles [R [π [Part [Speak]]]], [R [π [Part]]], and [R [π]] which are present on syntactic heads and are spelled out by the VIs.

for second cycle Agree. Now, if # is a *Chameleon probe*, viz. if it is valued by person features, it is indistinguishable for the morphological component from second cycle valuation of the regular person probe which is raised from *v* to *T* in order to be able to be valued by the subject. This prediction is indeed borne out: The exponents for person of the subject in the shaded cells are /n/, /k/, /nze/, all of which are second cycle morphemes, cf. (41). The only exception is the combination 3sg subject – 1sg object, where the first cycle morpheme /m/ shows up – an exceptional form which has already been discussed above. Recall that the *chameleon probe* references the person value of the object in this form instead of the subject’s person feature, as expected. I cannot explain why the subject is ignored in this derivation, but nevertheless we would still expect the first cycle morpheme for the object’s person, viz. /n/ instead of /m/. However, in the Moksha dialect of Mordvin, we find indeed the expected form /s-m-n/ (surface form -saman) (Raun 1988: 106).

Some comments on the VIs in (41) are in order. The difference between the second cycle person morpheme in the upper block (/n/, /k/, /nze/) and the second cycle person morphemes in the lower block (/nek/, /ŋk/, /Ø/) is that the latter show up when the subject is plural (hence in a context in which a plural subject acts on a 3rd person object, see (17)). In this case, the number value *plural* is present on *T*, hence the context features *T* and *plural* are located on the same head. The other person markers are inserted if the subject is singular. We have to rule out that person is spelled out by the second cycle morphemes /nek/, /ŋk/, /Ø/ if the subject is singular and the object is 3rd person plural. In this context, there is also a person value and plural valued on *T* and hence, these VIs are the most specific ones, but the less specific second cycle person morphemes are realized (/n/, /k/, /nze/). The difference between the combinations singular subject – 3rd plural object and plural subject – 3rd person object is that in the former case plural is valued on the second cycle, but on the first cycle in the latter case. Hence, the insertion of /nek/, /ŋk/, /Ø/ has to be further restricted to the context feature “plural” which is valued on the first cycle, but not on the second. The time of valuation may be distinguished by a diacritic for first cycle number valuation.¹⁵

¹⁵Three forms in the definite paradigm are special and do not show the expected exponent. The first is the combination 2sg subject – 3pl object in past and non-past. Here the person value of the subject is indicated by the first cycle morpheme /t/, although the second cycle morpheme /k/ is expected. The second complication is the combination of a 3sg subject acting on a 3sg object in the non-past definite paradigm. The person exponent /nze/ is expected to occur, but instead

A more general point is that the morphemes with the context restriction [T, pl] are always preceded by the plural marker /iz/. It is standardly assumed that features are deleted by vocabulary insertion, but in my analysis the feature *plural* must be accessible as a context feature for the insertion of the second cycle person morpheme although it has already been realized by /iz/. I thus have to assume that features realized by vocabulary insertion are inaccessible for further insertion but are still visible as context features (cf. the concept of secondary exponence Noyer 1992: 69).¹⁶

6.3. Exceptions to the Two-marker-generalization

The last issue I want to address is the alleged violation of the *two-marker-generalization* in the combination singular subject - 3rd singular object. There is only a single person exponent, although two person markers are expected under my analysis because both arguments are singular. But note that under

/i/ is realized. This is the 3rd singular subject marker of the indefinite non-past conjugation, see (18). Hence, it does express the required person value but in the wrong conjugation class. With respect to the occurrence of /i/ in the definite paradigm, Abondolo (1982: 17, 18) claims that the expected /(n)ze/ has indeed been the person suffix in this context in the definite conjugation, as textbooks from the 19th century show. It can still be found in the corresponding cell in the definite past paradigm. The marker has been replaced by the 3rd person marker /i/ from the indefinite conjugation. We could implement this by postulating a rule of referral Stump (2001). See the discussion of the indefinite paradigm below. The last form which is special is 3sg subject-2sg object. In this case the exponents are the expected ones, but in the wrong linear order. Usually, a VI expressing object features preceded a VIs which cross-references subject features (person on v > number on T in accordance with the Mirror Principle, cf. Baker (1988)), but in this combination it is the other way around (cf. Abondolo 1982: 16). The present approach cannot explain this fact, but neither can any of the other analysis of verbal inflection in Mordvin. Note that the order of person and number markers is generally reversed when a second cycle of person agreement obtains, hence, when the object is 3rd person. In this case the number exponent precedes the person exponent. Both the person and the number value are then present on T and the order cannot be determined by the Mirror Principle. I assume that the order of exponents is determined by a language-specific principle which orders number before person when two exponents spell out features of the same syntactic head.

¹⁶ Another option for the exponents of the definite conjugation would be to replace the context feature *pl* of the second cycle person morphemes by the morpheme /iz/. The second cycle person morphemes can then only be inserted if the plural morpheme /iz/ has already been inserted, they would be inwards-sensitive allomorphs of the first cycle person morphemes in the sense of Bobaljik (2000). In this way, secondary exponence is not necessary. However, I do not pursue this option here because it does not work for the exponents /ne/ and /de/ of the indefinite conjugation which are preceded by a different plural marker, cf. section 7.

the present analysis a situation arises in these contexts which does not occur elsewhere: There are two identical person values on the head T. One person value is the result of valuation of the regular person probe at the second cycle. π is raised to T because the object is 3rd person and cannot value π at the first cycle. The second person value is the result of the valuation of the former number probe (the *Chameleon probe*) because both arguments are singular. I propose that there is a constraint which rules out identical feature values on the same head and that one of the feature bundles is deleted before vocabulary insertion. Such an effect is also observed by Nevins and Sandalo (2010) in Kadiwéu and Georgian. They argue that “two instances of [+participant] within the same T domain is banned by a dissimilatory ‘OCP’ of two identical marked morphosyntactic feature values within the same domain” (Nevins and Sandalo 2010: 11). They propose that the more marked of these values is then deleted. I adopt their suggestion in an expanded version for Mordvin: Two instances of the feature person on the same head (T) are banned and one of these is deleted. It does not matter which one, because the values are identical anyway. This additional assumptions which is at work in other languages as well, explains the alleged violation of the *two-marker-generalization* in the combination singular subject – 3rd singular object. In the syntax, there are two valued probes, but one value is deleted in the morphological component prior to vocabulary insertion such that we cannot see the two valued probes at the surface.

7. The Indefinite Conjugation

So far I have focused on the derivation of the definite conjugation in which two arguments are potential goals for Agree with π and #. In this section I show that the system is also able to derive the indefinite conjugation. Remember that intransitives as well as transitives in non-perfective aspect and/or with an indefinite object fall in the indefinite conjugation in which there is only subject agreement. The subanalysed paradigms are repeated below:

(42) *Subanalysed indefinite
non-past paradigm:*

1sg	-n 1
2sg	-t 2
3sg	-i 3
1pl	-t -no ? 1
2pl	-t -do ? 2
3pl	-i -t 3 pl

(43) *Subanalysed indefinite
past paradigm:*

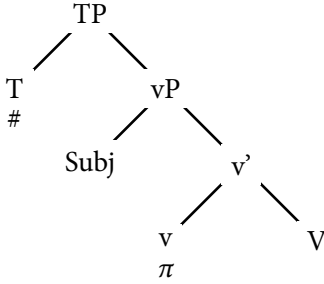
1sg	-n 1
2sg	-t 2
3sg	-∅ 3
1pl	-nek 1
2pl	-de 2
3pl	-t -∅ pl 3

Intransitives and transitives with an indefinite object or non-perfective aspect are treated in the same way. This suggests that they are structurally identical for the Agree process. As intransitives have only a single argument, the internal argument of a transitive verb must also be invisible as a goal which can explain pure subject agreement in both cases and makes intransitives and transitives with the properties listed above structurally alike. I won't speculate why indefinite objects or definite objects selected by a non-perfective verb are not accessible goals. This is a more general question about how the definiteness/aspect split comes about in the Uralic languages to which I have nothing new to add.¹⁷ What is crucial for the following analysis is that, obviously, these internal arguments of transitive verbs cannot be agreed with and they are thus like intransitives: There is only a single argument which is a potential goal for a probe. Assuming this to be the case allows us to derive the observed agreement patterns by the system developed for the definite conjugation. The tree in (44) repeats the initial distribution of the person and number probe in the syntactic structure with a single accessible goal. For reason of simplification, I omit the

¹⁷See Richards (2008) for a proposal about the distinction of definite and indefinite arguments, in which person features are not present on indefinite arguments. This idea is completely compatible with my analysis in which the person probe must not find a goal at the first cycle.

object in the representation also when the verb is indeed transitive, because the object is invisible for Agree.¹⁸

(44) *Probes in the structure:*



Starting with the person probe, it is clear that it cannot be valued at the first cycle: Either there is no object (intransitives) or it is invisible (transitives). Hence, π is raised to T, impoverished for [Part] and valued by the subject. There are two scenarios for the number probe on T. If the closest (and only available) goal is plural, it is valued. If, however, it is singular, the probe is impoverished for [Group]. But even then it cannot be valued by the singular subject because singular is absence of features. Neither can it be valued by person which is not a superset of the probe's features [R [Indiv]]. Hence, probe impoverishment applies for a second time, deleting [Indiv] and leaving only [R] on #. It is then a *Chameleon probe* and can be valued by person of the subject. Note that in the latter case, there are two identical person values on T, just as it was the case in the definite conjugation when a singular object acted on a 3rd singular object. Following Nevins and Sandalo (2010) I proposed that in such a situation one of the person values is deleted on T to avoid a kind of OCP effect.

These derivations which are in accordance with the analysis for the definite conjugation predict that there are two agreement patterns:

¹⁸I also abstract away from the difference between unaccusative and unergative intransitive verbs. I represent intransitives as unergatives with the single argument being an external argument. There are only few and very short descriptions of the syntax of Mordvin and other Uralic languages such that I did not find evidence for the split of intransitives into two groups. If this split exists, some complications arise for the present analysis, but as far as I see unaccusatives could still be integrated in my system with a few additional operations.

(45) *Expected agreement in the indefinite conjugation:*

- a. **singular subject:** π and # agree with the single argument in person; one value is deleted → one subject person marker on the surface
- b. **plural subject:** π valued by person of the subject; # valued by number of the subject
→ two suffixes: one person and one number marker for the subject

The first prediction is clearly borne out in the paradigms (42) and (43): There is a single marker cross-referencing person of the subject. Some complications arise for the second prediction. In accordance with it, there is always a person marker which indicates person of the subject if it is plural. If the subject is 3rd person plural, there is also the expected plural marker /t/. It differs from the plural marker /iz/ in the definite conjugation, but it is the regular plural marker in the nominal declension in Mordvin and according to Raun (1988: 234) and Zaicz (1988: 191, 199), 3rd person indefinite forms are diachronically derived nominal forms, hence the nominal plural marker. Note that this marker also shows up in the 1st and 2nd person plural of the non-past paradigm. Neither Zaicz (1988) nor Raun (1988) can tell us what this marker /t/ (/ta/ on the surface) means (hence the question marks in the glosses), but they agree in separating it as a further morpheme from the following suffix which is clearly a person marker (cf. the discussion of transparadigmatic syncretisms in section 3). Following the *Syncretism Principle*, I take this /t/ to be the same marker which marks plural in 3rd person – at least from a synchronic point of view; and as long as we do not know what exactly the function of this /t(a)/ was, this assumption does not contradict statements in the literature.¹⁹ The question marks can thus be replaced by the gloss “pl”, too. As a result, the second prediction is borne out in the non-past paradigm: There are two markers in a suffix string – one indicating person and the other expressing number of the subject. However, in the past paradigm the plural marker is only present for a 3rd person plural subject, although it should occur whenever the subject is plural. I do not have an explanation for its absence and take it to be accidental. In order to integrate this fact in the analysis, we might capture it by postulating an impoverishment

¹⁹Note that the marker /t/ is multiply ambiguous. We also postulated a /t/ as a present tense marker and as a 2nd person object marker in the definite conjugation. In principle, the /t/ of the indefinite paradigm could also be identical with one of those morphemes. But if the present approach is on the right track it should lead the learner to analyse /t/ as a plural marker, whatever other function it might have had in earlier stages of the language.

rule which deletes [plural] in the context of a [R [π [Part]]] subject in the indefinite past.

But on the whole, the predictions on exponence in the indefinite conjugation made by the analysis which was developed for the definite conjugation are borne out. I take this to support the extended version of the system proposed by Béjar (2003) with 3rd cycle agreement and a number probe that can be valued by person.

Finally, some VIs have to be added to the list in order to integrate the indefinite conjugation into the system. The second cycle morpheme /n/ which realizes 1st person occurs as expected. The second person marker is /t/ which was analysed as a first cycle morpheme above, but we expect the second cycle morpheme /k/. In order to capture the distribution of the 2nd person marker, I reanalyse /t/ as the default 2nd person marker which is not contextually restricted to v or T and thus fits whenever an argument is 2nd person. /k/ is more specific and blocks /t/ when person is valued on T. In order to avoid the insertion of /k/ in the indefinite conjugation, /k/ must be restricted to T in a transitive context. It is not clear which morpho-syntactic feature distinguishes T in a transitive and intransitive context. For the present purposes and for lack of a better solution I stipulate a feature [\pm trans] which is present on T, although it might not play a role in the syntactic derivation and is only needed for vocabulary insertion. /k/ is then restricted to T with the feature [+trans]. The remaining VIs which occur exclusively in the indefinite conjugation are specified for T [-trans] in order to prevent their insertion in transitive contexts. Note that /ne/, /i/, and / \emptyset / are additionally specified for tense because they only appear in the past or non-past indefinite paradigm, respectively.

(46) *More vocabulary items:*

- | | | |
|----|--|-----------------------------|
| a. | /k/ \leftrightarrow [2] / [T, +trans] | <i>replaces /k/ in (41)</i> |
| b. | /t/ \leftrightarrow [2] | <i>replaces /t/ in (41)</i> |
| c. | /i/ \leftrightarrow [3] / [T, -trans, -past] | |
| d. | / \emptyset / \leftrightarrow [3] / [T, -trans, +past] | |
| e. | /ne/ \leftrightarrow [1] / [T, pl, -trans, -past] | |
| f. | /de/ \leftrightarrow [2] / [T, pl, -trans] | |
| g. | /t ₂ / \leftrightarrow [pl] / [-trans] | |

8. Conclusion

In this paper I have provided an analysis of verbal inflection in Mordvin. Based on the analysis by Béjar (2003), I have shown that the double person marking pattern which is unexpected under her analysis with a single person and a single number probe can be derived with an independently motivated change in the structure of feature bundles. Including the feature [R] into these bundles, as proposed in the feature geometry by Harley and Ritter (2002), allows us to form a natural class between person and number. If the number probe is successively impoverished such that it only contains [R], which can only happen if both arguments of a transitive verb are singular, it can be valued by person features. In this way, double person marking obtains although the structure initially contains only a single person probe and a number probe. I called the number probe which is valued by person a Chameleon probe, indicating -descriptively spoken- that the probe can adapt the features it searches for to the syntactic context. In order to achieve this, no new mechanisms had to be assumed that have not also been part of Béjar's analysis. In general, I proposed a further repair strategy which may apply in order to avoid the crash of a derivation containing unvalued probes: valuation by features the probe originally did not search for.

Besides, I have shown that the system can also derive the indefinite paradigms, which are not taken into account in other formal analyses of Mordvin verbal inflection. Finally, it has been illustrated that a subanalysis of verbal suffixes is well-motivated and allows to derive transparadigmatic syncretisms across categories in Mordvin.

References

- Abondolo, Daniel (1982): 'Verb Paradigm in Erza Mordvinian', *Folia Slavica* 5, 11–24.
- Aranovich, Raúl (2007): 'Optimizing Verbal Agreement in Mordvin', *Studia Linguistica* 61(3), 185–211.
- Baker, Mark (1988): *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.
- Béjar, Susana (2003): *Phi-Syntax: A Theory of Agreement*. PhD thesis, University of Toronto.
- Bobaljik, Jonathan (2000): 'The Ins and Outs of Contextual Allomorphy', *University of Maryland Working Papers on Linguistics* 10, 35–71.

- Collinder, Björn (1957): *Survey of the Uralic languages*. Almqvist & Wiksell, Stockholm, pp. 228–246.
- Halle, Morris and Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale and S. J. Keyser, eds, *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.
- Halle, Morris and Alec Marantz (1994): Some Key Features of Distributed Morphology. In: A. Carnie, H. Harley and T. Bures, eds, *Papers on Phonology and Morphology*. Vol. 21 of *MIT Working Papers in Linguistics*, MITWPL, Cambridge, Mass., pp. 275–288.
- Harley, Heidi and Elisabeth Ritter (2002): ‘Person and Number in Pronouns: A Feature-Geometric Analysis’, *Language* 78, 482–526.
- Müller, Gereon (2004): A Distributed Morphology Approach to Syncretism in Russian Noun Inflection. In: O. Arnaudova, W. Browne, M. L. Rivero and D. Stojanovic, eds, *Proceedings of FASL 12*. University of Ottawa.
- Nevins, Andrew (2010): Unpacking Mordvin Portmanteaus. Talk given at DGfS annual meeting, Berlin 2010, Workshop on Portmanteaus.
- Nevins, Andrew and Filomena Sandalo (2010): Markedness and Morphotactics in Kadiwéu [+participant] Agreement. Ms., Harvard/UNICAMP.
- Noyer, Rolf (1992): Features, positions and affixes in autonomous morphological structure. PhD thesis, Massachusetts Institute of Technology, Cambridge, Mass.
- Raun, A. (1988): The Mordvin Language. In: D. Sinor, ed., *The Uralic Languages: description, history and foreign influences*. Brill, Leiden, pp. 96–110.
- Richards, Marc (2008): Defective Agree, Case Alternations, and the Prominence of Person. In: A. Malchukov and M. Richards, eds, *Scales*, Vol. 86 of *Linguistische Arbeits Berichte*. Universität Leipzig, Leipzig, pp. 137–161.
- Stump, Gregory (2001): *Inflectional Morphology*. Cambridge University Press, Cambridge.
- Zaicz, Gábor (1988): Mordva. In: D. Abondolo, ed., *The Uralic Languages*. Routledge, London, pp. 184–218.

