

An empirical test of the Agglutination Hypothesis¹

Martin Haspelmath

Max-Planck-Institut für evolutionäre Anthropologie, Leipzig

In this paper, I approach the agglutination-fusion distinction from an empirical point of view. Although the well-known morphological typology of languages (isolating, agglutinating, flexive/fusional, incorporating) has often been criticized as empty, the old idea that there are (predominantly) agglutinating and (predominantly) fusional languages in fact makes two implicit predictions. First, agglutination/fusion is characteristic of whole languages rather than individual constructions; second, the various components of agglutination/fusion correlate with each other. The (unstated, but widely assumed) Agglutination Hypothesis can thus be formulated as follows:

(i) First prediction: If a language is agglutinating/fusional in one area of its morphology (e.g. in nouns, or in the future tense), it shows the same type elsewhere.

(ii) Second prediction: If a language is agglutinating/fusional with respect to one of the three agglutination parameters (a-c) (and perhaps others), it shows the same type with respect to the other two parameters: (a) separation/cumulation, (b) morpheme invariance/morpheme variability, (c) affix uniformity/affix suppletion.

I report on a study of the nominal and verbal inflectional morphology of a reasonably balanced world-wide sample of 30 languages, applying a variety of measures for the agglutination parameters and determining whether they are cross-linguistically significant. The results do not confirm the validity of the Agglutination Hypothesis, and the current evidence suggests that “agglutination” is just one way of trying to capture the strangeness of non-Indo-European languages, which all look alike to Eurocentric eyes.

1. Agglutination and fusion: An ambiguous success story

One of the seemingly most successful stories in the history of linguistic typology is the creation of a holistic morphological typology in the first half of the 19th century, initially by the combined efforts of Friedrich von Schlegel (1808), August Wilhelm von Schlegel (1818), and Wilhelm von Humboldt (1822, 1836). As is widely known (e.g. Greenberg 1974:35-41), these three were responsible for the classical subdivision of languages into an **isolating** (or “analytic”) type, an **agglutinating** type, a **fusional** (or **flexive**) type,² and an **incorporating** type. This way of classifying languages was made popular especially by Schleicher (1850) and Müller (1871), and has been part of linguists’ textbook knowledge ever since. Almost every introduction to linguistics mentions the terms, and they are frequently used in the technical literature (at least the term *agglutinating/agglutination*). Three representative sentences from recent works by influential authors are given in (1).

- (1) a. Evans (1995:1): “Kayardild is a dependent-marking, agglutinating, entirely suffixing language with a free order of phrasal constituents and a rich system of case-marking...”

¹ Earlier versions of this paper were presented at the 3rd conference of the Association for Linguistic Typology (Amsterdam 1999) and at the 9th International Morphology Meeting (Vienna 2000). I am grateful to the audiences at these occasions, as well as to two reviewers, for useful comments. (The core ideas of this paper were first presented at the DGfS Summer School on Language Typology in Mainz, September 1998.)

² For the flexive type, other term variants such as (*in*)*flexional*/*inflectional* are also often used. As was often noted (e.g. Bazell 1958), the term *inflectional* is confusing because it also has a different sense: One also says that agglutinating languages have inflection (i.e. different word-forms belonging to a single lexeme), so Sapir’s term *fusional* has tended to supplant it in the typological sense. Plank (1999) retains the term *flexive* (deliberately differentiating it from *inflectional*), presumably because he feels that there is much more to the agglutination/flexion distinction than what Sapir meant by *fusion* (cf. note 7). In this paper, I use *fusion* as the opposite of *agglutination*, simply because it seems that this term is now better known.

- b. Slobin (1997:281): "... On this account, agglutinating languages like Turkish and Japanese have no closed-class morphemes..."
- c. Hyman (2001:1397): "... In other words, a highly developed paradigmatic system of tonal oppositions appears not to be very compatible with a highly developed syntagmatic system of agglutinative morphology"

But at least since Sapir (1921), it has been widely recognized that this four-way distinction is problematic, because it conflates three different parameters: (i) the **degree of synthesis** (isolating vs. non-isolating, or in other terms, analytic vs. synthetic); (ii) the **degree of stem combination** (incorporating vs. non-incorporating); and (iii) the **degree of fusion** (agglutinating vs. fusional/flexive). The notions of synthesis and stem combination are quite easy to describe and identify, as long as one agrees on word boundaries and a definition of "stem". But what exactly is meant by agglutination and fusion/flexion, and what we need such concepts for, cannot be so readily explained.

So with respect to the agglutination/fusion distinction, the success of the 19th century classification is ambiguous: While the classification is still widely known, it does not have an exact meaning and does not seem to be taken seriously. In fact, with a few notable exceptions (Skalička 1951/1979, Plank 1986, 1991, 1999, Plungian 2001, Testelefs 2001), linguistic theorists in the latter half of the 20th century have either ignored or severely criticized 19th century morphological typology. The quotations in (2) seem to be fairly characteristic of the mainstream view.

- (2) a. Anderson (1985:10): "... nothing much seems to follow from this classification: it has never been shown, for example, that languages with agglutinative properties share other features of a non-accidental sort that are not shared with non-agglutinating languages as well. For these reasons, the traditional terms do not seem to constitute any significant typology."
- b. Bauer (1988:170): "Basically, a typology is not of much value unless it predicts other things about the various types of languages... Now ... a typology in terms of isolating, agglutinative and fusional does not [seem to] correlate with anything else in the morphology at all... The value of the typology *qua* typology is thus very much in doubt."
- c. Spencer (1991:38) "This typology, though sanctioned by tradition, has been criticized for being both incoherent and useless. It is useless because nothing of any interest follows from classifying languages in this way."

But if classical morphological typology is "incoherent and useless", why is the terminology still with us? Moreover, most linguists find that the parameters of degree of synthesis (e.g. Greenberg 1954[1960], Bickel & Nichols 2005) and incorporation (e.g. Baker 1996) are of great typological interest. Is there also an "agglutination parameter"? In the following section I will argue that the agglutination/fusion distinction much more interesting than superficial textbook statements such as in (2a-c) make it appear, but we will eventually see that an empirical test largely comes to negative conclusions about its validity.

2. The Agglutination Hypothesis: Implicit claims made explicit

Anderson, Bauer and Spencer make seriously misleading statements when they claim that nothing of interest follows from classifying languages into agglutinating and fusional types, and that it implies no correlations. The use of this classification does imply testable claims about extremely interesting correlations, but these claims are usually not made explicit.

The reason why the classification implies a number of correlations is that "agglutinating language" is not a primitive notion. As we will see shortly, saying that a language or a pattern is agglutinating embodies a set of logically separate claims. By accepting the classification of languages or patterns into agglutinating and non-agglutinating types, one implicitly accepts the idea that the various properties that make a language agglutinating correlate with each other. Thus, we should think of agglutination not so much as a classificatory concept or as a parameter of variation, but as an empirically testable hypothesis. I call it the **Agglutination Hypothesis** here, and I distinguish two main components of it. (This is very similar to Plank's (1999:285) "Strong Homogeneity Hypothesis".³)

The first empirical claim (or universal prediction) follows from the fact that what is normally classified as agglutinating or fusional is whole languages, not just small subsystems. So implicit in this classification is the expectation that the First Prediction in (3) is correct.

(3) **First Prediction: Correlation among parts of the morphology**

If a language is agglutinating/fusional in one area of its morphology (e.g. in nouns, or in the future tense), it tends to show the same type elsewhere.

Logically speaking, languages could of course be agglutinating in their noun morphology and fusional in their verb morphology, or even isolating in the indicative mood and agglutinating in the subjunctive mood. But linguists tend to accept the notion that the morphological types apply to entire languages, thus implicitly claiming that languages are more homogeneous than they would have to be. This presupposition of course goes back to the early 19th century, and it could be that it is simply naive. In the early days of modern typology, what struck linguists was the differences between languages, and not so much the more abstract differences between patterns. In other areas of typology, linguists also tended to ascribe interesting differences to languages in the past (e.g. "ergative languages", "tone languages"), whereas more recently they have usually become more careful, recognizing that system splits (such as ergative in the perfective vs. accusative in the imperfective) are not a marginal phenomenon. Dividing languages into accusative, ergative and neutral languages is a meaningful enterprise only if one expects that different constructions tend to show the same patterns, or at least that one of the constructions is somehow predominant.

The second empirical claim follows from the fact that agglutination is not a primitive feature. But what exactly are the component features of agglutination? Unfortunately, it is very difficult to find a precise definition of agglutination and fusion in the literature. A few exemplary definitions are listed in (4) (emphasis is mine).

³ The only reason I do not adopt Plank's term is that my term "Agglutination Hypothesis" appears in the title of this paper, and I did not want the paper to be unrecognizable. Note that I developed these ideas independently of Plank (but under the influence of his earlier work, e.g. Plank 1986, 1991), at about the same time.

- (4) a. Bloomfield (1933:207): “*Inflectional* languages show a merging of semantically distinct features either in a **single bound form** or in **closely united bound forms**, as when the suffix *-o@* in a Latin form like *amo@* ‘I love’ expresses the meanings 1st, sg, pres, ind.”
- b. Hjelmslev (1963[1968:109]): “... *flektierenden* Typ, bei dem **die Grenze zwischen Wurzel und Suffix nicht klar ist**, bei dem jedes Suffix oftmals gleichzeitig mehrere verschiedene grammatische Verhältnisse ausdrückt, und bei dem **die Wurzel selbst bei der Flexion Veränderungen erfahren kann**.”⁴
- c. Anderson (1985:9): “(*in*)*flexional* languages have internally complex words which **cannot easily be segmented** into an exhaustive and non-overlapping string of formatives”
- d. Vance (1987:175): “[In agglutinative languages] **each morph represents only one semantic unit**... The other property is that morphs are simply stuck together (“agglutinated”); ideally, the **boundaries between morphs are clear**, and there is **no allomorphy**.”
- e. Whaley (1997:134): “A language is *fusional* **if the boundaries between its morphemes are hard to determine**. The effect is as if the morphemes were blending, or fusing, together.”

Most modern authors seem to agree that one of the ingredients of fusion is **cumulation**, i.e. the expression of several morphological subcategories in one affix (cf. the quotations from Bloomfield, Hjelmslev and Vance).⁵ Agglutinating structures, by contrast, exhibit **separation** (to use a terminology that was introduced by Frans Plank, cf. Plank 1986, 1999).

Another characteristic of fusional languages is that they exhibit **stem alternations**, i.e. the (co-)expression of morphological categories by changing, rather than adding to, the stem (cf. Hjelmslev's definition). This is actually the oldest criterion: Friedrich von Schlegel and his 19th century followers were primarily struck by the difference between, on the one hand, languages like Sanskrit and German with their salient stem vowel changes, and on the other hand, languages like Quechua, which exhibit complete **stem invariance**.

A very vague criterion that is often mentioned in definitions of fusion and agglutination is that segmentation of morphemes is “difficult” (cf. Hjelmslev's, Anderson's and Whaley's definitions). The question of course is what properties of the language cause these difficulties. It seems that they often arise from stem alternations (see the last paragraph) or from **affix alternations**. The basic idea is that in a typical agglutinating language, each affix not only stands for just one subcategory, but is also invariant in its shape, whereas in fusional languages, not only stems, but also affixes show considerable morphophonological allomorphy.⁶ Consider the partial paradigm of Hungarian noun inflection in (5).

⁴ “the inflectional type, in which the boundary between root and suffix is not clear, in which each suffix often expresses several different grammatical properties simultaneously, and in which the root can undergo changes even in inflection.”

⁵ Note that I use the term *subcategory* for concepts such as singular, dative, future, and *category-system* for concepts like number, case, tense. (Plank 1999 uses the terms *term/category* in the same sense.)

⁶ Greenberg (1954[1960:185]) uses morphophonological alternations of stems and affixes as the decisive criterion for his “index of agglutination”, i.e. he disregards cumulation, and he lumps stem alternations and affix alternations together. (In this he is followed by other linguists working in the tradition of quantitative typology, such as Krupa 1965 and Silnitsky 1993.)

(5) Hungarian			'house'	'table'	'river'
	Sg	Nom	<i>ház</i>	<i>asztal</i>	<i> folyó</i>
		Acc	<i>házat</i>	<i>asztalt</i>	<i>folyót</i>
	Pl	Nom	<i>házak</i>	<i>asztalok</i>	<i>folyók</i>
		Acc	<i>házakat</i>	<i>asztalokat</i>	<i>folyókat</i>

It is not immediately obvious how these forms should be segmented. For example, in the accusative singular form *házat*, the accusative suffix could be taken as *-at*, or as *-t* (as in *asztal-t* and *folyó-t*). In either case, we have to recognize an alternation, a stem alternation (*ház/háza*) or an affix alternation (*t/at*). Thus, it is the existence of alternations that makes segmentation difficult, and **affix invariance** makes it easy.

An even more radical kind of lack of one-to-one correspondence between meaning and form is **affix suppletion**, which is also commonly associated with fusional as opposed to agglutinating morphology (see especially Skalička 1951). Affix suppletion is allomorphy that cannot be described in phonological or morphophonological terms. It can be conditioned lexically, as in (6a); morphologically, as in (6b); or phonologically, as in (6c). Lack of affix suppletion will be called **affix uniformity**.

- (6) a. Kannada (Dravidian; India)
- | | | |
|--------|--------------|-------------|
| Plural | <i>-aru</i> | /humans |
| | <i>-gaɦu</i> | /inanimates |
- b. Latin
- | | | |
|-------------------|-----------|------------------|
| 1sg subject index | <i>-o</i> | /Present tense |
| | <i>-m</i> | /Imperfect tense |
- c. Lezgian (Nakh-Daghestanian; Caucasus; Haspelmath 1993:131)
- | | | |
|--------------|------------|-------------------|
| Aorist Part. | <i>-r</i> | /after high vowel |
| | <i>-ji</i> | /after low vowel |

It appears that **cumulation**, **stem alternation**, **affix alternation**, and **affix suppletion** are the key ingredients of non-agglutination or fusion. In the following, I will assume (somewhat counterfactually⁷) that no other morphological properties are relevant to defining the agglutination/fusion distinction.

⁷ In a few works that address the agglutination/fusion distinction in some detail (Skalic&ka 1951, Pöchtrager et al. 1998, Plank 1999, Plungian 2001), a number of further correlating properties have been mentioned:

agglutinating	fusional
affixes distinct	affixes often homonymous
always zero exponence	no/ sporadic zero exponence
only local exponence	also extended exponence
repeatable affixes	unrepeatable affixes
large paradigm size	small paradigm size (cf. Plank 1986)
loose bonding	tight bonding
optionality	obligatoriness

Thus, the number of claimed correlations that can be found in the literature is actually much higher. However, since these correlations are little known, one would not say that using the terms "agglutination" and "fusion" implies accepting these correlations as valid. Thus, testing these further correlations is less urgent than testing the correlations that are implicit in the definitions of "agglutination" and "fusion", and I will not say anything further about them here.

From the perspective of the individual properties, one could simply describe a language as cumulating, or as stem-alternating, or as affix-uniform, and so on. However, the fact that languages (or patterns) are typically described as "agglutinating" or as "fusional" shows that linguists expect the component properties to correlate with each other. So the Agglutination Hypothesis makes a second prediction:

(7) Second Prediction: Correlation among different component properties

If a language is agglutinating/fusional with respect to one of the component properties (cumulation, (stem/affix) alternation, affix suppletion), it will to behave similarly with respect to the other features.

Taken together, the two predictions made by the Agglutination Hypothesis are of course extremely interesting for the comparative linguist. But whether the terms "agglutination" and "fusion" are really useful depends also on whether the hypothesis is true. Somewhat surprisingly, this question has rarely been asked (Plank 1999, which came to my attention only after the first version of this paper had been finished, is the only exception known to me).

Apparently, it has not been generally recognized that the terms "agglutination" and "fusion" imply strong empirical claims, so linguists have not attempted to test these empirical claims. Another reason why such a test has not been carried out is probably that it encounters multiple difficulties, some of which will be addressed in the next section. However, I will also say how they can be overcome, so that in §4 I will present a first empirical test of the correlations.

3. Difficulties for an empirical test

If one wants to compare the morphological systems of widely different languages, one needs to make a large number of simplifications and idealizations, hoping that these do not introduce a bias into the investigation. In this section I discuss some of the difficulties I encountered in testing the two predictions of the preceding section, and make proposals for how to overcome them.

First, different languages show different morphological categories and vary significantly in morphological complexity. This is a difficulty especially for testing the First Prediction. For example, if this prediction is taken to imply that agglutination in nouns entails agglutination in verbs, it is impossible to test in languages that have inflectional morphology in verbs but not in nouns.⁸ To address this difficulty, I only looked at languages with a fair amount of inflection, and took into account only the core of nominal and verbal inflection. Thus, I left aside inflection of adjectives, adpositions and other word classes (which are difficult to compare across languages anyway). I also left aside non-finite verb morphology, which does not belong to the core of verbal inflection and differs more across languages than tense, aspect and voice morphology. The assumption behind this is that if the idea at the heart of the Agglutination Hypothesis is correct, the correlations should emerge even if only the core inflectional categories of nouns and verbs are considered.

Second, it is well-known that cumulation is extremely common everywhere with person-number affixes (cf. Cysouw 2003:296). Some examples of possessive person-number affixes in languages from different continents are given in (8).

⁸ Such languages could of course be taken to exhibit isolation in nouns, though this is rarely done.

(8)	Hungarian (Uralic; Europe)	O'odham (Uto-Aztecans; North America)	Tauya (Trans- New Guinea)	Lango (Nilotic; Africa)
Sg 1	<i>kez-em</i>	<i>ñ-kakkio</i>	<i>ya-neme</i>	<i>pàlà-ná</i>
2	<i>kez-ed</i>	<i>m-kakkio</i>	<i>na-neme</i>	<i>pàl'ÿ-ní</i>
3	<i>kez-e</i>	<i>kakkio-j</i>	<i>Ø-neme</i>	<i>pàlà-mE/rEfl</i>
Pl 1	<i>kez-üink</i>	<i>t-kakkio</i>	<i>se-neme</i>	<i>pàlà-wá</i>
2	<i>kez-etek</i>	<i>'em-kakkio</i>	<i>te-neme</i>	<i>pàl'ÿ-wú</i>
3	<i>kez-üik</i>	<i>ha-kakkio</i>	<i>ne-neme</i>	<i>pàlà-gí</i>
	'my hand etc.'	'my legs etc.'	'my head etc.'	'my knife etc.'

Thus, including person-number affixes would lead us to find some amount of cumulation in the great majority of languages, even in languages that otherwise would probably be regarded as perfect examples of the agglutinating type. Thus, in my empirical test, I counted "person-number(-gender)" as a single inflectional category-system, so that these cases do not count as cumulation.

Third, in tense-aspect-mood forms, cumulation also seems to be common, but it is extremely difficult to identify the categories that are cumulated and to distinguish cumulation from semantic complexity. Consider the examples in (9).

(9)		past + perfective	past + imperfective	
a.	Modern Greek	<i>é-γrap-s-e</i>	<i>é-γraf-Ø-e</i>	's/he wrote'
b.	Lezgian	<i>qaču-na</i>	<i>qaču-zwa-j</i>	'took'
c.	Italian	<i>port-ai</i>	<i>port-avo</i>	'I carried'
d.	Pipil	<i>chiw-ki</i>	<i>chiwa-ya</i>	'did'

In Modern Greek, one can readily isolate a perfective marker *-s* and a past marker *é-*. Both forms are fully compositional. In Lezgian (Haspelmath 1993), it is possible to identify an imperfective marker *-zwa* and a past marker *-j*, but the past perfective *-na* is not compositional. In Italian and Pipil (a Uto-Aztecans language of El Salvador), we have no compositionality at all. We might say that both Italian and Pipil are fully cumulative, but it is not even fully clear that "past" and "(im)perfective" are separate inflectional categories in these languages. To avoid these complications, I regarded "tense-aspect-mood" as a single category-system, unless there were very strong reasons to separate them.

Fourth, a problem with measuring the relevance of stem alternations is that in most languages, only a subset of lexical items show stem alternations (e.g. in German, only strong verbs alternate, and these are a small minority), and generally most lexical items do not alternate. We would probably want to quantify the importance of stem alternation: A language that has a stem change in one lexical item should not be regarded as equally fusional (with regard to the criterion of stem alternation) as a language that has stem changes in more than a hundred items. Thus, for each category-system separately, I assigned scores depending on the (estimated) absolute number of lexemes that exhibit stem alternations:

score of 1	fewer than 10 lexemes with stem alternation
score of 4	10-50 lexemes
score of 7	more than 50 lexemes
score of 10	the majority of (or all) lexemes

Fifth, morphophonological affix alternations are very difficult to distinguish from purely phonologically conditioned alternations. For example, the German plural suffix *-n/-en* (as in *Straße-n* 'streets', *Frau-en* 'women', *Partikel-n* 'particles') could either be described as morphophonological affix alternations (making German fusional in this regard), or as the same affix that happens to have different realizations for purely phonological reasons. After considering this problem thoroughly, I decided to exclude the property "affix alternation" from the empirical test, because it did not seem feasible to make the distinction in a coherent way for a greater number of languages.⁹

4. The empirical test

4.1. Data

Thus, in my empirical test of the two predictions I studied three properties of morphological systems: cumulation, stem alternation, and affix suppletion. I examined a sample of thirty languages which is somewhat biased in the usual way, with more languages from Europe than would be justified by the genealogical diversity in Europe. Not all the languages are unrelated, but each is from a different genus (in the sense of Dryer 2005). The thirty sample languages are listed in Table 1.

⁹ I also simplified the procedure by considering only up to fifteen subcategories per lexeme class, and only up to ten subcategories per category-system, on the assumption that more was not required to capture the type of the language.

Table 1: The 30 languages of the sample

Niger-Congo	Swahili	Dravidian	Kannada
Nilotic	Lango	Tibeto-Burman	
Kordofanian	Krongo	Bodic	Classical Tibetan
Afro-Asiatic		Newaric	Dolakha Newari
Semitic	Arabic	Trans-New Guinea	
Egyptian	Coptic	Adelbert R.	Tauya
Indo-European		Madang	Amele
Germanic	German	Oceanic	Ponapean
Italic	Latin	Pama-Nyungan	Martuthunira
Iranian	Ossetic	Yuman	Maricopa
Indic	Hindi-Urdu	Uto-Aztecan	
Uralic		Numic	Tümpisa Shoshone
Finnic	Finnish	Aztecan	Pipil
Ugric	Hungarian	Mayan	Tzutujil
Lezgic	Lezgian	Cariban	Hixkaryana
Turkic	Turkish	Paezan	Páez
Tungusic	Evenki	Quechua	Huallaga Quechua
Yukaghir	Kolyma Yukaghir		
Nivkh	Amur Nivkh		

For each of these thirty languages, I determined a Cumulation Index, an Alternation Index, and a Suppletion Index, separately for nominal and for verbal inflection. The value of these indices is always between 0 (maximal agglutination) and 100 (maximal fusion). The Cumulation Index is defined as the percentage of inflected forms that exhibit cumulation; the Alternation Index is defined as the average alternation score (cf. §3), i.e. the sum of stem alternation scores for each category-system divided by the number of category-systems; and the Suppletion Index is defined as the average percentage of subcategories (per category-system) that exhibit affix suppletion.

4.2. Testing the Second Prediction: Correlation among the three indices?

Let us begin by examining the Second Prediction: is there a correlation among the three indices? Do languages with a lot of stem changes also show a lot of cumulation, and do languages with a lot of cumulation also show a lot of affix suppletion? The values of the three indices for the sample languages are shown in Table 2, where the languages are listed in rank order.

Table 2: Index values and language ranking for each index

Ranking by Cumulation index:		Ranking by Alternation index:		Ranking by Suppletion index:	
1. Krongo	0	1. Hindi/Urdu	0	1. Nivkh	0
Lango	0	Hixkaryana	0	2. Tauya	3
Lezgian	0	Martuthunira	0	3. Ponapean	4
Newari	0	Nivkh	0	4. Quechua	10
Pipil	0	Páez	0	5. Páez	12
Ponapean	0	Swahili	0	6. Lezgian	12
Shoshone	0	Tauya	0	7. Tibetan	14
Tibetan	0	Turkish	0	8. Coptic	15
Turkish	0	9. Evenki	2	Krongo	15
Tzutujil	0	10. Coptic	5	10. Pipil	16
11. Swahili	0.1	11. Latin	7	11. Finnish	18
12. Amele	0.2	12. Tibetan	9	12. Maricopa	19
13. Maricopa	0.4	13. Krongo	12	13. Hixkaryana	22
14. Tauya	0.5	14. Yukaghir	13	14. Hungarian	23
15. Coptic	1.8	15. Maricopa	14	Turkish	23
16. German	2	Quechua	14	16. Swahili	28
17. Quechua	2.5	17. Lango	17	17. Martuthunira	29
18. Yukaghir	4	18. Newari	18	18. Lango	37
19. Hixkaryana	4.5	19. Tzutujil	24	19. Shoshone	38
20. Evenki	5	20. Amele	25	20. Evenki	39
21. Ossetic	6	21. Ossetic	26	21. Yukaghir	40
22. Nivkh	7	22. Shoshone	27	22. Newari	41
23. Arabic	8	23. Finnish	30	23. Hindi/Urdu	50
24. Finnish	13	24. Hungarian	36	24. Kannada	51
25. Kannada	14	25. Arabic	42	Ossetic	51
26. Hungarian	18	26. Kannada	43	26. German	56
27. Martuthunira	18	27. Lezgian	46	27. Arabic	62
28. Páez	30	28. Pipil	50	28. Amele	69
29. Hindi/Urdu	50	29. German	52	29. Tzutujil	77
30. Latin	66	30. Ponapean	75	30. Latin	84

Comparison between the three indices is made difficult by the fact that the numbers stand for very different things. The total range of the values is not very different: between 0 and 66 for Cumulation, between 0 and 75 for Alternation, and between 0 and 84 for Suppletion. Thus, no language shows 100% fusion with respect to any of the indices, and for each index, there is at least one language with 100% agglutination. However, it is easily seen that the Cumulation values tend to be much lower than the Alternation and Suppletion values. Cumulation is simply a rare phenomenon outside of the Indo-European family, if cumulation of person and number is disregarded and if tense and aspect are not counted separately (see the discussion in §3). Only the Indo-European languages Hindi-Urdu and Latin have figures over 20% of cumulative forms.

Now the crucial question is: Do the languages tend to exhibit similar values for each of the three indices?¹⁰ This is difficult to see by just inspecting Table 2, so in order to make the indices more comparable, I compared the rank values on the three indices. Table 3 lists languages in the order of average rank. The first three columns give the three indices, but the most important triple of figures is the rank values in the next three columns. These figures show, for example, that Turkish has rank 1 for Cumulation, rank 1 for Alternation, and rank 12 for Suppletion, and so on. By the criterion of average rank, Turkish turns out to be the "most agglutinating" language, while Arabic is the "least agglutinating" language, and Indo-European languages such as Latin and German are also close to the bottom. This accords well with our expectations, of course, but there are also some surprises, especially perhaps the low (and hence "relatively fusional") position of Kannada, Hungarian and Finnish. These languages have traditionally been labelled "agglutinating" (in Table 3, all languages which have been called "agglutinating" in the literature are preceded by an asterisk).

Of course, the confirmation of our expectations for Turkish and Latin does not mean that the Agglutination Hypothesis has been confirmed. On the contrary, it could well be that the models of these two well-known languages have been so powerful that linguists have unconsciously tended to define *agglutinating* as "Turkish-like" and *fusional* as "Latin-like".¹¹ In this way, various properties that happen to be combined in these languages would have become part of the typological prototype, although there is in fact no tendency for these features to cooccur in languages cross-linguistically.

¹⁰ That the different components of agglutination do not always go together has occasionally been noted in the literature. Thus, Vance (1987:176) notes for Japanese:

"Japanese morphology certainly tends to be agglutinative. The two properties involved in agglutination, however, do not correlate very well. Portmanteau morphs are rare... Allomorphy, on the other hand, is not at all uncommon..."

¹¹ A quotation such as the following is quite typical: "The ideal type of an agglutinating language, best represented in Turkish..." (Pöchtrager et al. 1998:57)

Table 3: Index values and rank numbers for each language

	Indices			Ranks			(Average Rank)
	Cum	Alt	Sup	Cum	Alt	Sup	
Tauya	0.5	0	3	14	4	2	6.7
*Turkish	0	0	23	5.5	4	14	7.8
Tibetan	0	9	14	5.5	12	7	8.2
Krongo	0	12	15	5.5	13	8.5	9.0
*Nivkh	7	0	0	22	4	1	9.0
*Swahili	0.1	0	28	11	4	16	10.3
Coptic	2	5	15	15	10	8.5	11.2
Hixkaryana	4.5	0	22	19	4	13	12.0
*Quechua	2.5	14	10	17	15.5	4	12.2
Páez	30	0	12	28	4	5	12.3
*Lezgian	0	46	13	5.5	27	6	12.8
Ponapean	0	75	4	5.5	30	3	12.8
Maricopa	0.4	14	19	13	14	12	13.0
Lango	0	17	37	5.5	17	18	13.5
Pipil	0	50	16	5.5	28	10	14.5
Newari	0	18	41	5.5	18	22	15.2
*Shoshone	0	26	38	5.5	22	19	15.5
*Martuthunira	18	0	28	27	4	17	16.0
*Evenki	5	2	39	20	9	20	16.3
*Yukaghir	4	13.5	40	18	14	21	17.7
Tzutujil	0	24	77	5.5	19	29	17.8
Hindi-Urdu	50	0	50	29	4	23	18.7
*Finnish	13	30	18	24	23	11	19.3
Amele	0.2	25	69	12	20	28	20.0
*Hungarian	18	36	23	26	24	14.5	21.5
Ossetic	6	26	51	21	21	24.5	22.2
Latin	66	7	84	30	11	30	23.7
German	2	52	56	16	29	26	23.7
Arabic	8	42	62	23	25	27	25.0
*Kannada	14	42	51	25	26	24.5	25.2

* = language has been referred to as "agglutinating" in the literature

As can be seen from Table 3, there are some "well-behaved" languages that show similar rankings with respect to the three indices, e.g. Maricopa (ranks 13, 14, 15), Ossetic (ranks 21, 21, 24.5), Arabic (ranks 23, 25, 27), and Kannada (ranks 25, 26, 24.5). If all languages were like these, we would say that the three indices really cluster with each other, and that the Second Prediction of the Agglutination Hypothesis has been confirmed.

But unfortunately, too many languages are not "well-behaved". For example, Nivkh has the ranks 22, 4, 1, Páez has the ranks 28, 4, 5, Ponapean has the ranks 5.5, 30, 3, and Hindi-Urdu has the ranks 29, 4, 23. In fact, there is no statistically significant correlation among the three rankings: the Kendall coefficient of concordance is $W = 0.37$, $X^2 = 32.19$, and the significance is $p < .35$, well above the significance level of .05. My empirical test thus provides no support for the Second Prediction of the Agglutination Hypothesis.

This does not mean, of course, that I have definitively disconfirmed this prediction. My assignment of the various indices has involved some simplifications and arbitrary choices that could have influenced the results in a negative way. But my results present a strong challenge to those linguists who still want to maintain the agglutination-fusion distinction.

4.3. Testing the First Prediction:

Correlation among nominal and verbal inflection?

But what about the First Prediction, concerning the correlation between different parts of the morphology? Since it is well-known that verbs and nouns sometimes behave differently (e.g. the Romance and Balkan Slavic languages lost their nominal case inflection, but retained the synthetic verb inflection of their ancestors), one might not have particularly high expectations with regard to this prediction. Already Greenberg (1954[1960:182]) noted that "a term like *agglutinative* applies primarily to a single construction. A language may well and indeed usually does contain some agglutinational as well as some nonagglutinational constructions." And Wolfgang U. Dressler, a consistent proponent of the agglutination-inflection typology along Skalička's lines (cf. Dressler 1985, Pöchtrager et al. 1998), has recently acknowledged that "noun inflection and verb inflection may have a different typological character within the same language and develop diachronically in typologically different directions" (Dressler 2005:7). Thus, it is not exactly expected that the First Prediction would fare better in my empirical test.

To test it, I looked at nouns and verbs separately, for all three indices. Table 4 shows the index values and rank numbers, separately for the three indices.

Table 4: Index values and rank numbers, separately for nouns (N) and verbs (V)

	Cumulation				Alternation				Suppletion			
	index		rank		index		rank		index		rank	
	N	V	N	V	N	V	N	V	N	V	N	V
Amele	0	0.5	11	14	0	50	7.5	28	91	47	28	22.5
Arabic	15	0	25	6.5	33	50	22	28	44	80	24	29
Coptic	0.6	3	22	17	10	0	17	7	6	24	7	13
Evenki	0	10	11	23	3	0	15	7	19	59	14	25
Finnish	19	7	26	20	42	18	24	17.5	21	14	17	9
German	0	4	11	18	35	70	23	30	67	45	26.5	21
Hindi/Urdu	100	0	27.5	6.5	0	0	7.5	7	100	0	29.5	2
Hixkaryana	0	9	11	22	0	0	7.5	7	20	24	15.5	13
Hungarian	1	36	23	28	60	13	28	15	13	33	11	18.5
Kannada	0	27	11	26	50	35	26.5	23.5	25	77	19	28
Krongo	0	0	11	6.5	0	25	7.5	19	20	10	15.5	7.5
Lango	0	0	11	6.5	0	33	7.5	21	27	47	20.5	22.5
Latin	100	33	27.5	27	0	14	7.5	16	100	69	29.5	26
Lezgian	0	0	11	6.5	50	42	26.5	25	8	17	10	10
Maricopa	0	0.7	11	15	17	12	18	14	6	32	7	17
Martuthunira	0	37	11	29	0	0	7.5	7	24	33	18	18.5
Newari	0	0	11	6.5	0	35	7.5	23.5	0	82	3	30
Nivkh	0	14	11	25	0	0	7.5	7	0	0	3	2
Ossetic	0	11	11	24	20	33	19	21	27	75	20.5	27
Páez	4	57	24	30	0	0	7.5	7	17	6	13	4.5
Pipil	0	0	11	6.5	66	33	29	21	7	24	9	13
Ponapean	–	0	–	6.5	100	50	30	28	0	7	3	6
Quechua	0	5	11	19	29	0	21	7	0	20	3	11
Shoshone	0	0	11	6.5	5	47	16	26	67	10	26.5	7.5
Swahili	0	0.1	11	13	0	0	7.5	7	50	6	25	4.5
Tauya	0	1	11	16	0	0	7.5	7	6	0	7	2
Tibetan	–	0	–	6.5	0	18	7.5	17.5	0	27	3	15
Turkish	0	0	11	6.5	0	0	7.5	7	15	31	12	16
Tzutujil	0	0	11	6.5	47	0	25	7	33	44	23	20
Yukaghir	0	8	11	21	27	0	20	7	31	50	22	24

We see that for Alternation and especially Suppletion, the rankings do correlate significantly (Spearman rank correlation coefficient, $p < 0.01$ (Suppletion), $p < 0.05$ (Alternation)). However, for Cumulation there is no significant correlation ($p < 0.20$).

Thus, we can say that languages tend to show (or lack) affix suppletion and stem alternations in both nouns and verbs simultaneously. In other words, languages can be meaningfully typologized as generally affix-supplementing (vs. affix-invariant) languages, and as generally stem-alternating (vs. stem-uniform) languages, at least across nouns and verbs. Thus, it is not completely meaningless to search for the kinds of correlations that we have been looking at here. But these positive correlations are hardly sufficient to justify a global distinction between "agglutinating" and "fusional" languages, because as we saw in the preceding section, affix suppletion and stem alternation do not correlate among each other.

5. Conclusions

The overall results of this study are mostly negative. The main positive point is that it is indeed possible, if difficult, to test the Agglutination Hypothesis empirically.

But the result of the preliminary empirical test carried out here is mostly negative. I have not found a statistically significant correlation between the Cumulation index, the Alternation index and the Suppletion index, which would have been expected if the correlations implicit in the composite types "agglutinating" and "fusional" existed.

Also, the evidence that nouns and verbs tend to behave alike is not overwhelming. There is no correlation with respect to cumulation, which is sometimes taken to be the primary defining property of "fusional" languages (especially in the more recent literature), and also for alternation, the evidence for a correlation is not particularly strong. However, with respect to affix suppletion, we can say that nouns and verbs tend to behave in the same way across languages.

Since the study is based on a relatively small number of languages and the sample is not truly representative of the diversity of the world's languages, these results are far from conclusive. I cannot say that I have shown that the Agglutination Hypothesis is wrong. However, the results cast sufficient doubt on the hypothesis to say that from now on, the burden of proof is on those who believe that it is correct.

As there are indications that stem alternation and affix suppletion characterize both nouns and verbs, one can legitimately characterize entire languages as *affix-supplementing* or *stem-alternating*. However, since cumulation does not tend to be similar across word classes, it is less meaningful to characterize entire languages as "cumulating" or "separatist". These terms are meaningful primarily when applied to particular morphological subsystems.

After this study, the terms *agglutination* and *fusion* have lost much of their legitimacy (unless they are given a technical sense that is at variance with the common usage). Of course, it is logically possible to define a composite type consisting of properties that do not correlate with each other (or only very partially), but it does not make much sense. While Turkish could perhaps still be characterized as "agglutinating" and Latin as "fusional", for many languages neither of these terms would apply, and it would not be possible to say that they are "intermediate" between these two extremes either. It is quite possible that the reason for the success of the agglutination/fusion distinction is that Latin and Turkish have been such prominent languages in Western linguistics over the last few hundred years. They differ strikingly in their morphological systems (cf. Plank 1991), and it is perhaps natural that from this point of view one would classify languages as more Latin-like or more Turkish-like. However, linguistics should move beyond Latinocentrism and Turkocentrism and try to do justice to each language, to describe and characterize it in its own terms, or in truly universal terms.

References

- Anderson, Stephen R. 1985. "Typological distinctions in word formation." In: Shopen, Timothy (ed.) *Language typology and syntactic description, vol. III*. Cambridge: Cambridge University Press, 3-56.

- Baker, Mark. 1996. *The polysynthesis parameter*. New York: Oxford Univ. Press.
- Bauer, Laurie. 1988. *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.
- Bazell, Charles. 1958. "Linguistic typology." In: Strevens, Peter D. (ed.) *Five inaugural lectures*. London: Oxford University Press, 29-49.
- Bickel, Balthasar & Nichols, Johanna. 2005. "Inflectional synthesis of the verb." In: Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.) *The World Atlas of Language Structures*. Oxford: Oxford University Press, 94-97.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt & Co.
- Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.
- Dressler, Wolfgang U. 1985. "Typological aspects of Natural Morphology." *Acta Linguistica Hungarica* 35: 51-70.
- Dressler, Wolfgang U. 2005. "Morphological typology and first language acquisition: some mutual challenges." In: *Morphology and Linguistic Typology: On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4) Catania, 21-23 September 2003*, edited by Geert Booij, Emiliano Guevara, Angela Ralli, Salvatore SgROI & Sergio Scalise, 7-20. (<http://mmm.lingue.unibo.it>)
- Dryer, Matthew S. 2005. "Genealogical language list." In: Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.) *The World Atlas of Language Structures*. Oxford: Oxford University Press, 584-644.
- Evans, Nicholas. 1995. *A grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Greenberg, Joseph. 1954 [1960]. "A quantitative approach to the morphological typology of language." In: *Method and Perspective in Anthropology: Papers in Honor of Wilson D. Wallis*, ed. by Robert F. Spencer, pp. 192-220. Minneapolis: University of Minnesota Press. [Reprinted in: *International Journal of American Linguistics* 26 (1960): 178-94.]
- Greenberg, Joseph. 1974. *Language typology: a historical and analytic overview*. The Hague: Mouton.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.
- Hjelmslev, Louis. 1963. *Sproget: En introduktion*. København: Berlingske Forlag. (Dt. *Die Sprache*. Darnstadt, Wiss. Buchgesellschaft, 1968)
- Humboldt, Wilhelm von. 1822/1985. "Ueber das Entstehen der grammatischen Formen, und ihren Einfluss auf die Ideenentwicklung", reprinted in: *Über die Sprache: Ausgewählte Schriften*, ed. by Jürgen Trabant, München: Deutscher Taschenbuch Verlag.
- Humboldt, Wilhelm von. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechtes*. Reprinted.
- Hyman, Larry. 2001. "Tonal systems." In: Haspelmath, Martin et al. (eds) *Language typology and language universals: An international handbook* (HSK), vol. II. Berlin: de Gruyter, 1367-1380.
- Krupa, Viktor. 1965. "On quantification in typology." *Linguistics* 12: 31-36.
- Müller, Friedrich Max. 1871. *Lectures on the science of language*. 2nd. Ed. New York: Scribner.
- Plank, Frans. 1986. "Paradigm size, morphological typology and universal economy." *Folia Linguistica* 20: 29-48.
- Plank, Frans. 1991. "Of abundance and scantiness in inflection: a typological prelude", in: Plank, Frans (ed.) 1991. *Paradigms: the economy of inflection*. Berlin: Mouton de Gruyter, 1-39.

- Plank, Frans. 1998. "The co-variation of phonology with morphology and syntax", *Linguistic Typology* 2.2:195-230.
- Plank, Frans. 1999. "Split morphology: How agglutination and flexion mix", to appear in *Linguistic Typology* 3.3: 279-340.
- Plungian, Vladimir A. 2001. "Agglutination vs. flection", In: Haspelmath, Martin et al. (eds) *Language typology and language universals: An international handbook* (HSK), vol. I. Berlin: de Gruyter, 669-678.
- Pöchtrager, Markus A. & Bodó, Csanád & Dressler, Wolfgang U. & Schweiger, Teresa. 1998. "On some inflectional properties of the agglutinating type illustrated from Finnish, Hungarian and Turkish inflection." *Wiener Linguistische Gazette* 62-63: 57-92.
- Sapir, Edward. 1921. *Language*. New York: Harcourt.
- Schlegel, August Wilhelm von. 1818. *Observation sur la langue et la littérature provençales*. Paris: Librairie Grecque-latine-allemande.
- Schlegel, Friedrich von. 1808. *Ueber die Sprache und Weisheit der Indier: Ein Beitrag zur Begründung der Alterthumskunde*. Heidelberg: Mohr & Zimmer.
- Schleicher, August. 1850. *Die Sprachen Europas in systematischer Übersicht*. Bonn: H. B. König.
- Silnitsky, George. 1993. "Typological indices and language classes: a quantitative study." *Glottometrika* 14: 139-160.
- Skalička, Vladimír. 1951. "Das Erscheinungsbild der Sprachtypen." In: Skalička, V. 1979. *Typologische Studien*. Braunschweig: Vieweg, 21-58.
- Slobin, Dan. 1997. "The origins of grammaticizable notions." In: Dan Slobin (ed.), *The Crosslinguistic Study of Language Acquisition, Volume 5: Expanding the Contexts*. Mahwah, NJ: Erlbaum.
- Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.
- Testelefs, Yakov G. 2001. "Russian works on linguistic typology in the 1960s-1990s", in: Haspelmath, Martin et al. (eds) *Language typology and language universals: An international handbook*. Berlin: de Gruyter, 306-323.
- Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany: SUNY Press.
- Whaley, Lindsay J. 1997. *Introduction to typology*. Thousand Oaks/CA: Sage.