SUFFIXAUFNAHME

MARCUS KRACHT
II. MATHEMATISCHES INSTITUT
ARNIMALLEE 3
D – 14195 BERLIN
KRACHT@MATH.FU-BERLIN.DE
HTTP://WWW.MATH.FU-BERLIN.DE/~KRACHT

ABSTRACT. Suffixaufnahme is a peculiar phenomenon of languages that has been discovered in the mid 19th century by Franz Bopp already, but which has only recently attracted attention within formal language theory. Essentially, what it amounts to is the intrusion of the recursion found in syntax into the morphology. This makes suffixaufnahme an interesting test bed for formal theories of language. For example, it can be shown that many syntactic formalisms (including the Minimalist Program as defined by Stabler) are too weak to generate languages with full suffixaufnahme. Though suffixaufnahme in its extreme form is perhaps limited to a handful of languages, there are a lot of phenomena even in European languages that involve suffixaufnahme to a considerable extent. Thus, suffixaufnahme is not an oddity; and it shows us which formalisms stand a chance of being adequate for human language and which ones do not.

1. HISTORY AND INTRODUCTION [1]

1.1. **Old Georgian.** In 1842, Franz Bopp in an address to the Academy of Sciences of Berlin noted the following curiosity about Old Georgian.

(1)	gwam-isa	krist-es-isa
	body-GEN	christ-GEN-GEN
	of the body of Christ	
(2)	c̣q̇oba-sa	mṭer-ta-sa
	attack-DAT	enemy-OBL.PL-DAT
	of the attack of the enemies	
(3)	qeli-ta	mocikul-ta-ta
	hand-OBL.PL	apostle-OBL.PL-OBL.PL
	through the hands of the apostles	
(4)	c̣inamsrbol-n-i	laškar-ta-n-i
	forerunner-PL-NOM	army-OBL.PL-PL-NOM
	the vanguard of the armies	

What we find is that the possessed NP additionally carries the case (and number) inflection of the possessor. Bopp analysed this as an adjectivized noun agreeing like an adjective in case and number. Thus, to Bopp Old Georgian looked like an Indo–European language. Finck, upon analysing the following construction was led to conclude that what we find is a repetition of the suffix of the higher head and he coined the name 'suffixaufnahme' (*taking up of suffixes*).

[1] This section is heavily based on Plank [29].

(5)	ra	turṗa	ṗrinvelia	c̣amoiʒaxa	erṭ-ma	bavšv-ṭa-gan-ma
	what	wonderful	bird-is,	exclaimed	one-ERG	child-OBL.PL-ABL-ERG

"What a wonderful bird!" exclaimed one of the children

Finck introduced eight types of language structure, one of which was the group inflecting type, to which he counted Old Georgian. Group inflecting means that every member of a phrase inflects for the same grammatical categories. In this he saw actually no difference with Greek or Latin. Once again, Old Georgian looked like an ordinary Indo–European language. Incidentally, Finck's classification was based mainly on morphological criteria.

1.2. **Hurrian.** In 1905, Ferdinand Bork already pointed at the fact that the phenomenon involved other Caucasian languages (Tsahkur (a Lezgian language), Bats (North–Eastern Caucasian)), and moreover could be found in some ancient languages of the Middle-East, such as the Hurrian–Urartian group and an isolate language, Elamite. For Hurrian this was noted already by Peter Jensen and Leopold Messerschmidt in 1899. Here is an example of three cases (plus extra suffixes) in a row. This example is from Hurrian texts of Boğazköy, quoted from [35].

(6) unni=ma ^{D}Teššub=va šarri=ne=va evre=n(i)=ne=na ...
en(i)=na=až=(v)a eğli=ve=NE=ve=NA=až=(v)a
šubri=ve=NE=ve=NA=až=(v)a ... un=a
now=PRT Teššub-DAT king-ART.SG-DAT lord-INDIV-ART.SG-DAT ...
god-ART.PL-PL-DAT salvation-GEN-CARR:SG-GEN-CARR:PL-PL-DAT
šubri-GEN-CARR.SG-GEN-CARR:PL-PL-DAT ... come-INTRNS
Now he comes to Teššub, to the king, to the lord ..., to the gods of the saviour (lit. of the one of salvation), of the one of the šubri ...

There are other suffixes in addition to case markers. The sequence falls naturally into a sequence of what we call *shells*. Here is a division of a word into shells:

šubri	=ve	=NE=ve	=NA=až=(v)a
šubri	-GEN	-CARR:SG-GEN	-CARR:PL-PL-DAT
root	1	2	3

We find three shells, each ended by a case suffix, containing some plurality markers and some extra suffixes. The translation does not reveal where to associate the plural markers. Typically, number marking is inside case marking. This is the case in Hurrian, too. Evidence that the plural marker must be put into the same shell with the following case is given for example by

(7)	en(i)=n(a)=až=už	attani=ve=n(a)=až=už
	god-RLT:PL-PL-ERG	father-GEN-RLT:PL-PL-ERG

the gods of the father

The plural is in the shell of the ergative, not that of the genitive, since it follows the genitive. In fact, Diakonoff and Starostin have argued in [7] that there is a genetic relationship between Proto–Urarto–Hurrian and North–East Caucasian, but that view is not universally accepted.

It was Bork and Heinrich Winkler who gave a different interpretation of the facts. According to them, the stacked suffixes show the syntactic affiliation of the words. This is, in a nutshell, what is believed today as well. This constituted a turning point. Clearly, the possessum–as–adjective story did no longer work for the newly found languages. In 1909, Henri Bourgeois argued also that Old Georgian was different. He claimed that it was in fact agglutinative, not inflecting, and that it lacked an inherent distinction between adjectives and nouns. Cases made roots into nouns, which could be turned into attributes. Only when they followed the higher NP, they indicated their relation with it being resuming the suffixes. So, suffixaufnahme was in fact order sensitive. (However, as Boeder [5] shows, this is contrary to fact.) Bourgeois believed that it was not genuine to Georgian but arose under the influence from neighbouring Greek and Armenian, which showed adjectival agreement.

(8) hai thugatéres hai tōn
the.NOM.PL.FEM daughter-NOM.PL the.NOM.PL.FEM the.GEN.MASC.PL
polit-ōn
citizen-GEN.PL
the daughters of the citizens

1.3. **Kayardild.** Igor Mel'čuk, in an attempt to clarify the situation, drew attention to the fact that case on adjectives is different from case on nouns. The latter was *casus rectus* (governed case) while the former was *casus concordatus* (agreement case). He thus assimilated the Old Georgian genitive once again to agreement case, though admitting that syntactically the possessed NPs were still nouns. It led him to propose a number of different types of cases, which however formwise were mostly identical to each other. He was also the first to bring Australian languages such as Ngarluma and Kayardild into the spotlight. (See [24].)

In fact, Mel'čuk got his data from Nick Evans, who together with Alan Dench has extensively researched the phenomenon of case stacking in Australian languages (see [6]). In Australian languages there is typically not much in terms of agreement, there are only cases. Case stacking is in some languages an exuberant phenomenon. Here is an example from Kayardild, showing up to four cases in a row (see [9]).

(9) Maku-ntha yulawu-jarra-ntha yakuri-naa-ntha
woman-OBL catch-PAST-OBL fish-MABL-OBL
dangka-karra-nguni-naa-ntha mijil-nguni-naa-ntha
man-GEN-INST-MABL-OBL net-INST-ABL-OBL
'The woman must have caught fish with the man's net.'

(Here, OBL stands for *oblique* and MABL for *modal ablative*.) The oblique signals some aspect of the sentence. Notice that the verb also carries a case marker for OBL. This may be taken as a sign that these are not really case suffixes, but this is a matter of theoretical decision. We follow here the established tradition and call them cases. The complement of the verb is marked by the ablative, and additionally it shows case concord with the subject and the verb. The instrument however also takes the modal ablative as well as the oblique, and finally the possessor takes all these three suffixes and the genitive.

2. What is happening here?

2.1. **The combinatorics of marking.** Intuitively, the matter is as follows. Suppose that X assigns instrumental to its sister constituent, [Y Z]. And suppose that Y assigns genitive to Z.

$$\begin{array}{ccccc} & \text{INST} & & \text{GEN} & \\ & \curvearrowright & & \curvearrowright & \\ [\mathsf{X} & & [\mathsf{Y} & & \mathsf{Z}]] \end{array}$$

Then two basic choices exist.

(1) The case marker is found only once at the phrase, typically right peripherally. Case is a phrasal affix. These languages are called **group marking**. (Turkish, Japanese, Hungarian)
(2) The case marker is iterated at every word of the phrase. Case is a **word affix**. Languages of this kind are called **word marking**. (German, Latin, Finnish)

Let us review (1) first. If additionally case assignment is to the left, we get

[Z-GEN Y]-INST X

This pattern we find in Turkish (see [16]). Turkish is consistently head final and suffixing.

(10) Hasan Ali-nin araba-sın-ı yak-mış
Hasan Ali-GEN car-POSS:3.SG-ACC burn-REP.PAST
Hasan is said to have burned Ali's car.

Now suppose that case assignment is to the right. Then two subpatterns arise. (a) The case marker is peripheral. Then we get the following structure, where a single word carries two case suffixes.

X [Y Z-INST]-GEN

Thus, the phenomenon of **double case** is technically distinct from **suffixaufnahme**. Such a language is Sumerian (see [29]).

(11) é lugal-ak
house king-GEN
house of the king
(12) é lugal-ak-a
house king-GEN-LOC
in the house of the king
(13) é šeš lugal-ak-ak-a
house brother king-GEN-GEN-LOC
in the house of the brother of the king

Here, the case is a phrasal suffix, and the genitive complement also follows the head noun. Hence the last example is to be bracketed as follows.

(é (šeš (lugal)-ak)-ak)-a

We find in the abovementioned source also examples from Late Elamite and Kanuri (Nilo–Saharan).

(b) The case marker attaches to the head. Then we get

X [Y-INST Z-GEN]

Finally, we turn to (2). We assume for simplicity head initial order, with suffixing case morphology. This scenario leads to case agreement within a phrase.

X [Y-INST Z-GEN-INST]

Notice the difference in bracketing: we do not view the case markers as phrasal affixes (pace Mel'čuk, who would probably distinguish the second occurrence of the instrumental marker as being different from the first)). Kayardild is an example of a language where most case markers survive. It also offers free word order, so that one cannot speak of that language as a head initial or head final language at all. Thus, suffixaufnahme in its regular form is nothing but the consistent execution of a marking regime that spells out the case on every part of the phrase rather than once. Notice that the order of the suffixes is always inside out. The closer the suffix is to the root, the smaller the constituent that it reflects. The only known exception to this rule (again involving Kayardild) is discussed in Evans [10].

Mostly, morphology forbids stacking one case on top of another (eg in Indo–European languages). In those languages the innermost case survives, leading to

X [Y-INST Z-GEN]

This is the basic pattern of Indo–European languages. (However, exceptions to this pattern exist (see below).)

(14) im Palast des Königs
IN.DAT palace DEF.GEN.MASC.SG king-GEN
dans le palais du roi
domō rēgis
in the palace of the king

Some variants exist. For example, in German, genitives may be preposed (in des Königs Palast) but this already sounds a bit archaic. The identity with this pattern and the one discussed under 1(b) is only superficial, since we have not displayed agreement within the NP.

3. Scope of the Phenomenon

3.1. **Double Case.** We say a language has **double case** the moment that it allows for a word to carry two case markers. Suffixaufnahme is not to be equated with double case. As we have seen, agreement markers can (and sometimes must) be iterated in addition to case.

It is clear that inflecting languages disfavour suffixaufnahme. If suffixaufnahme is nothing but agreement then there are many phenomena that may be and have been subsumed under it. One phenomenon is agreeing postpositions in Hindi, and possessor/possessum agreement in general. We give examples from Finnish and Hungarian.

(15) minun auto-ssa-ni
1.SG.GEN car-INE-POSS:1.SG

(16) az én kocsi-m-ban
DET 1.SG car-POSS:1.SG-INE
in my car

Both languages use head marking to show agreement with the possessor, which may or may not show up, in Finnish in the genitive, in Hungarian in the dative or nominative (for personal pronouns). Moreover, in Hungarian the possessor agreement is obligatory.

3.2. **Case Attraction.** Although Indo–European languages allow only one case per word, there are phenomena that show sensitivity to the case one level up. The phenomenon is referred to as **case attraction**. Here is an example from Armenian.

(17)	i	knoǰ-ê t'agawor-i-n
	by wife-ABL.SG	king-GEN.SG-DEF
(18)	i	knoǰ-ê t'agawor-ē-n
	by wife-ABL.SG	king-ABL.SG-DEF
	by the wife of the king	

Other examples are found in Russian and Classical Greek. It was Hans Vogt in 1906 who drew attention to the parallel between case attraction and suffixaufnahme. There are other constructions that show similar effects, for example free relatives. The relative pronoun shows sensitivity both to the case assigned inside the relative clause (which is to be expected) but also to the case assigned to higher (but empty) head which the relative clause modifies (see Vogel [**?**] for a discussion). Notice also that in Hurrian suffixaufnahme is found especially when a head has been omitted, showing that the need to express two cases may either lead to suffixaufnahme (Hurrian) or in absence of this possibility to a kind of case conflict which more or less ends in the requirement that the two cases are identical (but not quite, see [**?**]).

3.3. **Failure of Word Marking.** As is clear from the data in Boeder, it is mostly nouns that show suffixaufnahme in Old Georgian. Determiners on the other hand do not. They consistently show only one case. This is reminiscent of Indo–European. Adnominal modifiers do not always show agreement. Adverbs and PPs form exceptions. So, unless we stipulate underlying (abstract, invisible) agreement, the plain rule whereby every member of a phrase agrees with every other is false even if it contains no embedded phrase.

It should also be noted that the question whether we have word marking or not depends on the way in which the case is expressed. I have argued elsewhere ([19]) that PPs are syntactically no different from case marked DPs, and to define cases as morphologically bound affixes is to fail to distinguish between case morphology and syntactic function. However, we know of no rule of phrasal concord for cases expressed by propositions. So, the typical scenario is that the preposition assigns case to its DP argument, and this case is spread over the entire DP without a concomitant spreading of the P head. Thus, agreement (and therefore suffixaufnahme) are limited to morphological affixes, and this may explain why in one language sometimes there is agreement and sometimes there isn't. (One may count this as an argument against my analysis.)

3.4. **Conditions on Case Sequences.** In their study of case stacking in Australian languages, Dench and Evans looked at the sequence of cases that are allowed to appear. It turns out that not only do we find portmanteau morphs etc, there are also peculiar restrictions on the sequences of cases (case markers). A particularly frequent condition is that adjacent cases may not be identical. Though this retrsiction is not universal it holds in other contexts too. Within Optimality Theory, Tara Mohanan ([26]) has proposed a constraint that forbids two successive NPs bearing identical case. In Finnish, this restriction applies to the partitive and the elative. It turns out that, besides other conditions, in iterated partitive constructions, the partitive case must alternate with the elative case (data from Anttila and Fong [3]).

(19) Sointu-sta tul-i munkki.
Sointu-ELA become-PAST monk
Sointu became a monk.

(20) Kolmasosa-sta $\left\{\begin{array}{l}\text{*miehistä}\\ \text{mieh-i-ä}\end{array}\right\}$ tul-i munkke-j-a.
one third-ELA $\left\{\begin{array}{l}\text{man-PL-ELA}\\ \text{man-PL-PART}\end{array}\right\}$ become-PAST monk-PL-PART
One third of the men became monks.

(21) Kolmasosa $\left\{\begin{array}{l}\text{mieh-i-stä}\\ \text{??mieh-i-ä}\end{array}\right\}$ ryhty-i munki-ksi.
one third $\left\{\begin{array}{l}\text{man-PL-ELA}\\ \text{man-PL-PART}\end{array}\right\}$ choose.to.be-PAST monk-TRANS
One third of the men chose to be monks.

Once again, Old Georgian proves to be a hard nut: iteration of the genitive is not ruled out but often encourages haplology (see [5]).

4. Semilinearity and Complexity

Let us look closer at suffixaufnahme. Let us sketch a very simple language, WME (**word marking English**). It has a verb run, a noun son, and a case marker gen. The verb assigns nominative to its subject, which is however empty. A noun may optionally take a genitive to its right, which is marked by the case suffix gen. Marking is on each member of the NP. So, the following sequences are licensed.

(22) son run.
son songen run.
son songen songengen run.
son songen songengen songengengen run.

Now, given that these are the only grammatical sentences, how many occurrences of the elements do we find in a sentence? Let's count them.

run	1	1	1	1	1	1
son	1	2	3	4	5	n
gen	0	1	3	6	10	$n(n-1)/2 = \binom{n}{2}$

What we find is that while the number of verbs is constant, the number of genitives grows quadratically with the number of nouns. This seemingly innocent fact has far reaching consequences, as we shall see. Notice by the way that we are counting not occurrences of words but occurrences of morphemes.

Let us associate with a sequence of elements a vector of numbers in the following way. The vector $\langle n_1, n_2, n_3\rangle$ is associated to a sequence if it contains run exactly n_1 times, son n_2 times, and gen n_3 times. The vector tells us how often an item occurs but not where. We denote by $\pi(\vec{x})$ the vector associated with the string $\vec{x}$. π is known as the **Parikh map**. A formal definition is as follows.

Definition 1. *Let $A = \{\mathsf{a}_i : i < n\}$ be an alphabet. Further, let $\vec{e}_i \in \mathbb{N}^n$ be the n the unit vector. Then let*

$$\begin{array}{rcl} \pi(\mathsf{a}_i) & := & \vec{e}_i \\ \pi(\vec{x}^\frown\vec{y}) & := & \pi(\vec{x}) + \pi(\vec{y}) \end{array}$$

This map is called the ***Parikh map****. It is unique up to the enumeration of A.*

The Parikh–images of the sentences in (22) are $\langle 1,1,0\rangle$, $\langle 1,2,1\rangle$, $\langle 1,3,3\rangle$, $\langle 1,4,6\rangle$ and $\langle 1,5,10\rangle$. In the Minimalist Program we find these vectors under the name *numeration.* Thus, if $\vec{x}$ is a string $\pi(\vec{x})$ is its numeration in the terminology of the MP. The vectors form an abelian semigroup, which is isomorphic to $\mathbb{N}^n$ for some n. In our case, $n = 3$; $\mathbb{N}$ is the set of natural numbers, including zero. Obviously, counting occurrences leads to a drastic simplification of the language. Parikh's Theorem will give evidence. Take a context free grammar for a language. It has rules of the form $X \to \vec{\alpha}$ for nonterminals X and sequences $\vec{\alpha}$. Collect all productions that have X on their left. Using the $\mid$ symbol we can collapse these productions as follows.

$$(\dagger) \qquad X \to \vec{\alpha}_1 \mid \vec{\alpha}_2 \mid \vec{\alpha}_3 \mid \ldots \mid \vec{\alpha}_p$$

Now, let $[X]$ be the set of all strings that form a constituent of category X. The map $[-]$ can be extended to sequences of nonterminals and terminals.

$$\begin{array}{rcl} [a] & := & \{a\} \qquad\qquad (a \in A) \\ [\vec{\alpha}\vec{\beta}] & := & [\vec{\alpha}] \odot [\vec{\beta}] := \{\vec{x}\vec{y} : \vec{x} \in [\vec{\alpha}], \vec{y} \in [\vec{\beta}]\} \end{array}$$

Then $(\dagger)$ is equivalent to the following equation:

$$(\ddagger) \qquad [X] = [\vec{\alpha}_1] \cup [\vec{\alpha}_2] \cup \ldots \cup [\vec{\alpha}_p]$$

For example, with $S \to AB \mid C\mathtt{x}$ we have the equation $[S] = ([A] \odot [B]) \cup ([C] \odot \{\mathtt{x}\})$. This is to say: a string is an S–string if and only if it possesses a decomposition $\vec{x} = \vec{y}\vec{z}$ where $\vec{y}$ is an A–string and $\vec{z}$ a B–string, or else $\vec{y}$ is a C–string and $\vec{z} = \mathtt{x}$. It is convenient to denote $[X]$ by the nonterminal X (or a suitable different variable).

Now we shall apply the Parikh map to this equation. Let $\vec{u}_i$ be the Parikh–image of $\vec{\alpha}_i$. Write $A \oplus B := \{\vec{u} + \vec{v} : \vec{u} \in A, \vec{v} \in B\}$. Then $(\ddagger)$ becomes

$$(\S) \qquad X = \vec{u}_1 \oplus \vec{u}_2 \oplus \vec{u}_3 \oplus \ldots \oplus \vec{u}_p$$

Notice that for every context free grammar there is a weakly equivalent grammar in which for every production $X \to veca\ \vec{\alpha}$ contains X at most once. (This is not hard to show.) Given this, the general form of these equations is $X = A \cup (B \oplus X)$ or $X = A$, where A and B do not contain X. While the latter provides a direct solution for the (unknown) X, the former still has to be solved. Let $\mathbb{N}B := \{k \cdot \vec{x} : k \in \mathbb{N}, \vec{x} \in B\}$.

Lemma 2. *The equation $X = A \cup (B \oplus X)$, where A and B do not contain occurrences of X, has as its least solution $X = A \oplus \mathbb{N}B$.*

Call a set $U \subseteq \mathbb{N}^n$ **linear** if there are $\vec{u}, \vec{v}_i \in \mathbb{N}$, $1 \leq i \leq \alpha$, such that

$$U = \vec{u} \oplus \mathbb{N}\vec{v}_1 \oplus \mathbb{N}\vec{v}_2 \oplus \ldots \oplus \mathbb{N}\vec{v}_\alpha$$

Call U **semilinear** if it is a finite union of linear sets.

Definition 3. *A language is* ***semilinear*** *if its Parikh–image is semilinear. Languages $L, L' \subseteq A^*$ are* ***letter equivalent*** *if and only if they have the same image under the Parikh–map.*

Theorem 4 (Parikh)**.** *The following holds.*

(1) *Regular languages are semilinear.*
(2) *Every context free languages is letter equivalent to some regular language.*
(3) *Context free languages are semilinear.*

WME, however, is not semilinear. Old Georgian has a homomorphic image that is letter equivalent to WME. Hence we have

Theorem 5 (Michaelis & Kracht)**.** *Old Georgian is not semilinear. Hence it is not context free.*

This already shows us that there is no GPSG grammar of Old Georgian. However, there is more in stock. Almerindo Ojeda [28] has shown that if one keeps the constituency but allows for nodes to be unordered with respect to each other, GPSG can generate Swiss German, which is known not to be context free. It is now immediately clear that such relaxation will not do for Old Georgian: no matter how we define ordering principles, if the constituency skeleton remains the same, the languages generated are semilinear.

In response to the proof that Swiss German is not context free, Joshi and others have argued for somewhat stronger grammar formalisms, which became known as Linear Context Free Rewrite Systems (LCFRSs, see [32]) on the one hand, and Multicomponent Tree Adjoining Grammars (MCTAGs) on the other. These formalisms define the same class of languages, which are called **LCFRSL**s or **MCTAL**s, depending on taste. LCFRSs are like context free grammars, but constituents are k–tuples of strings, where k is a number which is fixed for every individual LCFRS. Thus, for given k, one speaks of a k–LCRFS if constituents are at most k–tuples. As regards strong generative capacity, a k–LCFRSs generates a language in which constituents are broken into at most k continuous segments. The Linear Indexed Grammars of Gerald Gazdar and the Combinatory Categorial Grammars of Mark Steedman are weakly equivalent to 2–LCFRSs. Thus everything that we say about LCRFLs will apply to them, too.

Theorem 6 (Vijay–Shanker)**.** ***LCFRL**s and **MCTAL**s are semilinear.*

The proof is a straightforward extension of the original idea by Parikh. Somewhat surprisingly, it turned out that the Minimalist Grammars in the sense of Stabler [34] can be mimicked rather closely by LCFRSs. This led to the following result, independently established by Jens Michaelis ([25]) and Henk Harkema ([13]).

Theorem 7 (Michaelis, Harkema)**.** *The languages that can be generated by Minimalist Grammars in the sense of Stabler are exactly the LCFRLs.*

As a corollary we obtain

Corollary 8. *Old Georgian is not a LCFRL, nor a MCTAL. Neither can it be generated by a Minimalist Grammar in the sense of Stabler.*

It might be deemed that WME and languages with fully fledged suffixaufnahme are complex. But this is not so. To the contrary, they are very easy. Take once again WME. It is perfectly easy to parse this language. Moreover, suppose that we change the syntax in such a way that any sentence letter equivalent to a WME sentence is grammatical sentence (this is the permutation closure of WME). Call this language WME^c. In other words, $\mathrm{WME}^c = \pi^{-1}[\pi[\mathrm{WME}]]$. Moreover, we assume that if $\vec{x}$ is letter equivalent to $\vec{y}$, $\vec{y}$ and $\vec{x}$ have the same meaning. So, the following are sentences of WNE^c, and all have the same meaning.

songengen run son songen., run songen son songengen., son songen run songengen.

Theorem 9 (Ebert & Kracht). *It takes time $O(n^{3/2} \log n)$ on a 3–tape Turing machine to parse a language with full suffixaufnahme but free word order. Moreover, these languages are uniquely readable.*

Thus the complexity is slightly more than linear and definitely better than the best known for context free languages (which has exponent $\log_2 7$, which is greater than 2.5). The reason is easy to spot: the case markers allow for an easy reconstruction of the structure, and the structure in turn determines the meaning. If we eliminate the case markers, massive ambiguity will arise. It is this fact that lies behind the idea that case stacks are substitutes of structure. An investigation of this idea can be found in [20]. Notice that this provides theoretical evidence against the typology of Baker and others whereby free word order arises only in head marking languages. Empirical evidence comes from languages such as Jiwarli (Austin), which has completely free word order but is consistently dependent marking. (See also Nordlinger [27].)

The consequences of all these facts are far reaching. Not only does this demonstrate that the MP is not as mighty as it seems at first sight, there also is something distinctly odd about it: case and agreement features are the driving force of the MP, so one would expect that it has a good story to tell about them. However, it does seem that the mechanism of feature checking, because it displayed resource sensitivity in much the same way as other frameworks, was on the wrong track. This calls for a deeper analysis of the nature of features, to which we now turn.

5. Syntactic and Semantic Analysis

Up until the mid 20th century linguistics has concentrated on morphology, and suffixaufnahme has been seen mainly as a morphological feature. The emergence of transformational grammar has shifted the attention from morphology to syntax. While Finck's classification of languages was purely morphological, nowadays the main criteria by which languages are distinguished are syntactic (eg word order). Morphology was just a language particular means to introduce certain features. However, although it is not unplausible to do that, it should be noticed that case can only be a property of a phrase, not a word, but morphologically it is associated with a word. This creates a tension that makes the integration of morphology and syntax so difficult.

Before we continue, we have to get clear about the nature of features in syntactic theory. In linguistics, it is customary to assume that case, number and class are *features*, just like phonological features. This comes out clearest in GPSG. Here, accusative case is nothing but a boolean constant, which is interpreted as a set of nonterminal labels. Thus, it is a property of constituents that either is true of a given constituent or is not. Latin poetam has accusative, consulibus does not have it. Although it makes little sense to speak of the case of an adverb, the boolean view is perfectly consistent; features are never undefined. Thus, the adverb cras (*tomorrow*), for example, is (under this view) not accusative. Actually, GPSG employs the notation of attribute value logic, so that rather than writing NP[acc] we write

$$\begin{bmatrix} \text{CAT} & : & n \\ \text{BAR} & : & 2 \\ \text{CASE} & : & acc \end{bmatrix}$$

However, as long as the recursion was basically nonexistent, there was no difference between attribute value notation and plain boolean notation, except that the former allowed a more concise statement of feature (co)occurrence restrictions. GPSG explained agreement by feature sharing. This idea is constitutive for unification based formalisms (UCG, CUG, HPSG, LFG) and is therefore employed for agreement, too. GPSG had feature percolation rules that basically constrained the way features coexist in a labelled tree.

Transformational grammar oscillates between different views. On the one hand, features are often considered very much the same way as in GPSG. They are seen as properties of nonterminal labels, or constituents, whichever. The idea that representations are in te mind makes all the difference, though. The minute they are assumed to be in the human mind, the feature itself turns into something tangible, just like the ink spots that one finds on this paper to denote them. Essentially, a feature (occurrence) is a thing that is manipulated. In this respect it is much like the case ending itself: (a) it can be there, or (b) a different case ending can be there, or else (c) no ending can be there. Likewise, a feature can be present positively, or negatively, or it can simply be absent. That a feature, say accusative, is absent does no longer mean that the concrete item fails to be morphologically accusative; only the computational procedure has 'forgotten' about it.

We distinguish the two views as follows. The first we dub the **features as properties**, the second **features as substance** view. In this respect it is helpful to recall the distinction made by Mel'čuk between casus rectus and casus concordatus. The need for distinguishing between these two only arises under the features as substance view. The case, being assigned by the higher head, cannot be multiplied at will. It is assigned only once. However much sense this makes historically, for a synchronic description it is not helpful and tends to obscure the matter. The analogy between proof theory and semantics may help here. Suppose you have proved that p is the case. Even if you write it down only once, nothing forbids you from using it several times over. Nevertheless, proof theory does use resource sensitive constructions (discharge of assumptions, for example). However, nothing of substance follows from this. This being said we turn to actual analyses.

Boeder in [5] develops a transformational (pre–GB) account. Case suffixes of Old Georgian are clitics. When a constituent moves out of a higher constituent, it picks up the suffix of the higher constituent.

$$[\mathsf{X}\text{-}\textsc{gen}\ \mathsf{Y}]\text{-}\textsc{gen} \quad\leadsto\quad [[t\ \mathsf{Y}]\text{-}\textsc{gen}\ \mathsf{X}\text{-}\textsc{gen}\text{-}\textsc{gen}]$$

Repeated application of this operation yields several suffixes in a row. The reason why determiners do not show suffixaufnahme is readily explained by the fact that they are not moved alone, only within a phrase. Movement picks up only one copy of the clitic. Boeder asserts that suffixaufnahme involves restructuring but not necessarily a change in order (so, movement may be rightward or leftward). A different account has been given by Ken Hale on that matter ([12]). It is worth quoting the original passage.

> I assume, speculatively, that this [ie agreement in case in split constituents, MK] is accomplished by a rule of concord which marks each constituent of a noun phrase with an abstract case feature appropriate to the case category

> of the noun phrase as a whole. Whether the actual case ending appears once or repeatedly depends upon whether the noun phrase constituents, at the time the features are given phonological shape, are dominated by a common NP-node – if they are, then the case will be spelled out once, on the final constituent of the noun phrase; but if they are not, the case will be spelled out separately on each of the constituents.

Thus, Hale differs from Boeder in that he employs abstract features (not simply clitics). This view has advantages when it comes to morphological restrictions on case sequences. If it was just a matter of picking up a clitic it is hard to understand why such restrictions are in force. On the other hand, the Finnish data given above shows that such restrictions do exist even when morphology is not involved. Another difference is that Boeder is actually inserting phonological material during the derivation while Hale (my exegesis) takes the sequences of cases to be there underlyingly and deletes them when appropriate. What is left unclear, however, is the nature of features and what happens if we have several conflicting features.

Actually, at this point it is worth looking at the Minimalist Program. In the MP, the features as substance view has completely taken over. In accordance with categorial grammar, agreement features are viewed as resources that the computational system consumes during its computation. The standard version of the MP assumes that a feature is a single ticket: once the element has taken a ride with that ticket, it is no longer valid. A further point of notice is the geometry of features. While the original writings by Chomsky remain unclear about the way features constitute the feature part of the lexical entry, Stabler aligns them in a sequence. Otherwise, one cannot distinguish between different occurrences of the same feature. To wit, there are different possibilities (for example, Anderson [1]), but essentially without linear or hierarchical organisation nothing works. The notion of a flat phonological matrix (which Chomsky seems to presuppose) is unworkable. Stabler assumes that the features are linearly ordered, and that checking is sensitive to the ordering. Two features are responsible for the result that his version of the MP is semilinear and therefore forbids the suffixaufnahme languages. The first is the restriction that a feature that triggers raising always finds exactly one eligible subconstituent. The second is that features are cancelled from the list once they have done their duty. Lifting only the latter restriction will not help; the language will remain semilinear. Lifting the first is an unexplored option, but my hunch is that it too will not help. One might lift both, or one might propose that the case markers are added in the morphology anyway, thus generating an infinite lexicon to start with. This seems to be the least dramatic change and can even be motivated linguistically. What case markers do, then, is to add features to the lexical item, roughly in the way Anderson [1] has outlined it. (If suffix order is strictly iconic, we might simply view addition of a feature as concatenation with the feature sequence of the word to which the suffix is attached.)

In her book [33], Jane Simpson tried to deal with case stacking by resorting to a distinction between syntactic and semantic case (following Mel'čuk and Blake [4]). The grammar is couched in LFG, which allows to distinguish at least two levels of construction: c–structure (the structure that determines the syntactic shape) and f–structure (the structure that determines functional relationships). These levels are built up in tandem, and rules within LFG are at the same time rules for building up c–structure and f–structure. (There are other levels, too, but they don't play a role here.) Andrews [2] drew attention to the fact that in certain situations we have two successive cases which add contradictory features into the f–structure. Since the structures are grammatical, and the f–structure is inconsistent, something must have gone wrong. Clearly, this problem is exacerbated in languages that unlike Warlpiri (which has only two case shells) allow for unbounded stacking of case suffixes, since then the space of basic feature values is necessarily finite. The problem is easy to spot: the sequence of case markers is mapped into a flat feature structure, while it clearly reflects something of a syntactic hierarchy of constituents. This indication of the hierarchy is thrown away in the process of translation. However, formalisms based on attribute value structures allow for recursion. Treating CASE as a recursive feature, we may actually allow for feature structures such as the following.

$$\left[\begin{array}{lll}\text{CASE} & : & \left[\begin{array}{lll} acc & & \\ \text{CASE} & : & \left[\begin{array}{lll} nom & & \\ \text{CASE} & : & gen \end{array}\right]\end{array}\right]\end{array}\right]$$

It is this basic insight on which the unification based approaches rest (see Kracht [17] for a discussion of exactly this aspect of the case–feature within AV–logics). The paper by Andrews is not specifically clear how the structures are being built up, but essentially the proposal is that the stacked cases create a morphological structure on the side which is used to identify (by means of inside out unification) the place into which the f–structure for the head shall be put. Nordlinger [27] modifies this proposal as follows. She assumes that case markers may (in certain languages) project their own f–structure. This is to say in effect that the recursive structure shown above is part of the f–structure of a corresponding noun, and moreover, by standard rules of LFG the case markers create this f–structure. Second, independent of this the c–structure may be flat (as in Jiwarli and Kayardild) or articulate (as in Martuthunira). Thus, although case stacking allows in principle free word order, languages may (but need not) additionally have configurational syntax. It is perhaps this aspect of her proposal that shows the strength of LFG. Having the possibility of developing c–structure and f–structure independently turns into an advantage when it comes to analysing these languages. As we have seen above, suffixaufnahme allows for free word order; but, as has been observed, languages may have suffixaufnahme without free word order.

Nordlinger makes the following general claims.

> *Generalization A.* Case morphology can construct grammatical relations on a par with and independently from, phrase structure.
> *Generalization B.* Case morphology can construct the larger syntactic context, including providing complex information about the clause.

This model is designed for languages that have only cases. However, it can be straightforwardly extended to cover other agreement suffixes. Thus, the AVSs provide a format to integrate case stacks into syntax.

The previous models were all heavily syntactically oriented. It is perhaps enlightening to see a fairly simple semantical model that uses the case stacks to do the variable management, proposed by Ebert and Kracht [8]. Here, the basic structures are DRSs enriched by a so–called referent system which says in what ways the variables are identified once two structures are merged.

Definition 10. *A **variable** over F is a sequence $x = \sigma^\frown\circ$, where σ is a sequence of case markers. If y is a variable, then y^x denotes the result of replacing in y the symbol $\circ$ by x. For a set Δ of formulae, $\Delta^x := \{\delta^x : \delta \in \Delta\}$.*

Definition 11. *A **semantic structure** is a pair $A = [\circ : x, [V : \Delta]]$, where x is a variable and $[V : \Delta]$ a DRS, that is, V is a set of variables and Δ a set of formulae. A is **plain** if $x = \circ$. Let $A = [\circ : x, [V : \Delta]]$ and $B = [\circ : y, [V' : \Delta']]$. The **merge** of A and B, denoted by $A \oplus B$ is defined if and only if either A or B is plain.*

- *If A is plain, $A \oplus B := [\circ : \circ, [V^y \cup V' : \Delta^y \cup \Delta']]$.*
- *If B is plain, $A \oplus B := [\circ : \circ, [V \cup V'^x : \Delta \cup \Delta'^y]]$.*

To see the mechanism work, let us write down a lexical entry for a simple noun and a case marker:

$$\begin{array}{c} /\mathsf{doctor}/ \\ \begin{array}{|c|}\hline \circ : \circ \\ \hline \varnothing \\ \hline \mathsf{doctor}'(\circ) \\ \hline \end{array} \end{array} \qquad \begin{array}{c} /\textsc{nom}/ \\ \begin{array}{|c|}\hline \circ : \textsc{nom}^\frown\circ \\ \hline \varnothing \\ \hline \varnothing \\ \hline \end{array} \end{array}$$

Now we shall compute the merge with the semantic structure for nominative:

$$\begin{array}{c} /\mathsf{doctor}/ \\ \begin{array}{|c|}\hline \circ : \circ \\ \hline \varnothing \\ \hline \mathsf{doctor}'(\circ) \\ \hline \end{array} \end{array} \oplus \begin{array}{c} /\textsc{nom}/ \\ \begin{array}{|c|}\hline \circ : \textsc{nom}^\frown\circ \\ \hline \varnothing \\ \hline \varnothing \\ \hline \end{array} \end{array} = \begin{array}{c} /\mathsf{doctor} + \textsc{nom}/ \\ \begin{array}{|c|}\hline \circ : \circ \\ \hline \varnothing \\ \hline \mathsf{doctor}'(\textsc{nom}^\frown\circ) \\ \hline \end{array} \end{array}$$

To see why this is so, notice that the symbol $\circ$ in the first structure is replaced by $\textsc{nom}^\frown\circ$ according to the laws of merge. Now, in order to understand the potential of this proposal let us repeat this example with a relational noun, teacher:

$$\begin{array}{c} /\mathsf{teacher}/ \\ \begin{array}{|c|}\hline \circ : \circ \\ \hline \varnothing \\ \hline \mathsf{teach}'(\circ, \textsc{gen}^\frown\circ) \\ \hline \end{array} \end{array}$$

If we add the suffix NOM, we get

$$\begin{array}{c} /\mathsf{teacher} + \textsc{nom}/ \\ \begin{array}{|c|}\hline \circ : \circ \\ \hline \varnothing \\ \hline \mathsf{teach}'(\textsc{nom}^\frown\circ, \textsc{gen}^\frown\textsc{nom}^\frown\circ) \\ \hline \end{array} \end{array}$$

Notice that by the mechanics of replacement, it is not only the main variable that changes its name but also the variable of the complement. This is what we will make use of.

Cases may or may not have a semantics. This actually does not make much of a difference for this calculus. Take the genitive, which in many languages is used for marking possession:

$$\begin{array}{c} /\textsc{gen}/ \\ \begin{array}{|c|}\hline \circ : \textsc{gen}^\frown\circ \\ \hline \varnothing \\ \hline \mathsf{belong\text{-}to}'(\circ, \textsc{gen}^\frown\circ) \\ \hline \end{array} \end{array}$$

So, when a genitive is attached, it says that the thing to which it attaches owns something. Here, $\circ$ represents the thing that is possessed, while $\textsc{gen}^\frown\circ$ is the thing that owns it. To see how this works, we shall turn to a real example. The following is a construction of Old–Georgian.

(23) saxel-ita mam-isa-jta
name-INST father-GEN-INST
with father's name

The first part is clear: saxel-ita is

$$(A) \qquad \begin{array}{c} /\mathsf{saxel\text{-}ita}/ \\ \hline \multicolumn{1}{|c|}{\circ : \circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{name}'(\textsc{inst}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{instr}'(\textsc{inst}^\frown\circ).} \\ \hline \end{array}$$

Here, $\mathsf{inst}'(x)$ means that x is an instrument. Actually, it would be more accurate to use a binary predicate $\mathsf{inst}'(x, e)$, with the intended interpretation that x is used as an instrument in the event e. We shall generously ignore this detail, however. Now let us turn to mam-isa-jta. First we attach the genitive to mam:

$$\begin{array}{c} /\mathsf{mam}/ \\ \hline \multicolumn{1}{|c|}{\circ : \circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{father}'(\circ, \textsc{gen}^\frown\circ)} \\ \hline \end{array} \oplus \begin{array}{c} /\mathsf{isa}/ \\ \hline \multicolumn{1}{|c|}{\circ : \textsc{gen}^\frown\circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{belong\text{-}to}'(\circ, \textsc{gen}^\frown\circ)} \\ \hline \end{array}$$

$$= \begin{array}{c} /\mathsf{mam\text{-}isa}/ \\ \hline \multicolumn{1}{|c|}{\circ : \circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{father}'(\textsc{gen}^\frown\circ, \textsc{gen}^\frown\textsc{gen}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{belong\text{-}to}'(\circ, \textsc{gen}^\frown\circ).} \\ \hline \end{array}$$

Next we attach the instrumental suffix:

$$(B) \qquad \begin{array}{c} /\mathsf{mam\text{-}isa\text{-}jta}/ \\ \hline \multicolumn{1}{|c|}{\circ : \circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{father}'(\textsc{gen}^\frown\textsc{inst}^\frown\circ, \textsc{gen}^\frown\textsc{gen}^\frown\textsc{inst}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{belong\text{-}to}'(\textsc{inst}^\frown\circ, \textsc{gen}^\frown\textsc{inst}^\frown\circ).} \\ \hline \end{array}$$

Finally, the two structures (A) and (B) are merged to derive the final representation (C).

$$(C) \qquad \begin{array}{c} /\mathsf{saxel\text{-}ita\ mam\text{-}isa\text{-}jta}/ \\ \hline \multicolumn{1}{|c|}{\circ : \circ} \\ \hline \\ \hline \multicolumn{1}{|c|}{\varnothing} \\ \hline \multicolumn{1}{|c|}{\mathsf{name}'(\textsc{inst}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{father}'(\textsc{gen}^\frown\textsc{inst}^\frown\circ, \textsc{gen}^\frown\textsc{gen}^\frown\textsc{inst}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{instr}'(\textsc{inst}^\frown\circ);} \\ \multicolumn{1}{|c|}{\mathsf{belong\text{-}to}'(\textsc{inst}^\frown\circ, \textsc{gen}^\frown\textsc{inst}^\frown\circ).} \\ \hline \end{array}$$

(C) is true in a model under an assignment for the variables if $\textsc{inst}^\frown\circ$ is instantiated to a thing x that is a name, and $\textsc{gen}^\frown\textsc{inst}^\frown\circ$ is instantiated to a thing y that is a father, and x belongs to y. And that y is a father of the value of $\textsc{gen}^\frown\textsc{gen}^\frown\textsc{inst}^\frown\circ$. This is exactly as it should be.

This model allows to assign a semantics for languages with full suffixaufnahme regardless of any syntactic word order constraints. These may (and must be) imposed over and above the semantics. However, this can be done without making reference to agreement marking, since the latter is semantic. Notice that the semantics works also for languages that have no suffixaufnahme, provided that the parts that fail to exhibit suffixaufnahme are contiguous. Therefore, this model is fairly close to Hale's analysis (though it does not make use of movement). Difficulties arise in particular with head–inflecting languages.

It is worth pointing at an important advantage of these kinds of approaches. It is predicted that constituent boundaries arise in order to restrict the space of possible analyses of a sentence. Namely, if N is a cse marker, $A \oplus (B \oplus N)$ is different from $(A \oplus B) \oplus N$, the latter being the same as $(A \oplus N) \oplus (B \oplus N)$. Moreover, unlike Hale we do not propose underlying features that may or may not be spelled out. This drastically reduces the number of possible readings. In fact, not only is agreement in discontinuous constituent under our assumptions, it is predicted to be optionally present in the continuous phrase (which, to the best of my knowledge, is partially correct). Basically, the LFG analyses still stand out as the best ones available.

6. Further Cases

The preceding discussion centered around the extreme form of suffixaufnahme. The theoretical results showed that such languages cannot be generated by many grammar formalisms. The reason is simply that a feature that is constitutive of the basic labels cannot be recursive. However, even if full recursion is not present, this does not mean that these formalisms are appropriate. We shall first present some evidence. Locative cases are semantically, syntactically and often morphologically layered. Their structure is as follows.

$$[M\ [L\ DP]]$$

Here, $[L\ DP]$ is a constituent denoting a (time dependent) location, while the enture phrase denotes a movement of an entity with respect to that location. (This layering has been proposed by Jackendoff (see eg [14] and earlier work) and Koopman [15].) What is interesting here is that heads may select either $M + L$ or just M (see Kracht [22]). Finnish is a very interesting case in point. Not only does it make heavy use of directional locatives (selection of M alone), it also has DP internal agreement. These two together are very hard to reconcile in a theory. I have argued in Kracht [19] that

(1) Cases are nothing but morphemes. A morpheme is a case if and only if it is selected by a higher head. Heads may select any number of morphemes in a row. What makes a morpheme a case is the fact that it is selected by some head.
(2) There are two kinds of merge: semantic merge and syntactic merge. Syntactic merge does nothing but to put the case marker on top of the case stack of that item.

Suffixaufnahme arises when the case that is stacked is stacked also on top of the cases of the arguments. So, when a DP that has a genitive marked argument is put into the instrumental in Old Georgian, the instrumental morpheme is added so that it now expects a genitive+instrumental DP. Such a DP is produced by attaching first genitive and then instrumental to the DP.

Let us take the case of phrase internal agreement. Within MP this is very delicate already for a simple reason. Every act of checking a feature eliminates that feature. Thus it cannot be used again for checking. Then, how is agreement between more than two items organized if checking is only between two? The answer is: this is impossible without creating higher order objects. This is reminiscent of an observation by Charles Peirce that polyadic relations cannot be generated from binary relations. At first blush there seems to be an easy way out: we just allow for one feature to check any number of agreement features (in linear logic terms: we change from a premiss p to the premiss $p!$, where the latter as opposed to the former can be used any number of times). This however is not really enough, as we have seen above. Additionally, binary checking is not really sufficient anyway. In [18] I have given arguments that number in coordinated structures is inherently ternary, showing that no binary checking whatsoever can generate the desired facts (unless, or course, certain 'ghost features' without obvious meaning are introduced). But this is a side issue.

Let us return to the NP internal agreement. There is reason to suspect that the "Pennsylvania Grammars", LCFRGs and MCTAGs, are not adequate in their handling of agreement. The reason is as follows. Recall that an LCFRS is a k–LCFRS for some k. This means that every constituent has at most k continuous parts. Agreement markers are iterated as many times as needed, without a priori bound. Thus, in these grammars it is impossible to view the multiple copies of the agreement morphemes as parts of a single constituent. Linguistically, this is the most satisfactory analysis, however. (It allows for an easy syntax–semantics translation algorithm, for example.)

7. Copying

Annius Groenink, building on work by Bill Rounds [31], has given a simple characterization in terms of grammars of languages that can be recognized in polynomial time. Without going into details, it can be said that the grammars defined in Groenink [11] (called **simple literal movement grammars**, or **SLMG**s, for short), extend the LCFRSs by two additional devices:

(1) copying
(2) multiple analyses

Multiple analyses means that may make the application of a rule conditional on the fact that a certain string possesses different analyses at the same time. This feature of SLMGs need not concern us further, though. What is of importance is copying. Using copying one can easily generate languages that are not semilinear.

Theorem 12 (Groenink)**.** *A language is recognizable in polynomial time on a deterministic Turing machine if and only if it can be generated by an SLMG.*

Indeed, by this theorem and Theorem 9 above, the languages with full suffixaufnahme can be generated by SLMGs. Unfortunately, the grammars are anything but transparent. Moreover, they suffer from the same deficit as we have noted above for LCFRSs with respect to agreement: it is not possible to assign the copies of the agreement morpheme a single rule.

In Kracht [21] I have proposed a different line of attack. The basic mistake, in my view, is to always assume that a language sign has a specific substance associated with it. For example, plural in English (exceptions aside) is realized by a suffix `s`, possibly using some morphophonological rules; past tense is signalled by `ed`, and so on. Copying, however, does not fit this picture. But it does exist (some claim, as does Manaster–Ramer ([23]), that it exists even in English). However, there are languages where plural is formed simply by reduplicating the head noun (Malay/Bahasa Indonesia). Mandarin Chinese forms a yes–no–question by copying an entire phrase (see [30]). In all these cases we are hard pressed to say what the exponent of the sign for plural or yes–no–question is. SLMGs do not answer the question: for them, copying is the effect of using a syntactic construction. I think that this is an illusion. Rather, reduplication simply *is* the exponent of the sign. This insight leads to the following type of grammar. We call a **de Saussurean sign** a pair $\langle E, M \rangle$ a pair, where E is a (typed) λ–term over (typed) strings with no empty binding and M a meaning. For example, plural in Malay is the sign $\langle \lambda x.x^\frown\text{-}^\frown x, \lambda \mathcal{P}.\{x : \mathcal{P}(x)\}\rangle$, where $^\frown$ denotes concatenation. This sign is not pronounceable, but once applied to a noun the whole can be pronounced (orang-orang (*men*), buku-buku (*books*)). The existence of unpronounceable units is nothing revolutionary. Roots in inflectional languages belong to this kind. Indeed, we may simply analyse a Latin noun as something that is pronounceable as soon as it has case and number. Similarly with an adjective. This analysis has the advantage that case (and other agreement features) is inserted only after the production of the constituent, and then distributes itself automatically across all parts according to the morphological laws. This solution, however, easily becomes abstract in the sense that it now relegates the constituency to the morphology. An analysis of suffixaufnahme, by the way, is still extant.

8. Conclusion

Suffixaufnahme, as we hope to have shown, is not as marginal a phenomenon as it may appear at first sight. Moreover, it proves to be a very intricate phenomenon as soon as one starts to take it seriously. It is not hard to understand what suffixaufnahme actually is: it is a (perhaps extreme) form of agreement. However, it is extremely difficult to provide an insightful analysis of suffixaufnahme within a theory of agreement. The problem lies less with suffixaufnahme itself but rather with our poor understanding of agreement, which in my view results from a preoccupation among formalisms of our era with syntactic questions. They fare in fact quite badly when it comes to agreement, although notable exceptions exist. The syntactician's idea of 'agreement features' is not in itself mistaken; but it needs a lot of sophistication to execute it properly.

References

[1] S. Anderson. *A–morphous Morphology.* Cambridge University Press, Cambridge, 1992.

[2] Avery C. Andrews. Semantic case–stacking and inside–out unification. *Australian Journal of Linguistics*, 16:1 – 55, 1996.

[3] Arto Anttila and Vivienne Fong. The Partitive Constraint in Optimality Theory. *Journal of Semantics*, 17:281 – 314, 2001.

[4] Barry J. Blake. *Case.* Cambridge Textbooks in Linguistics. Cambridge University Press, 1994.

[5] Winfried Boeder. Suffixaufnahme in Kartvelian. In Frans Plank, editor, *Double Case. Agreement by Suffixaufnahme*, pages 151 – 215. Oxford University Press, 1995.

[6] Alan Dench and Nicholas Evans. Multiple case–marking in Australian languages. *Australian Linguistics*, 8:1 – 47, 1988.

[7] Igor M. Diakonoff and S. A. Starostin. *Hurro–Urartian as an Eastern Caucasian Language.* Kitzinger, 1986.

[8] Christian Ebert and Marcus Kracht. Formal syntax and semantics of case stacking languages. In *Proceedings of COLING 2000*, 2000.

[9] Nicholas D. Evans. *A Grammar of Kayardild. With Historical–Comparative Notes on Tangkic.* Mouton de Gruyter, Berlin, 1995.

[10] Nick Evans. Multiple Case in Kayardild: Anti–iconic Suffix Ordering and the Diachronic Filter. In Frans Plank, editor, *Double Case. Agreement by Suffixaufnahme*, pages 396 – 428. Oxford University Press, 1995.
[11] Annius Groenink. *Surface without Structure. Word Order and Tractability Issues in Natural Language Analysis.* PhD thesis, University of Utrecht, 1997.
[12] Kenneth Hale. The adjoined relative clause in australia. In R. M. W. Dixon, editor, *Grammatical Categories in Aboriginal Languages*, pages 78 – 105. Australian Institute of Aboriginal Studies, Canberra, 1976.
[13] Henk Harkema. A Characterization of Minimalist Languages. In Philippe de Groote, Glyn Morrill, and Christian Retoré, editors, *Logical Aspects of Computational Linguistics (LACL '01)*, number 2099 in Lecture Notes in Artificial Intelligence, pages 193 – 211, Berlin, Heidelberg, 2001. Springer.
[14] Ray Jackendoff. *Semantic Structures.* MIT Press, Cambridge, Massachussetts, 1990.
[15] Hilda Koopman. Prepositions, Postpositions, Circumpositions and Particles: The Structure of Dutch PPs. unpublished, 1997.
[16] Jaklin Kornfilt. *Turkish.* Routledge, London and New York, 1997.
[17] Marcus Kracht. Is there a genuine modal perspective on feature structures? *Linguistics and Philosophy*, pages 401 – 458, 1995.
[18] Marcus Kracht. On Reducing Principles to Rules. In Patrick Blackburn and Maarten de Rijke, editors, *Specifying Syntactic Structure*, pages 45 – 73. CSLI, Stanford, 1997.
[19] Marcus Kracht. Against the Feature Bundle Theory of Case. In *New Perspectives on Case.* CSLI, 2001.
[20] Marcus Kracht. The combinatorics of cases. *Journal for Language and Computation*, 2, 2001.
[21] Marcus Kracht. Mathematics of language. unpublished, 2002.
[22] Marcus Kracht. On the semantics of locatives. *Linguistics and Philosophy*, 2002. to appear.
[23] Alexis Manaster-Ramer. Copying in natural languages, context–freeness and queue grammars. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 85 –89, 1986.
[24] Igor Mel'čuk. Toward a definition of case. In A. D. Brecht and J. S. Levine, editors, *Case in Slavic*, pages 35 – 85. Columbus, Ohio: Slavica, 1986.
[25] Jens Michaelis. Derivational minimalism is mildly context sensitive. In Michael Moortgat, editor, *Proceedings of LALC '98.* Springer, 2000.
[26] Tara W. Mohanan. Case OCP. A constraint on word order in Hindi. In Miriam Butt, Tracy Holloway King, and Gilliam Ramchand, editors, *Theoretical Perspectives on Word Order in South Pacific Languages.* CSLI, Stanford, 1994.
[27] Rachel Nordlinger. *Constructive Case. Evidence from Australian Languages.* Dissertations in Linguistics. CSLI, Stanford, 1998.
[28] Almerindo E. Ojeda. A Linear Precedence Account of Cross-Serial Dependencies. *Linguistics and Philosophy*, 11:457 – 492, 1988.
[29] Frans Plank. (Re-)Introducing Suffixaufnahme. In Frans Plank, editor, *Double Case. Agreement by Suffixaufnahme*, pages 3 – 110. Oxford University Press, 1995.
[30] Daniel Radzinski. Unbounded Syntactic Copying in Mandarin Chinese. *Linguistics and Philosophy*, 13:113 – 127, 1990.
[31] William C. Rounds. LFP: A Logic for Linguistic Description and an Analysis of its Complexity. *Computational Linguistics*, 14:1 – 9, 1988.
[32] Hiroyuki Seki, Takashi Matsumura, Mamoru Fujii, and Tadao Kasami. On multiple context–free grammars. *Theoretical Computer Science*, 88:191 – 229, 1991.
[33] Jane Helen Simpson. *Warlpiri Morpho–Syntax.* Kluwer, Dordrecht, 1991.
[34] Edward P. Stabler. Derivational minimalism. In Christian Retoré, editor, *Logical Aspects of Computational Linguistics (LACL '96)*, number 1328 in Lecture Notes in Artificial Intelligence, pages 68 – 95, Heidelberg, 1997. Springer.
[35] Ilse Wegner. Suffixaufnahme in Hurrian: Normal Cases and Special Cases. In Frans Plank, editor, *Double Case. Agreement by Suffixaufnahme*, pages 136 – 147. Oxford University Press, 1995.