Morphological perspectives on the Georgian verb

Maël Gautier

Abstract

This paper explores in detail Georgian verbal agreement/tense morphology and aims at uncovering regularities and natural classes in this paradigm, using in turn the tools of Distributed Morphology and Channel Theory, or artificial learning. After reviewing part of the morphological literature on the topic of affix blocking in transitive verbs in the present, I turn to a decomposition of all non-perfect person/tense combinations. I offer a preliminary analysis in the Distributed Morphology framework, allowing to establish meaning-form regularities in the paradigm. Admitting difficulties linked to the latter framework (many operations are needed and seem to be ad hoc), I turn to a Channel (or Accessibility-based) analysis that does away with these operations. It is not evident that the latter account is radically superior, though, since many manipulations not mentioned in Keine's (2012) seminal paper seem to be necessary in this special case, calling for an extension/discussion of this model. The last part deals with a different approach, that of inflectional learning, where affixes are selected by an artificial learner in a serialist optimality setting. The latter approach reveals interesting findings, notably with regard to a form-meaning pair not previously considered as forming a natural class, namely the *-s* suffix of Georgian found not only in the third singular, but also the third plural.

1. Introduction

In this paper I provide a comparison of several accounts by examining their capacity to explain morphological exponence of Georgian verbs conjugated in some of the language's most common tenses, roughly the non-perfect ones. The central phenomena here form a superset of the data analyzed in previous accounts, which are either morphology-oriented (e.g. Anderson 1986, 1992, Carmack 1997, Stump 2001, Halle, Morris and Alec Marantz 1993) or syntax-oriented (Béjar 2003, Lomashvili 2011). Generally, the questions raised by these data amount to (sometimes sophisticated, as in Carmack 1997 or Béjar 2011) analyses aimed at providing well-grounded reasons for the behavior of Georgian plural suffixes, i.e. why a third person plural subject argument (suffix

Topics at InfL, 121–152
A. Assmann, S. Bank, D. Georgi, T. Klein, P. Weisser & E. Zimmermann (eds.)
Linguistische Arbeits Berichte 92, Universität Leipzig 2014

-*en*) is enforced to the detriment of the object's plurality (suffix -*t*), and why the latter is realized to the detriment of third person singular subject (suffix -*s*). Another question often raised in these accounts is why the prefix (or clitic) corresponding to a second person object (*g*-) systematically takes precedence over a first person subject's one (*v*-).

There I take up these issues to show some of the morphologically-oriented means that have been devised, i.e. the original Halle, Morris and Alec Marantz (1993) approach and the Extended Word and Paradigm approach of Anderson (1986, 1992), and I quickly review their results as far as Georgian is concerned. I further aim at giving an account of each Georgian verbal affix in several tense-aspect-mood combinations. In section 2 I offer an analysis fully in line with Halle, Morris and Alec Marantz's (1993) Distributed Morphology approach, with the addition of a set of data interacting with the previous smaller set.

In section 3 I contend that the relatively high number of morphological (known as postsyntactic) operations needed in Distributed Morphology (Vocabulary Insertion, Fusion, Fission, Impoverishment, Readjustment) renders worthwhile exploration of an alternative approach, the Accessibility-based theory developed in Keine (2012), which explicitly argues against such a variety of operations. The exact same set of data receives a slightly different segmentation and is not subject to as many operations as in Distributed Morphology, but nevertheless turns out to be particularly hard to handle with regard to the treatment of features and position classes.

Section 4, on inflectional learning (Bank and Trommer 2012), tackles higher-level analysis possibilities, whereby artificial creation of affix sets is more central than inter-affix relationships and, e.g. blocking. This optimality-based, serialist approach, gives insight into the role of constraint rankings, for a given language, in selecting affix hypotheses by checking their numerical accuracy with regard to their distribution in a whole paradigm.

Overall, it will be seen that neither of the two morphological analyses fares perfectly well with this set of data, either because of many operations that are often criticized and considered as ad hoc devises (Distributed Morphology), or because of a high number of diverse manipulations on features and position classes (Accessibility). The learning approach has a different status in that it has to occur before any morphological analysis proper; no special criticism is offered, since it is essentially a means of constituting affix sets which are themselves the matter onto which analyses are performed, i.e. the affix sets are potential inputs to virtually any morphological analysis in the sense of Distributed Morphology or Accessibility.

2. Georgian verbal agreement

One often studied aspect of the language is its verbal agreement morphology. Purely morphological analyses are among others Anderson (1986) and Halle, Morris and Alec Marantz (1993). The central data in these accounts are represented by the pattern of regular transitive verb forms in the present tense[1] including both subject and object markers. The paradigm is shown in table 1. The fact noted in the cited approaches is known as blocking: in some places, markers that are expected do not show up. This is the case in:

- the forms with first person subject and second person objects, where the first person subject prefix *v*-, which is seen with a 3 object, fails to co-occur with the second person object prefix *g*-
- the forms with 3sg subject and 2pl object, where the 3sg subject suffix -*s* is overridden by the plural suffix -*t* (in this case marking plural of the 2 object)
- the forms with 3pl subject and 2pl object, where the same suffix -*t* fails to co-occur with the 3pl subject suffix -*en*

In this section, I first briefly review the approaches by Anderson (1986) and Halle, Morris and Alec Marantz (1993). It will be seen that the latter, morphemic approach, allows to derive the blocking effects more explicitly than the former.[2] Then I present a larger set of verbal forms that have not been systematically studied before (transitive verbs in five screeves). To conclude, I will propose a morphemic analysis of this data.

2.1. Previous analyses

Anderson (1986) developed a word-based model of affixation where the traditional notion of a meaningful inflectional morpheme is rejected and a system of

[1] Or present *screeve*, as tense-aspect-mood (TAM) paradigms are called in the traditional literature.

[2] Although for this it is necessary to accept a set of operations not postulated in Anderson's (1986) work.

Present

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…	v-…		g-…-t	v-…
2sg	m-…		…	gv-…		…
3sg	m-…-s	g-…-s	…-s	gv-…-s	g-…-t	…-s
1pl		g-…-t	v-…-t		g-…-t	v-…-t
2pl	m-…-t		…-t	gv-…-t		…-t
3pl	m-…-en	g-…-en	…-en	gv-…-en	g-…-en	…-en

Table 1: *Georgian transitive verb (present)*

interacting rules is preferred. These *Word-formation rules* (WFRs) are organized in blocks so as to correspond to position classes, i.e., each block of rules can be seen as adding material at a specific location in a word-form, which does not exclude application of rules from other blocks. Inside each block, rules apply disjunctively: an applicable rule more specific than another one applies and prevents application of the latter. I summarize below how Anderson's (1986) approach captures the mentioned cases of blocking.

***v-/g-* competition:** Both prefixes belong to the same rule block. *g-* is specified as [[..+2..]][3] and *v-* as [+1..]. It is not clear here how the specificity effect could arise; rather, it is an instance of the stipulated ordering Anderson assumes. It is simply stated that in the prefix list, *g-* must precede *v-*, since only this can yield the attested result.

***-t/-s* and *-en/-t* competition:** Again, a rule block is established for suffixes. An ordering suggested by the facts would be *-en* > *-t* > *-s*. This is precisely how the list is ordered in Anderson (1986). The relative order of *-t* and *-en* is given by specificity, since both realize third person, but the latter realizes a plural feature in addition. However, that *-t* is located in the middle of the list is again stipulated, since it realizes only plural and isn't more specific than *-s*.

In the syntax-based realizational model of Distributed Morphology (Halle, Morris and Alec Marantz 1993), fully specified syntactic positions (*morphemes*) are the target of vocabulary insertion, which provides them with a phonological form. The Vocabulary Items are pairs of such a phonological string and a

[3]In Anderson (1986), embedding in such morphological structures stands for subjectness/objectness: here the features +2 have two levels of embedding, meaning that they signal the object.

possibly underspecified morphosyntactic specification. The basic requirement for insertion is that the items' features have to be non-distinct from that of the morpheme. Also, of two items competing for insertion, the most specific one is the one which is inserted, i.e. the one with more features to realize. The framework also makes use of four distinct operations that all find an illustration in this Georgian excerpt.

***v-/g-* competition:** In presence of a first person subject and a second person object, the syntactic features of both these arguments are each dominated by a pronominal clitic head. What prevents insertion of both prefixes is *fusion*, which gathers the features of two heads under only one head. Once this is done, and since vocabulary insertion occurs only once for a given head, a unique exponent can be inserted. *v-* expressing solely (first) person, and *g-* bearing a case feature in addition to its person feature – 2, ACC –, the latter systematically wins the competition.

***-t/-s* and *-en/-t* competition:** *-t* is a plural suffix reserved for first and second person arguments, but the latter are typically realized in prefix position. It is taken for granted that the plural feature of such arguments (except 1pl object) is generated there, and relocated in the suffix position through the operation of *fission*. Once this is done, the feature may interact with other features in postverbal position. Blocking of *-s* by *-t* is due to a *readjustment* rule which deletes the segment /s/ if followed by a plural feature. Blocking of *-t* by *-en* is due to an *impoverishment* rule which deletes the plural feature whenever preceded by the features 3 and pl.

From what precedes, Distributed Morphology fares better when it comes to explaining such blocking effects: in one case the principle of specificity is respected (contrary to Anderson 1986), and in the other the effect can be derived by means of special operations not available to Anderson (1986). I now turn to an extended set of data, adopting the analysis developed in Halle, Morris and Alec Marantz (1993).

2.2. Screeves in Distributed Morphology

In the descriptive and/or traditional literature (e.g. Tschenkéli 1958), the total number of screeves is eleven. They are grouped into a super-category *series*, of which there are three. Series I has six screeves and is traditionally called *present series*, although it also includes three future screeves; Series II has two screeves, which are characterized by perfectivity; Series III has three screeves and is

[−Pa]	[+Pa]	[−Pa]	[+Pa]	[−Pa]
[−P]	[−P]	[−P]	[+P]	[+P]
[−S]	[−S]	[+S]	[−S]	[+S]
present	*imperfect*	*subjunctive*	*aorist*	*optative*

Table 2: Feature specification of the five screeves

called the *perfect series*. The latter series exhibits the property of inversion for transitive (and unergative) verbs: verbal agreement and case-marking are subject to a redistribution whereby subjects are marked as indirect objects and direct objects as subjects. Since these changes are not directly relevant to the kind of morphological explanation I am pursuing here, I restrict myself to Series I and II.

As indicated, I tentatively take for granted the Distributed Morphology analysis exposed so far. That is, I assume the same as Halle, Morris and Alec Marantz (1993) about the prefixes (as pronominal clitics) and the suffixes *-s*, *-t* and *-en* (and its past aorist counterpart *-es*). This decision has direct descriptive consequences: the amount of data to analyze will be reduced in that the difference is abolished:

– between first and second person, since they never differ by anything else than the prefix

– between 1/2 sg/pl, since the plural suffix in these cases is normally *-t* (except in the case of a 1pl object, which calls for a specific *prefix*)

Only three distinct person/number (PN) combinations are then needed. These are all concerned with the specifications of the subject, since the only possible object affixes are *g-*, *m-*, *gv-*, as well as *-t*, have already been accounted for; moreover, the only subject prefix, *v-*, is the only difference between first and second person in all circumstances and thus doesn't need to be taken into account here. The relevant combinations are the following: 1/2 sg/pl, 3sg and 3pl.

As for the TAM component of the conjugation patterns, choosing Series I and Series II should lead to analyze eight paradigms, but examining Series I, it is clear that its future subpart is a morphological by-product of the present

	present [−Pa−P−S]	imperfect [+Pa−P−S]	subjunct. [−Pa−P+S]	aorist [+Pa+P−S]	optative [−Pa+P+S]
2sg	c'er	c'erd-i	c'erd-e	c'er-e	c'er-o
3sg	c'er-s	c'erd-a	c'erd-e-s	c'er-a	c'er-o-s
3pl	c'er-en	c'erd-nen	c'erd-nen	c'er-es	c'er-o-n

Table 3: TAM/PN exponence with the verb *c'era* ‚write'

subpart.[4] Thus the future tense is not taken into account for the morphological analysis, which reduces the number of screeves to five. The total number of forms is then 15 (3 PN combinations × 5 screeves). The next move is to attribute featural content to the screeves. I assume three binary features [±(Pa)st], [±(P)erfect] and [±(S)ubjunctive]. Table 2 represents all relevant screeves with their TAM-feature specifications. It can be seen that the most prominent natural classes revolve around the P(erfect) feature, since all screeves of the so-called *present* series (I) are −P while the two screeves of series II are +P. Table 3 represents the surface forms corresponding to these PN/TAM combinations with the regular verb *c'era*, "write".[5] I also assume that two positions can follow the verb stem; this is suggested by the possibility of, e.g., either *-o* alone (optative), or *-s* alone (3sg present and subjunctive), or *-o-s* (3sg optative). Thus a position 1 is dedicated primarily to TAM fetaures, while a position 2 is dedicated primarily to PN features (in fact 3sg/pl features).

The issue is to attribute a feature content to the affixes of the paradigm. Interesting characteristics are: 2sg imperfect *-i* stands isolated and has as a 3sg counterpart *-a*, which also appears in the aorist; the 2sg counterpart of 3sg *-a* in the aorist, however, is not *-i*, but *-e*, which also marks 2sg *and* 3sg in the subjunctive; 3sg *-s* is in complementary distribution with *-a*, but not with *-e* or *-o*; 3pl present *-en* seems to carry over to the other imperfective columns, but with an added segment: *-nen*, and appears as simple *-n* after optative *-o*; the latter appears to have the fullest distribution among the TAM affixes.

This pattern can be captured by positing two lists of vocabulary items, as in

[4] It differs only by the addition of a preverb.

[5] The 3pl pattern seems to be the most involved one with many /n/'s and an aorist /es/ reminiscent of 3sg subjunctive /-e-s/. The choice of segmentation, here plain *-es*, is a tentative one and could have been *-e-s* under different assumptions.

(1) and (2). Subscripts correspond to position 1 or 2 and serve for the indication of contexts. The two readjustment rules in (3) are dedicated to the 3pl affix /-n/.

(1) **TAM-position**

a.	[+P+S]	↔	/-o/	
b.	[+Pa+P]	↔	/-e_1/	/___$[-3]_{P2}$
c.	[+Pa–P]	↔	/-i/	/___$[-3]_{P2}$
d.	[–Pa–S]	↔	∅	
e.	[+Pa]	↔	/-a/	/___$[+3-pl]_{P2}$
f.		↔	∅	/___$[+3+pl]_{P2}$
g.	[+S]	↔	/-e_2/	

(2) **P/N-position**

a.	[+3–pl]	↔	/-s/	/ $[-Pa]_{P1}$___
b.	[+3+pl]	↔	/-es/	/ $[+Pa+P]_{P1}$___
c.	[+3+pl]	↔	/-n/	

(3) **/-n/-readjustment**

a.	∅	→	/<e>n/	/ <$[[-Pa], [-P], [-S]]_{P1}$> ___ $[[3],[pl]]_{P2}$
b.	∅	→	/<ne>n/	/ <$[[+Pa], [-P]]_{P1}$> ___ $[[3],[pl]]_{P2}$ ∨
				/ <$[[-P], [+S]]_{P1}$> ___ $[[3],[pl]]_{P2}$

This approach effectively accounts for the distribution of the affixes. Note that, in the TAM position, the first four VIs need not be ordered as they are featurally divergent. To avoid the insertion of any element in the 3pl TAM position apart from optative /-o/, an empty element is extrinsically ordered before the subjunctive affix /-e_2/. Two /-e/ VIs are posited since they appear in featurally disjoint environments. Context features crucially restrict the distribution of affixes in both sets. Thus, for instance, the generalization that 3sg *-s* only ever surfaces in the present, the subjunctive and the optative is captured by its contextual feature [–Pa], which makes reference to the adjacent TAM position, and the two [+Pa] items /-e_1/ and /-i/ are banned from third person contexts. The surface variation of 3pl *-n* is captured by two readjustment rules which extend the affix by one (→ *-en*) or two (→ *-nen*) segments in the relevant contexts (present and imperfective paradigms).

Let's recall that the above analysis is performed on a subpart of the general paradigm, since I adopt Halle, Morris and Alec Marantz's (1993) analysis of the prefixes and plural *-t*. If both analyses are to be integrated in this way, then, again, the operations of fusion, fission, readjustment and impoverishment have to be adopted, too. A number of objections, be they linked to Halle, Morris and Alec Marantz's (1993) analysis or to the present one, can be raised with regard to the machinery needed. First there is the need for some empty elements, as materialized in (1) by the present marker and the marker with a 3pl contextual feature. Such devices are driven only by the necessity of blocking some other affix. The case of the present zero affix doesn't seem to be too problematic if it can be seen that the present tense is effectively an unmarked one where absence of a marker is in complementary distribution with presence of a marker in other tenses. But, second, whereas acceptance of this marker might be helped by the fact that it realizes a characteristic set of features (−P−S), this is not the case with the other empty element. Namely, the latter does not realize any features – it simply says that the TAM position may not be filled in a given context. This is because it is impossible to write a TAM entry with 3pl substantial features. Moreover, and this is the third point, the fact that this blocking zero affix has to be extrinsically ordered before another affix (/-e/) is not warranted in a system whose explicit goal is to rely fully on specificity. It is not clear how this important issue could be solved. Finally, concerning the 3pl marker proper, the use of readjustment to account for its different shapes (provided it is one affix), has the flavour of a set of ad hoc rules, since apparently arbitrary phonological modifications are modelled as the direct consequence of some morphosyntactic features' presence[6]. In addition, as this account of the distribution of *-nen* is based on the featural specifications [+Pa−P−S] (imperfect) and [−Pa−P+S] (subjunctive), it faces the problem of the disjunction seen in (3-b), specifying that the rule has to apply to divergent feature sets.

Overall, then, the present DM analysis has some shortcomings which, albeit not fatal (a weakened version of the theory emerges from allowing specificity and extrinsic ordering to be intermingled, from allowing disjunctive rules etc.), call for a theoretical cross-comparison. The next section is devoted to an analysis of the same data in an *accessibility-based* framework (Keine 2012).

3. An accessibility-based analysis of Georgian

The accessibility-based framework developed in Keine (2012) (also known as Channel Theory) is distinct from Distributed Morphology in an important

[6] It is not obvious that listing three different affixes (/-n/, /-en/ and /-nen/) would be a less arbitrary option, though.

respect: it seeks to do away with the assumption that the operation of vocabulary insertion is only conditioned by the morphosyntactic features of exponents. The argument is based on apparent morphology-syntax mismatches that come about whenever some exponent is inserted in an environment with a conflicting feature specification. Classically, such cases are treated by resorting to postsyntactic operations (e.g. impoverishment, feature-changing, enrichment). Since Keine (2012) argues against the multiplication of such operations, a system is developed there to dispense with them.

The system relies on pairwise accessibility relations holding among exponents of a language. This kind of relation is formulated in the following statement: "The exponent chosen at step n affects the set of exponents competing for insertion at step n + 1". Thus, informally, markers select which other markers can be inserted after them. The set of these relations is one of the two idiosyncratic sets relevant for a given language – the other one being the set of exponents. The insertion algorithm postulated here makes use of the notion of *state*: the derivation of morphological forms goes through successive states, starting with an initial state. Each state is conceived of as a triple: an exponent, a set of morphosyntactic features, and a set of phonological features. The initial state (represented as ℵ) is the location of root insertion, where no exponent is present, but two fully specified sets of features are present as in Distributed Morphology.

Successive insertion of exponents is as a sequence of transitions. A transition from one state to the following is one where insertion of an exponent subtracts morphosyntactic features from the corresponding set of the former state and adds phonological feature to its corresponding representation. For a transition to be well-formed, the exponent of state n+1 has to: be accessible from the exponent of state n ("exp.1 → exp.2"); the morphosyntactic features of the exponent to be inserted have to be a subset of those of the relevant state (Subset Principle); among the accessible exponents, the most specific one is inserted (Specificity).

The relevant version of Specificiy assumed in the system not only makes reference to the number of features that the exponents possess, it also makes reference to classes of features. The set of classes is a hierarchical object to be defined on a language-specific basis. The impact this has on the specificity of individual exponents, informally, is that, of two exponents with the same number of features of class F, the one that possesses more features from a higher-ranked class G is more specific (Müller 2004).

Another crucial assumption is that once features have been discharged by insertion, they become unretrievable (the Strict Feature Discharge Theorem). In this way, it is impossible for the morphosyntactic specification of an exponent to make reference to features that have previously been deleted by insertion of another marker. Finally, the system does not make use of contextual features: if features are mentioned in the specification of an exponent, they automatically discharge the corresponding features of the state – the latter cannot remain unaffected, as would be the case for contextual features.

3.1. Illustration

The empirical ramifications of this framework are explored in some detail in Keine (2012), basing on phenomena drawn from different languages. It is shown for instance how multiple exponence in Archi can be handled with the accessibility framework. Multiple exponence is problematic for frameworks which make the two assumptions that only the features of the exponents are relevant to insertion and that features on the input can only be active once. Namely, the effect of these combined assumptions is that a means has to be devised to insert an exponent specified for a feature that has already been discharged by a previous exponent. In Archi, this takes the form of an apparently doubly marked plural feature on oblique nouns (4). Concretely, the language distinguishes nominative nouns and nouns marked by any other case by means of a stem extension on top of which more specific case markers are added. A very minimal example is shown in (4).

(4)

	aInš		**dab**	
	SG	PL	SG	PL
NOM	aInš	aInš-um	dab	dab-mul
ERG	aInš-li	aInš-um-čaj	dab-li	dab-mul-čaj

In this small excerpt, two nouns, belonging to different classes, each exhibit a class-specific but case-independent plural suffix *-um* or *-mul*. In the singular, the difference between a nominative and an oblique noun is signaled by the marker *-li*. In the oblique plural, however, *-li* is superseded by the class-specific plural suffix, while a plural/oblique-specific suffix *-čaj* follows the class-dependent plural suffix. The pattern is clearly a case of multiple exponence, since *-čaj*, which occurs only in the oblique plural, repeats the plural feature already expressed. Keine (2012) reviews briefly some feature-based accounts that have been entertained. Under the secondary exponence approach of Noyer (1997), *-um*/*-mul* could be primary exponents of plural, while *-čaj* would realize primarily the oblique feature and only secondarily the plural feature (as a bracketed diacritic). Under the enrichment approach of Müller (2007), a specific rule creates a plural feature in the environment of a plural *and* oblique specification (plural duplication), allowing the insertion of both suffixes.

However, a slightly divergent set of facts from Archi is brought to light by Keine (2012). Archi has numerous nominal classes, as well as lexical exceptions, and not all of them conform to the above pattern. Two instances of this are shown in (5).

(5)

	haˁtəra		**χˁon**	
	SG	PL	SG	PL
NOM	haˁtəra	haˁtər-mul	χˁon	buːc'i
ERG	haˁtər-čaj	haˁtər-mul-čaj	χˁini	buːc'i-li

The word *haˁtəra*, "river", has *-čaj*, and not *-li*, as a general oblique marker, i.e. also in the singular. Conversely, the word *χˁon*, "cow" (where suppletion plays the role of an overt plural marker), exhibits *-li* as an oblique marker in the plural, but not in the singular. It seems then that specifying *-li* and *-čaj* as either singular or plural suffixes is not sufficient and leads to a dilemma. The solution proposed by Keine (2012) relies on a radical kind of exponent underspecification whereby both *-li* and *-čaj* are pure oblique markers; the concept of accessibility is the corollary of this assumption in explaining the distribution of such affixes (Figure 1).

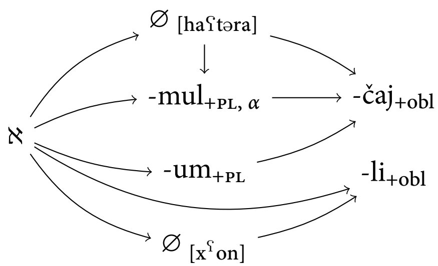

Figure 1: Archi oblique and plural markers

Figure 1 reveals a peculiar conception of exponence in that some exponents are zero ($\varnothing_{[ha^{ʕ}təra]}$): these lexical class-bound empty affixes, along with the accessibility relations they are involved in, are here to ensure that the right exponents (in this case either the suffix *-čaj* or the suffix *-um* are attached). Although I will not be concerned further with nominal inflection or with exceptional classes, I go on in the following section to develop a tentative analysis of Georgian verbal agreement within the accessibility framework. Modifications and related comments will be provided in the discussion.

3.2. Georgian

A tentative analysis of the Georgian data in an accessibility-based framework is found in figure 2.[7,8]

I make the assumption that all person, number and case features are decomposed. Any such features present in the morphosyntactic heads carry with them their negative counterpart: a nominative feature is represented as +Nom−Acc, a third person feature as −1−2+3, a plural feature as −sg+pl etc. This has consequences for the insertion process and is built into the hierarchy.

The initial state ℵ hosts three feature structures of the form Subject Argument – Object Argument – TAM. For instance, the verbal form *v-xedav*, "I see him", starts out with the features: [[+Nom−Acc+1−2−3][−Nom+Acc−1−2+3][−Pa−P−S]].

Of the two affix sets potentially encountered first after the initial state, one comprises all and only prefixes. It is possible to go to this set and then to the first set of suffixes, but not the other way around. Among the four prefixes, three are accusative, and one, *v*-, is nominative. To ensure insertion of the prefixes whenever they are attested, the hierarchy specifies that positive case

[7] Contrary to Keine (2012), where four different phenomena from different languages serve to illustrate the empirical coverage of the framework, here the data is constituted of all (regular) markers of all (non-inverse) tenses of one language. Numerous accessibility relations are not shown in the graphic representation. Instead, affixes are grouped together, (roughly) following co-occurrence possibilities. Thus readability improves, but this doesn't mean that all affixes of a block are accessible from all affixes of the previous block (e.g., *-a* is not accessible from *v*-). In such cases, it suffices to observe that their features are either identical or contradictory, making insertion of the second one impossible.

[8] Here prefixes are mixed with suffixes, although their behavior is different from the latter, as evidenced by the entire TAM paradigm, where it can be seen that they are completely insensitive to the TAM category. Nevertheless, I see no a priori reason to exclude them, since they realize feature-categories (person, number and case) that are relevant for suffixes as well.

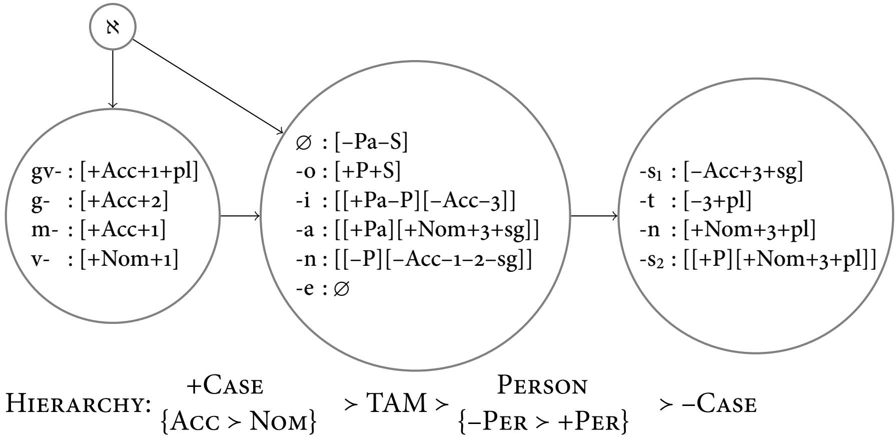

Figure 2: Georgian verbal agreement in Accessibility

should be the top-ranking feature class. In addition, (positive) accusative case takes priority over (positive) nominative case and both take priority over negative case in order to prevent too early insertion of a suffix; this is because some suffixes (*-a*, *-n*) do carry a case specification plus a number of additional features that would otherwise make them more specific. For example, were +Nom and +Acc on the same level, *-a* would be selected before, say, *m-*, because it possesses a TAM feature, a person feature and a number feature in addition. As suffixes normally only express features of the subject (with the exception of *-t*), *v-*, as the only nominative prefix, is the only one that could be in competition with suffixes. This is the case with *-i*, which does co-occur with *v-*. But since *-i*, carrying person features, must also specifiy the case of the relevant argument (the case of a subject), it would take priority over *v-* if it were specified as +Nom, since it possesses more features than it in addition to shared +Nom. The solution here is to specifiy *-i* as −Acc rather than +Nom, and rank −Case (short for −Acc > −Nom) lower than +Case.

Turning to the first set of suffixes, two general markers, specified only for TAM-features, are apparent: a ∅-element for present and *-o* for the optative. The similarity between the two sets of forms is clear when seeing that the present has nothing wherever the optative has *-o*, and this distribution suffers no exception whatever the person-number combination. Next comes *-a*, which is restricted to 3sg in the past, that is, imperfect or aorist. *-i* is even more restricted, realizing 1/2 person (−3) only in the imperfect, picked out by the features +Pa−P. The case of *-n* is more intricate, since it depends largely on the morphemic analysis that is chosen. Indeed, some form containing *-n* characterizes third person plural in all tenses apart from the aorist, and the form of the postradical string is not uniform: *-en* in the present, *-nen* in the imperfect and the subjunctive, *-n* in the optative. The choice here is to postulate two different *-n* belonging to two different affix sets, allowing for a sequence of two *-n*. The *-n* of the first suffix group is the one of the imperfect/subjunctive: its specification as −P restricts it to the three imperfective tenses. A case feature has to be specified to indicate which argument it encodes (the subject). For the ∅ of the present to take priority over it, leaving it only to the two latter tenses, *-n*'s case specification cannot be +Nom, since it would then be inserted also in the present (I take the *-n* of the present to be the one of the second suffix set). To solve this, *-n* is given the specification −Acc instead, which belongs to the lowest-ranking feature class. *-e* is the elsewhere suffix. Its range is constituted of the subjunctive of all persons apart from 3pl and the aorist of all persons apart from 3sg.

In the second set of suffixes, *-t* signals plurality of either a first or second person subject or of a second person object. This distribution is captured by the features −3+pl; lack of a case specification ensures this result, while the impossibility of *-t* to signal plurality of a 1 object follows simply from the feature content +Acc+1+pl of the dedicated affix *gv-* (in contrast to *g-* with only +Acc+2, +pl is already discharged). This affix set comprises two *-s*. $-s_1$ is the 3sg exponent in the present, the subjunctive and the optative. Again, as for *-i*, a case feature has to be provided to indicate which argument the affix realizes, and again, specifying it as +Nom would wrongly make it more specific than another affix (*-t* doesn't bear any case feature). Step one of the solution is to associate it with a −Acc feature instead. As −Case is the lowest-ranking class, the difference between the two affixes is neutralized – they both contain as many PN features. Step two consists in ranking −Person above +Person: as *-t* bears a −3 feature, it is more specific than $-s_1$, which bears a +3 feature. This derives one of the two well-known blocking effects involving *-t* ($-t > -s_1$). Accounting for the non-occurrence of $-s_1$ in the past tenses imperfect and aorist, affix *-a*, which is specified as +3, discharges this feature, rendering insertion of $-s_1$ impossible afterwards. The second *-n* is the most general 3pl affix of the inventory; it is designed to occur in all 3pl forms except that of the aorist. Thus in the present and the optative, it shows up preceded by either ∅ or *-o*, while in the imperfect and the subjunctive, it is preceded by the first *-n*.[9] $-s_2$ is the specific 3pl aorist affix, whose feature content differs from *-n* only by the +P feature. It may only be inserted after the elsewhere affix *-e* since all others block it: for instance, optative *-o* discharges the +P feature, making insertion of $-s_2$ impossible. The other well-known blocking effect involving *-t* (*-n, (-e)-s* > *-t*) is accounted for straightforwardly if both 3pl exponents are specified for a positive case feature, namely +Nom: these will always take priority over *-t* since +Case is higher-ranked than person and number features.

3.2.1. Summary

Thus the system developed here offers the beginning of an account of the Georgian data in an accessibility-based framework. The reasons why it is only a beginning are as follows.

First, the account exhibits one visible technical drawback in the interaction of exponents *-i* and *-t*. *-i* cannot be left without a case/person/number specification, since it is not a *general* imperfect marker: it has to be restricted to first and second person; with a +Pa–P specification, it would be incorrectly inserted in 3pl imperfect/subjunctive contexts instead of *-n*. The problem has to do with the feature content of *-t*, which also has to be specified as –3. As it stands, *-t* couldn't ever follow *-i*. A putative, and non-standard, solution would be to specify one of the two suffixes as realizing some "conjunction" of the first and second person features – something like +1,2. But even this wouldn't be satisfactory, since a full sequence *v-*…*-i-t* is part of the well-formed expressions of the language. In such a case, *v-* would discharge +1, and no person specification – positive or negative – of subsequent *-i* and *-t* would allow both of them to be inserted. This is clearly the beginning of a puzzle for the way an accessibility-based theory handles features.

Second, the interaction of the notions of position class (cf. the graphical representation) and of feature classes is not yet entirely clear. True, the paradigm appears to have a templatic character in that certain positions seem to be dedicated to a restricted set of suffixes, moreover allowing sequences like *-o-s/-o-t, -e-s/-e-t* where some autonomous affixal material is shared. True,

[9] Of course, these choices do not yield exactly the right exponence in the present/imperfect/subjunctive, since the attested forms really contain *-en* and *-nen*. Ultimately, some phonological process should account for the apparent epenthesis of /e/ between the affix and the right stem boundary.

inspection of this pattern reveals that TAM features play a prominent role in the position immediately following the stem, while PN features are more prominent in the following position. The ideal picture would be that there are no exceptions to these statements. Still, two affixes pose a problem. The elsewhere affix *-e* is located in the first suffix group because of its obvious morphotactic resemblance to the other vocalic suffixes *-a, -i* and *-o*, although, as an elsewhere, it realizes no morphosyntactic feature at all. The question is then why such a suffix should be grouped with the TAM suffixes. The $-s_2$ suffix is located in the second suffix group because it has to occur after *-e*, but the fact that it (obligatorily) bears a TAM feature (+P) should lead, under the above generalizations, to locate it in the precedent suffix group. The present accessibility-based analysis of Georgian thus gives no conclusive evidence that "position classes" and feature classes correlate in a coherent way.

Third, it is not clear how to handle feature specifications like +Nom−Acc or +1−2−3. There is no theoretical motivation that I know of that would call for decomposing case features into something else than unique (binary or privative) features. It is nevertheless true that the present analysis would not function without this addition to our inventory, for two reasons. First, as in the case of $-s_1$ vs. *-t*, this is the only way (in combination with the hierarchy) to ensure that the attested blocking relations are captured. Second, and this is an important point, this assumption allows to mimick multiple exponence. At an intuitive level, this solution distributes "identical" feature realization among exponents as if there were "primary" and "secondary" exponents. But as was shown with *-i* and *-t*, use of binary features only reaches this aim to the extent that the language has no more than two exponents realizing the "same" feature.

Thus, as in the case of Distributed Morphology, the possibilities provided by Channel theory suffer from some apparent drawbacks when applied to a full inflectional paradigm like the present Georgian one. The complexities required to explain this case and the paradoxes that it brings are not necessarily fatal, but suggest that each of the above points should be taken as separate problems to be considered in further research.

4. On Inflectional Learning

This section is concerned with computer-learning possibilities of the paradigms studied so far. The specific algorithm and assumptions developed in Bank and

Trommer (2012) form the basis of the discussion to come. I present here yet another analysis of the above Georgian data, but from this different point of view. In Bank and Trommer's (2012) model, the theoretician doesn't control directly the assignment of a meaning to affixes as they do in, e.g., Distributed Morphology or Channel Theory – rather this operation is mediated by the learner hypothesizing multiple form-meaning pairs (*morpheme hypotheses*) and checking them against patterns of distribution within full-scale inflectional paradims. Manipulation of learning possibilities is done in an optimality-theoretic setting, which allows the learner, according to a given constraint ranking, to "decide" what morphosyntactic content is best suited for a given phonological string on the basis of its distribution in the paradigm. Typical constraints refer to actual occurrence/predicted occurrence ratios of strings in (sets of) featurally fully specified paradigm cells, and establish a score that determines which features an affix should realize; re-ranking of course potentially gives different results. The first subsection gives indications as to how this is implemented in the case of Ainu, an example taken from Bank and Trommer (2012). The following subsections are concerned with Georgian and adopt the exposed point of view in taking whole paradigms as a subject of study. The goal will be to get insights into potentially different choices of affixes and meanings thereof, as compared to the previous two accounts, Distributed Morphology and Channel Theory.

4.1. Illustration with Ainu (Bank and Trommer 2012)

The language studied in Bank and Trommer (2012), Ainu, has the property of possessing affixes with a problematic distribution. Table 4 is an Ainu transitive paradigm in simplified version: blank cells actually display affixes, which have been omitted here for illustration (gray cells stand for reflexive forms which are not expressed in the same way as transitive forms and are not included). It is used to exemplify what are called the *imperfect distribution* and *meaning assignment* problems in inflectional learning.[10]

This simplified paradigm has only two affixes, *e-* and *eci-*, which are present in most cases where a second person argument is present, be it a subject or an object. It can be seen that *e-* signals a 2sg argument when the other argument is third person; a different prefix, not shown here, is used when the subject is

[10] A third problem, the *subsegmentation* problem, which is tackled in Bank and Trommer (2012), will not be discussed here.

S \ O	1sg	1pl	2sg	2pl	3sg	3pl
1sg			eci-	eci-		
1pl			eci-	eci-		
2sg					e-	e-
2pl	eci-	eci-			eci-	eci-
3sg			e-	eci-		
3pl			e-	eci-		

Table 4: Ainu simplified transitive paradigm (Bank and Trommer 2012)

2sg and the object first person; in all other cases, *eci-* is used. A puzzle arises when considering what these prefixes actually stand for: it cannot be said with perfect accuracy whether they each mean 2sg or 2pl. That *eci-* is present in all 2pl cells is clear, but assigning it this meaning leaves its presence in the two first person subject/2sg object cells unexplained. Choosing the reverse option, namely assigning it the general meaning second person is also problematic, since it wrongly predicts its occurrence in a number of cells, the ones with a 2sg argument (except for the 1/2sg combination), which actually display *e-*.

The treatment adopted in Bank and Trommer (2012) relies on the following reasoning. An artificial learner is provided with the paradigm and makes multiple hypotheses about the meaning of affixes, i.e., it associates phonological strings and morphosyntactic feature structures into pairs, just as the *eci*-2 or *eci*-2pl[11] mentioned above. In compliance with optimality, there is no upper bound to the number of possible candidates – these are just the two most plausible ones that can be generated. It is then checked how a given affix hypothesis fares with regard to the paradigm cells thanks to an accuracy measurement: rather informally, a cell is a true positive for the features of a morphosyntactically compatible affix if the former's phonological string contain the latter's one (properly or not) – this is the case of any "perfect" affix, occurring in all and only the cells compatible with it. True negatives represent the exact reverse situation where an affix shares neither features nor phonological material with a cell. The more difficult facts, a subset of those involving *eci-*/*e-* above, are

[11] Henceforth, all features will be notated with square brackets and a binary specification: e.g. [+2+pl].

called false negatives/positives. In that example, the cells with a 1 subject and a 2sg object are false negatives for the affix hypothesis *eci*-[+2+pl]: they are not compatible with its features, still they contain the string /eci/. Conversely, assuming an affix *eci*-[+2], mispredictions arise in all cells exhibiting *e*-, since such an affix is so general that it should be present in them, but is not: these cases are then called false positives.

This is the core of the imperfect distribution problem, and to show the implications of it, Bank and Trommer (2012) develop a range of optimality-based analyses. A crucial aspect is that each constraint ranking corresponds to a different grammar. What I'll be most interested in here is the interaction of two constraints, *Underinsertion and *Overinsertion, which are most directly linked to the above remarks about *eci*-/*e*-. Below I illustrate briefly how the simplified paradigm in table 4 is treated in Bank and Trommer (2012). The first step consists in having the learner be "presented" with the paradigm and make hypotheses. The paradigm is shown again on the left, and some affix possibilities are parallelly shown on the right, with the two possible rankings.

S \ O	1sg	1pl	2sg	2pl	3sg	3pl
1sg			eci-	eci-		
1pl			eci-	eci-		
2sg					e-	e-
2pl	eci-	eci-			eci-	eci-
3sg			e-	eci-		
3pl			e-	eci-		

	*U	*O
☞-eci:[+2]		*6
-eci:[+2+pl]	*!2	

	*O	*U
-eci:[+2]	*!6	
☞-eci:[+2+pl]		*2

Table 5: Learning of *eci*-

What this says is that the affix hypothesis *eci*-[+2+pl] incurs two violations of the constraint *Underinsertion, because it is predicted not to occur in the two 1-2sg cells, but still does, and that the affix hypothesis *eci*-[+2] incurs six violations of *Overinsertion since it is predicted to occur in the four cells with a 2sg subject and the two with a third person subject and a 2sg object, but doesn't. The two tableaux shown in table 5 represent different grammars in that if the ranking *Underinsertion ≫ *Overinsertion is chosen, the *eci*-[+2] hypothesis is kicked out of the run by the learner in favor of *eci*-[+2+pl], while the reverse ranking yields the reverse result. This is how the meaning assignment problem is addressed. Both rankings automatically give rise to *primary distributions* of affixes: the primary distribution of *eci*-[+2+pl], for instance, is made of all cells where the features [+2+pl] are found. As mentioned, both affixes have imperfect distributions, because these are either larger or smaller than the distribution attested in the real language. I will return shortly to ways of coping with this problem in the discussion of what Bank and Trommer (2012) call *paradigmatic readjustment*.

I turn in the next subsection to the method for building the complete affix set of a given language, which is the concern of the harmonic-serialist side of the optimality account. In this I leave Ainu and get directly to the Georgian data. The problems of meaning assignment, imperfect distribution and paradigmatic readjustment will all be discussed in this subsection.

4.2. A case study of Georgian

First, it is necessary to decide what paradigm the learner works with. In the Georgian case, my assumptions dictate that all possible subject-object combinations in all considered tenses be taken into account. Although prefixation, as mentioned, does not vary with TAM categories, it does express person and case features, which are also relevant to suffixation. It is thus fully included in the paradigm, and the above drastic simplification from six to three forms within each TAM subparadigm is not pursued here. Table 6 is an instance of this use of a full-scale paradigm meant especially for inflectional learning.[12]

Incidentally, the fact that prefixes don't change across TAM categories allows for a very efficient illustration of the system. I briefly show how this works in the Present screeve (table 7), ignoring all others for the time being.

In table 7, according to the preceding descriptions, the prefixes stand intuitively for the following meanings. *v*-: 1 subject, *g*-: second person object, *m*-: 1sg object and *gv*-: 1pl object. The number of cells of the paradigm in table 7 is 28 (as in Ainu, reflexive forms are omitted; indeed, they require use of special anaphors and are not directly linked to the distribution of affixes).

[12] Another difference from the accounts presented above is that all strings corresponding to suffixal material are decomposed; concretely, this gives us sequences like -*n*-*n* for 3pl and -*e*-*s* for 3pl aorist, whose parts will have to be learnt separately.

Present [–Pa–P–S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…	v-…		g-…-t	v-…
2sg	m-…		…	gv-…		…
3sg	m-…-s	g-…-s	…-s	gv-…-s	g-…-t	…-s
1pl		g-…-t	v-…-t		g-…-t	v-…-t
2pl	m-…-t		…-t	gv-…-t		…-t
3pl	m-…-n	g-…-n	…-n	gv-…-n	g-…-n	…-n

Imperfect [+Pa–P–S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…-i	v-…-i		g-…-i-t	v-…-i
2sg	m-…-i		…-i	gv-…-i		…-i
3sg	m-…-a	g-…-a	…-a	gv-…-a	g-…-a-t	…-a
1pl		g-…-i-t	v-…-i-t		g-…-i-t	v-…-i-t
2pl	m-…-i-t		…-i-t	gv-…-i-t		…-i-t
3pl	m-…-n-n	g-…-n-n	…-n-n	gv-…-n-n	g-…-n-n	…-n-n

Subjunctive [–Pa–P+S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…-e	v-…-e		g-…-e-t	v-…-e
2sg	m-…-e		…-e	gv-…-e		…-e
3sg	m-…-e-s	g-…-e-s	…-e-s	gv-…-e-s	g-…-e-t	…-e-s
1pl		g-…-e-t	v-…-e-t		g-…-e-t	v-…-e-t
2pl	m-…-e-t		…-e-t	gv-…-e-t		…-e-t
3pl	m-…-n-n	g-…-n-n	…-n-n	gv-…-n-n	g-…-n-n	…-n-n

Aorist [+Pa+P–S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…-e	v-…-e		g-…-e-t	v-…-e
2sg	m-…-e		…-e	gv-…-e		…-e
3sg	m-…-a	g-…-a	…-a	gv-…-a	g-…-a-t	…-a
1pl		g-…-e-t	v-…-e-t		g-…-e-t	v-…-e-t
2pl	m-…-e-t		…-e-t	gv-…-e-t		…-e-t
3pl	m-…-e-s	g-…-e-s	…-e-s	gv-…-e-s	g-…-e-s	…-e-s

Optative [–Pa+P+S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…-o	v-…-o		g-…-o-t	v-…-o
2sg	m-…-o		…-o	gv-…-o		…-o
3sg	m-…-o-s	g-…-o-s	…-o-s	gv-…-o-s	g-…-o-t	…-o-s
1pl		g-…-o-t	v-…-o-t		g-…-o-t	v-…-o-t
2pl	m-…-o-t		…-o-t	gv-…-o-t		…-o-t
3pl	m-…-o-n	g-…-o-n	…-o-n	gv-…-o-n	g-…-o-n	…-o-n

Table 6: Full Georgian paradigm

Present [–Pa–P–S]

	1sg	2sg	3sg	1pl	2pl	3pl
1sg		g-…	v-…		g-…-t	v-…
2sg	m-…		…	gv-…		…
3sg	m-…-s	g-…-s	…-s	gv-…-s	g-…-t	…-s
1pl		g-…-t	v-…-t		g-…-t	v-…-t
2pl	m-…-t		…-t	gv-…-t		…-t
3pl	m-…-n	g-…-n	…-n	gv-…-n	g-…-n	…-n

Table 7: Full present paradigm

Cyc. *n*		*Und.	*Ov.
☞	gv-[Acc+1+pl]		
	gv-[Acc+1]		*4
☞	m-[Acc+1–pl]		
	m-[Acc+1]		*4
	g-[Acc+2]	*4	
	g-[+2]	*2	*6
	v-[Nom+1]-[Acc–2]	*4	
	v-[Nom+1]	*4	*4
	v-[+1]		*8

Table 8: Prefixes with *Und. > *Ov. ranking (Cyc. *n*)

4.2.1. Cyclic iteration and affix selection

Table 8 show how a learner can treat these prefixes. The ranking is tentatively *Underinsertion ≫ *Overinsertion.[13] More or less obvious possible meanings are indicated for each affix. *gv*-[Acc+1+pl], as opposed to *gv*-[Acc+1], is called a "perfect" marker because it doesn't incur any violation: the string

[13] Another constraint not shown here, *Portmanteaux, is violated whenever an affix hypothesis realizes more than one feature structure. Although never crucial to the evaluations presented here, where *Over/Underinsertion are always ranked higher, it nevertheless reminds that other rankings are possible which would exclude all such hypotheses, requiring special, dedicated morphological analyses. It is mentioned in the examples that will follow the prefixes' one.

Cyc. $n+1$		*Und.	*Ov.
☞	g-[Acc+2]		
	g-[+2]		*6
☞	v-[Nom+1]-[Acc–2]		
	v-[Nom+1]		*4
	v-[+1]		*8

Table 9: Prefixes with *Und. ≫ *Ov. ranking (Cyc. $n+1$)

gv- is present in all and only the cells with the features [Acc+1+pl], while its concurrent is overinserted (i.e. *gv*-[Acc+1+pl] wins under both rankings). The same can be said about *m*-[Acc+1–pl] as opposed to *m*-[Acc+1]. The situation is different in the case of the two remaining prefixes. The intuitively more correct meaning for *g*- is [Acc+2]: it occurs wherever there is a 2 object; thus it would seem that its primary distribution is a perfect one. But the model presented here has more to say about strings, since another string present in the paradigm, *gv*-, *has g*- as a substring. As a consequence, *g*-[Acc+2] incurs violations of *Underinsertion in all cells with a 1pl object. As for *v*-, *v*-[Nom+1]-[Acc–2] is more accurate than *v*-[Nom+1] since it doesn't incur the *Overinsertion violations linked to the absence of *v*- in contexts with a 2 object – truly a reflex of what morphological theories attempt to block. But again, *gv*- also has *v*- as a substring, which makes the two *v*- hypotheses also incur *Underinsertion violations.

A crucial aspect of the analysis exhibited by this small excerpt is that several affixes (in fact, all of them) compete together. In table 8, *gv*- and *m*- win together with zero violations. This is the serial component: after this first *cycle* is completed, step one, *freezing*, applies. Freezing "deletes" from the paradigm strings that are matched by winning affix hypotheses; thus *gv*-/*m*- are not considered anymore. Step two, cyclic iteration, initiates a new "round" of evaluation: the paradigm is considered again with removed strings, and evaluation of the candidate set is operated again, without the previous winners (table 9). The removal of *gv*- has the effect that the violations of *Underinsertion previously incurred by *g*- and *v*- disappear. The whole procedure has clear effects on affix selection. For instance, under the given ranking, an affix hypothesis *v*-[+1] would seem to win on the second cycle by only incurring violations of *Overinsertion (it would not be underinserted

		*Und.	*Ov.	*PM
☞	-i[Nom−3]-[+Pa−P]			*
	-i[Nom−3]		*64	
☞	-a[Nom+3−pl]-[+Pa]			*
	-a[Nom+3−pl]		*18	
	-s[Nom+3+pl]-[+P−S]	*15		*
	-s[Nom+3]		*39	
	...			

Table 10: Suffixes -*i* and -*a*

as it is coherent with the 1pl object cells, which contain *gv*-). However, its *Overinsertion violations would make it a "less perfect" marker than *gv*- and *m*- and by the time the second cycle is reached, *gv*- is gone, eliminating the *Underinsertion violations of the hypotheses *v*-[Nom+1] and *v*-[Nom+1]-[Acc−2]. Only because of this does the latter win, and not *v*-[+1].

Examination of each and every affix hypothesis under both rankings, taking into account the cyclic nature of the procedure, would take us at some length given the amount of data (it can be mentioned that -*o*, as could be expected, is also a "perfect marker" if it is given the meaning [+P+S]). Instead, I will present a couple of illustrative examples. I maintain the ranking *Underinsertion ≫ *Overinsertion here, plus lowest-ranked *Portmanteaux. It is useful to keep in mind, though, that no ranking is more "legitimate" than another. Of course, departing from the simple example of TAM-insensitive prefixes, figures must range over the whole paradigm and refer specifically to how affix hypotheses fare with regard to it (i.e. violations appear to be more numerous).

The vocalic suffixes -*i* and -*a* seen in the preceding accounts are among the easiest ones here too (table 10). -*i*-[Nom−3]-[+Pa+P] cannot be superseded whether the ranking is *Underinsertion ≫ *Overinsertion, or the reverse. The same holds for -*a*-[Nom+3−pl]-[+Pa]. Both indeed have a perfect distribution, their only disadvantage being that of violating *Portmanteaux.

The "-*n*-like" 3pl suffixes that occur in all subparadigms but the aorist are a further interesting example. This is shown in table 11, which represents all

Obj. / TAM	1sg	2sg	3sg	1pl	2pl	3pl
Pres. [−Pa−P−S]	m-…-*n*	g-…-*n*	…-*n*	gv-…-*n*	g-…-*n*	…-*n*
Imp. [+Pa−P−S]	m-…-*n*-n	g-…-*n*-n	…-*n*-n	gv-…-*n*-n	g-…-*n*-n	…-*n*-n
Subj. [−Pa−P+S]	m-…-*n*-n	g-…-*n*-n	…-*n*-n	gv-…-*n*-n	g-…-*n*-n	…-*n*-n
Aor. [+Pa+P−S]	m-…-e-s	g-…-e-s	…-e-s	gv-…-e-s	g-…-e-s	…-e-s
Opt. [−Pa+P+S]	m-…-o-*n*	g-…-o-*n*	…-o-*n*	gv-…-o-*n*	g-…-o-*n*	…-o-*n*

Table 11: Nominative 3pl subparadigm

and only the cells of the paradigm in table 6 which have a 3pl argument; it is somewhat uniform as properties of the object (plural *-t*, apart from prefixes) are never expressed. Facts are treated here in a way similar to the Channel analysis above in that 3pl exponence in the imperfect and subjunctive is decomposed into two distinct segments spelled [n]. This is shown in table 12, which features two distinct sets of *-n* candidates with parallel feature specifications, plus an affix *-nn*, a potential candidate for the imperfect and subjunctive tenses. Not so many choices are plausible for the simple *-n* affixes: either they range over all 3pl cells and incur *Overinsertion violations in the aorist, or they range over 3pl non-present imperfective ([−P]) cells and incur the same number of *Underinsertion violations in the optative ([+P+S]) cells. The two sets of candidates being parallel, one, the overinserted *-n*-[Nom+3+pl], wins and freezes one instance of a string /n/ in each relevant cell (shown in italics), leaving unfrozen *-n*'s only in the imperfect and the subjunctive. *-nn* candidates are discarded since the best faring candidate, *-nn*-[Nom+3+pl]-[−P], is overinserted in the present *and* is a portmanteau affix. Given such results, the following cycle (table 13) leads to the selection of a second isolated *-n* (no *-nn* hypothesis is plausible anymore because no such string is present in the paradigm), the violation figures being as before modified by the outcome of the preceding cycle. One, *-n*-[Nom+3+pl] is overinserted but suited for the most general case of 3pl exponence, while the other, *-n*-[Nom+3+pl]-[−P], also overinserted, suits to the imperfect and subjunctive tenses.

The third person *-s* and *-es* affixes are something of a puzzle for many theories. As a reminder, in the Distributed Morphology analysis presented above, 3pl *-es* was seen as a fully independent suffix realizing unequivocally the features "3pl aorist": as it blocks *-t* the same way "*-n*-like" suffixes do, it is easily thought

Cyc. *n*		*Und.	*Ov.	*PM
☞	-n_1[Nom+3+pl]		*6	
	-n_1[Nom+3+pl]-[−P]	*6		*
	-n_2[Nom+3+pl]		*6	
	-n_2[Nom+3+pl]-[−P]	*6		*
	-nn[Nom+3+pl]		*18	
	-nn[Nom+3+pl]-[−P]		*6	*

Table 12: 3pl -*n*-suffixes with *Und. ≫ *Ov. ranking (Cyc. *n*)

Cyc. $n+1$		*Und.	*Ov.	*PM
	-n_2[Nom+3+pl]		*18	
☞	-n_2[Nom+3+pl]-[−P]		*6	*

Table 13: 3pl -*n*-suffixes with *Und. ≫ *Ov. ranking (Cyc $n+1$)

to be an affix of this very "family" and to be unrelated to either -*e* or -*s* (3sg present, subjunctive and optative). A different analysis was offered in the Channel framework as -*e* was treated as an elsewhere while two distinct -*s* suffixes were postulated: one for 3sg, including the -*e*-*s* sequence found in the subjunctive, while the second -*s* was specifically marked for 3pl *and* perfective, allowing it, through specificity, to occur only in the aorist. I show here how letting a learner select the affixes can lead to different choices. Table 14 shows the competition between -*s*, -*e*, and -*es* and has, again, to be checked against the full paradigm in 6. It appears that the latter suffix is eliminated quickly whatever its features, because it always incurs violations of *Underinsertion. As for -*e*, only one of its most plausible hypotheses doesn't incur such violations, and this is the one with no specification at all and the greatest number of *Overinsertion violations – a pure elsewhere marker. Still, the winner of the first cycle is a general -*s*-[Nom+3] hypothesis; all others somehow draw a line between 3sg and 3pl and thus incur fatal *Underinsertion violations, while the winner only incurs *Overinsertion violations. The following cycle is vacuous since only -*e* remains as a candidate.

The preceding explanations thus mention how a given portion of the affixes of the language are selected in a way that is not directly guided by preferences of

		*U	*O	*PM
	-s[Nom+3–pl]-[–Pa]	*6	*3	*
	-s[Nom+3+pl]-[+P–S]	*15		*
	-s[Nom+3–pl]	*6	*15	
☞	-s[Nom+3]		*39	
	-e[+P–S]	*22	*6	
	-e[–P+S]	*22	*6	
	-e[–P]	*6	*6	
	-e[ ]		*96	
	-es[Nom+3+pl]-[+P+S]	*5		*
	-es[Nom+3+pl]	*5	*24	

Table 14: *-s/-e/-es* with *Und. ≫ *Ov. ranking

[

gv-	:	[Acc+1+pl]
m-	:	[Acc+1–pl]
-o	:	[+P+S]
g-	:	[Acc+2]
v-	:	[Nom+1]-[Acc–2]
-i	:	[+Pa–P]-[Nom–3]
-a	:	[–Pa]-[Nom+3–pl]
$-n_1$	:	[Nom+3+pl]
$-n_2$	:	[–P]-[Nom+3+pl]
-t	:	[–3+pl]
-s	:	[Nom+3]
-e	:	[ ]

]

Table 15: Pool of affixes after cyclic iteration (*Und. ≫ *Ov.)

the theoretician. The choice made by the latter may only be an initial one, a pool of affixes being afterwards incrementally constructed by cyclic iteration. These affixes are presented in table 15.

4.2.2. *Summary: rankings and morphological theories*

The preceding discussion was entirely dedicated to a learning procedure which leads to the creation of lists of affixes, as developed in detail in Bank and Trommer (2012). The list in table 15 is the outcome of ranking *Und. before *Ov.; reversing the ranking can give different results, which I won't delve into here (it is notable, for instance, that under the reverse ranking, *-s* would not be selected as a general third person affix, but rather *-es* would be first selected, as in many analyses, as a 3pl aorist affix, and only then would *-s* be chosen as a specific non-plural, non-past third person affix). Moreover, Bank and Trommer (2012) discuss a number of additional constraints which force a more fine-grained selection procedure.

Most importantly here, as mentioned in the introduction to this section, there is a persistent problem of *imperfect distributions*, to which the proposed answer lies in *paradigmatic readjustment*. Many affixes from the pool above are indeed imperfect; mostly, these are affixes with violations of the constraint *Overinsertion. As a consequence, affixes like n_1, *-s* or *-e*, given their specifications, should be in a lot of places in the paradigm where in fact, they are not. This is considered a normal consequence of the optimality model which doesn't seek perfection. Yet, some "post-optimality" means must ensure that the selected affixes stick to the real distributions.

The relevant tools are to be found in existing or yet to be invented morphological theories. The typology found in Bank and Trommer (2012) comprises two large families: the *retractionist* and *expansionist* theories. Retractionist theories are devised for sets of overinserted affixes, and vice-versa. That is, retractionist theories will include special tools to render application of affixation narrower when necessary, e.g., by manipulating the featural representation of paradigm cells. The latter case is characteristic of the impoverishment rules postulated in Distributed Morphology, which delete features. It is also the approach I have pursued in the short demonstration above, which featured exclusively the *Underinsertion ≫ *Overinsertion ranking. As a matter of fact, approximately half of the affixes selected in this way are in need of a retractionist treatment: all the perfect markers would fare equally well under both rankings, but the cases of textitn_1, n_2, *-t*, *-s*, and *-e* cannot fit directly the picture offered by the real language.

A first step in managing these overinserted affixes would be, as above with Distributed Morphology and Channel theory, to reflect upon the problem of position classes (or any approximation to this notion), since it cannot be the case that these affixes (peculiarly the suffixes) are randomly inserted in the verbal complex (e.g. both *-n*'s realize 3pl features and are spelled the same, so there is a need for two slots; the order of other consonantal and vocalic suffixes must be clearly determined; etc.). Bank and Trommer (2012) explain clearly that these issues are of a different nature than the study of how rankings are organized to create various pools of affixes. A look into options of paradigmatic readjustment related to this problem of position classes might then be an crucial side question.

5. Conclusion

In this paper I have tackled the morphology of Georgian affixes. I have first inspired myself from work like that of Halle, Morris and Alec Marantz (1993) or Anderson (1986), who among the first have described and analyzed puzzling facts in the distribution of prefixes/suffixes like the plural suffix *-t* as against other plural suffixes (blocking) or the competition between first person subject and second person object prefixes. Keeping track of the Distributed Morphology account of Halle, Morris and Alec Marantz (1993), I suggested an extension of it to a larger set of data, including all the tense-aspect-mood affixes found outside of the more complicated perfect series, which remains a challenge. The system of Distributed Morphology allows a treatment of most affixes, but still raises questions as to the status of some apparently unmotivated empty elements, that, moreover would seem to have to be extrinsically ordered before other items, disregarding specificity.

The Channel, or Accessibility-based, analysis proposed by Keine (2012) has also been introduced and considered as an option to possibly overcome shortcomings of Distributed Morphology (in this particular instance). The globally more attractive simplicity of the model, where it is an explicit part of the stipulations that affixes may follow some affixes but not others, is potentially tied to the notion of position classes, making the model apparently suited to the observed nature of Georgian suffixes, which respect a certain order. It is however difficult to see what, for instance, determines that elsewhere exponents, by definition empty of features, will stick to a given position class. Moreover, the accessibility relationships proposed in this paper are often very intricate and require massive manipulation of features, like negative specification or a complicated hierarchy, to allow otherwise impossible co-occurrence facts, among other problems. I do not conclude that the theory is bad (or the language), but rather that a separate examination of each of these problems should be carried out.

Finally, the learning algorithm proposed in Bank and Trommer (2012) has been exposed in relation to the same set of data. The fundamental difference between this last part and the two preceding ones was that no particular morphological analysis was at play. It was shown that it is possible to rely on predetermined parameters to establish lists of affixes, whereby intervention of the theoretician is somewhat confined to the initial choice of the constraint ranking. I have tentatively followed a ranking among many others, that lets the learner select affix hypotheses which are in many cases "overinserted" in the relavant paradigm. One interesting finding was that under this ranking, it is possible to derive without much difficulty the third person plural *n-n* (or *-nen*) of the imperfect and subjunctive, which are not natural classes in my system; above all, it was possible to show that the almost omnipresent third person *-s* can be generalized to third person plural aorist *-e-s*, instead of being confined to third person singular, thus revealing correspondence between a natural class "third person" and the exponent *-s*.

References

Anderson, Stephen (1986): 'Disjunctive Ordering in Inflectional Morphology', *Natural Language & Linguistic Theory* 4(1), 1–31.

Anderson, Stephen (1992): *A-Morphous Morphology*. Cambridge University Press, Cambridge, chapter Some Complex Inflectional Systems, pp. 137–179.

Bank, Sebastian and Jochen Trommer (2012): Inflectional Learning as Local Optimization. manuscript.

Béjar, Susana (2003): Phi-Syntax: A Theory of Agreement. PhD thesis, University of Toronto, Toronto.

Béjar, Susana (2011): 'Remarks on omnivory and complementarity: a commentary on the paper by Andrew Nevins', *Natural Language & Linguistic Theory* **29**(4), 973–997.

Carmack, Stanford (1997): 'Blocking in Georgian Verb Morphology', *Language* 73(2), 314–338.

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*. Cambridge MA: MIT Press, pp. 111–176.

Keine, Stefan (2012): 'Syntagmatic Constraints on Insertion', *Special Issue of "Morphology"* .

Lomashvili, Leila (2011): 'Phases and templates in Georgian agreement', *Studia Linguistica* **65**(3), 233–267.

Müller, Gereon (2004): On Decomposing Inflection Class Features: Syncretism in Russian Noun Inflection. *In:* G. Müller, L. Gunkel and G. Zifonun, eds, *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 189–277.

Müller, Gereon (2007): . Extended Exponence by Enrichment: Argument Encoding in German, Archi, and Timucua. *In:* T. Scheffler, J. Tauberer, A. Eilam and L. Mayol, eds, *: Proceedings of the 30th Annual Penn Linguistics Colloquium*. University of Pennsylvania, Penn Working Papers in Linguistics 13.1, Philadelphia, pp. 253–266.

Noyer, Rolf (1997): *Features, Positions and Affixes in Autonomous Morphological Structure*. Garland Publishing, New York. Revised version of 1992 MIT Doctoral Dissertation.

Stump, Gregory (2001): *Inflectional morphology: A theory of paradigm structure*. Cambridge University Press, Cambridge.

Tschenkéli, Kita (1958): *Einführung in die georgische Sprache. Zwei Bände*. Amirani Verlag, Zürich.