

Finnish as an Agglutinating Language

F Karlsson, University of Helsinki, Helsinki, Finland

© 2006 Elsevier Ltd. All rights reserved.

A theoretically ideal agglutinative language would satisfy five criteria: there are no inflectional classes and all words of the same part of speech are inflected in the same way; there are several morphotactic positions for affixes of composite word forms, especially those of nouns and verbs; every morphological element (stem or affix) is clearly segmentable; the affixes convey one rather than several grammatical meanings; there are no morphophonological alternations in any element due to morphological processes such as affixation. As a corollary, every element has exactly one phonological shape (disregarding low-level phonetic processes) and no fusion of several meaning elements into one unsegmentable whole.

Morphotactic Structure

Finnish is a suffixing language with a 14-member case system. As for basic inflectional and cliticized morphotactic positions, the surface structures of Finnish nominal (nouns, adjectives, pronouns, and numerals), finite, and nonfinite verb forms are as follows (INE = inessive case, INS = instructive case, COND = conditional mood) (see Table 1, Table 2, and Table 3).

Inflectional Classes of Nominals

There is no consensus on how many inflectional classes there are for nominals and verbs. Traditional Finnish lexicography as manifested in *Nykysuomen sanakirja* (Dictionary of modern Finnish, 1951-1961) postulates 82 inflectional classes for nominals, whereas, at the other extreme, a generative description such as Wiik (1967) operates with none but a wealth of ordered (morpho)phonological rules. A surface-oriented morphological approach would recognize at least 10 nominal inflectional classes. The most important ones are those ending in *Bi*, *Be*, and *Bs* in the nominative singular. Four case forms (nominative,

genitive, and partitive singular, partitive plural) unfold the allomorphic variation in the stems of each class, closely linked to the selection of particular ending allomorphs (NZ = nominalizer) (Table 4).

Class 1 is the largest and most productive one with the least amount of stem allomorphy, minimally only the stem vowel *Bi* alternates with *Be* before the plural *Bi*. At least 10 000 nominals are inflected according to class 1 and this is the pattern of most borrowings and other neologisms.

Classes 2 and 3 are closed: class 2 has some 220 words and class 3 around 40. Class 2 is more complex than class 1 as the stem vowel alternates also in the singular, and on top of that class 3 is more complex than class 2 by further eliding the stem vowel in SG.PART and also having alternations in the medial stem consonant. Many of the words in classes 2 and 3 are high-frequency words, for class 3, for example, *kuusi* '6,' *uusi* 'new,' *vesi* 'water,' *viisi* '5.'

For nominals in *Be*, class 5 is in principle closed (disregarding certain derivatives) even if it has more than 1000 members. Class 4 is small in contemporary Finnish, with less than 100 items, but this class is productive. Comparison of classes 1-3 and 4-5 discloses, not surprisingly, that the productive inflectional classes are those that have a minimal amount of stem allomorphy.

For nominals in *Bs*, class 6 (>4000 items) is simpler and more productive than class 7 (some 800 items). Class 8 covers a common type of derivatives.

Inflectional Classes of Verbs

Nykysuomen sanakirja postulates 45 inflectional classes for verbs. A more generalizing approach would do with six classes, here presented by way of four central inflectional forms, the endingless (NOM) infinitive, the first person singular present tense indicative form, the third person singular past tense form, and the past tense participle in the nominative singular form (Table 5).

Table 1 Nominals

Stem	Number	Case	Possessive	Clitic
talo		ssa		
house		INE		
'in (a/the) house'				
talo		ssa	ni	kin
house		INE	POSS.1.SG	also
'also in my house'				
talo	i	sta	si	ko
house	PL	EL	POSS.2.SG	O
'from your houses?'				

Table 2 Finite verb forms

Stem	Indefinite	Tense/mood	Person	Clitic
sano			n	
say			PERS.1.SG	
'I say'				
sano		i	t	ko
say		PAST	PERS.2.SG	O
'did you say?'				
sano	tta	isi	in	pa-
say	INDEF	COND	PERS.1	rtuvn
'even if one would say'				

Class 1 is by far the largest class with some 10 000 members. Around 2000 verbs belong to class 2, but this is presently the most productive verb class, obviously because overall there is less stem allomorphy in class 2 than in class 1; cf. especially the indefinite forms (called passives in traditional Finnish grammar) where the last morpheme is a personal ending for the indefinite 'fourth' person.

Class 4 is also a strong class with 4000 verbs. Class 3 has fewer than 20 monosyllabic verbs but around 1000 polysyllabic ones.

Morphophonological Alternations in Stems

As demonstrated by the example words of the nominal and verbal inflectional classes, Finnish word structure is characterized by considerable allomorphy, both in stems and suffixes, which detracts from the theoretical agglutinative ideal. Part of the allomorphy is most conveniently described in terms of item-and-arrangement morphophonological alternations, partly in terms of item-and-process directional rules.

Vowel harmony is an overriding constraint on stems and suffixes. The Finnish vowels form three groups, the harmony vowels /y ö ä/ (front) and /u o a/ (back) plus the neutral vowels /i e/. The three vowel pairs from the harmony sets are often denoted by morphophonemic symbols: U, O, A. Vowels from the front and back harmony sets cannot co-occur in native words whose vowels are drawn either from /i e ä y ö/ or /i e u o a/. Suffixes with harmony vowels have

one front and one back variant occurring after front and back stems, respectively. Stems with neutral vowels only count as front. Thus:

talo-ssa
house-SG.INE
'in (a/the) house'
kylä-ssä
village-SG.INE
'in village'
venee-ssä
boat-SG.INE
'in boat'
auto-lla
car-SG.ADESS
'by car'
pyörä-llä
bike-SG.ADESS
'by bike'
veitse-llä
knife-SG.ADESS
'with knife'
tule-vat
come-PRES.INDIC.3.PL
'(they) come'
määrää-vät
decide-PRES.INDIC.3.PL
'(they) decide'
mene-vät
go-PRES.INDIC.3.PL
'(they) go'
anne-ta-an
give-INDEF-4
hala-ta-an
embrace-INDEF-4

Table 3 Nonfinite verb forms

Stem	Indef	Nonfinite	Number	Case	Possessive	Clitic
sano		a				
say		INF				
'to say'						
sano		a		kse	ni	
say		INF		TRANSLV	POSS.1.SG	
'in order for me to say'						
sano		e		ssa	si	ko
say		INF		INE	POSS.2.SG	Q
'when you are saying?'						
sano	tta	e		ssa		kin
say	INDEF	INF		INE		also
'also when one is saying'						
sano		va				
say		PART.PRES				
'saying'						
sano		nee	t			
say		PART.PAST	PL			
'said (pl.)'						
sano		v	i	en		
say		PART.PRES	PL	GEN		
'of the saying (ones)'						

Table 4 Inflectional classes of nouns

Class	SG	SG.GEN	SG.PART	PL-PART
(1)	lasi <i>glass</i>	lasi-n	lasi-a	lase-j-a
(2)	ovi <i>door</i>	ove-n	ove-a	ov-i-a
(3)	käsi <i>hand</i>	käde-n	kät-tä	käs-i-ä
(4)	nalle <i>bear</i>	nalle-n	nalle-a	nalle-j-a
(5)	vaje <i>lack</i>	vajee-n	vaje-tta	vaje-i-ta
(6)	varis <i>crow</i>	varikse-n	varis-ta	variks-i-a
(7)	vieras <i>guest</i>	vieraa-n	vieras-ta	viera-i-ta
(8)	rakka-us <i>love-NZ</i>	rakka-ude-n	rakka-ut-ta	rakka-uks-i-a

Vowel harmony is basically a phonological phenomenon. There are tens of other, more strongly morphologically conditioned alternations. The most profound one is consonant gradation, which concerns both nominals and verbs. Under complicated phonological and (partly opaque) morphological conditions, the long voiceless stops *pp*, *tt*, *kk* (which constitute two-phoneme combinations) are shortened to *p*, *t*, *k*, and the short voiceless stops *p*, *t*, *k* are weakened in various ways, e.g., *p* → *m* (after *m*), *p* → *v* (between vowels), *t* → *d* (between vowels), *t* → *l*, *r*, *n* (after an identical consonant), *k* → *ŋ* (after *ŋ*), *k* → zero (between vowels). The weak grade is triggered by suffixation processes, in particular the occurrence of a suffix closing the syllable in the beginning of which the strong grade occurs (Table 6).

However, the weak grade also occurs in many purely morphological contexts without suffixes closing the stem syllable, e.g., in certain imperative and indefinite verb forms and in the base forms (nominative singulars) of nouns belonging to inflectional class 5: *ker.ro* tell.IMP.2.SG, *ker.ro-ta-an* tell-INDEF-4, *sa.de* rain.SG.NOM, *sa.tee-n* rain-SG.GEN.

About 20% of the Finnish vocabulary is subject to consonant gradation: 8000 nouns, 1000 adjectives, and 6000 verbs. Of the 1000 most frequent words, 30% participate in consonant gradation. This is a profound characteristic of Finnish.

Another set of typical morphophonological alternations are the vowel mutations in front of certain suffixes starting with *Bi*. Long stem vowels are shortened, stem diphthongs are simplified, stem-final short *Bi* changes to *Be*, stem-final short *Ba* might change to *Bo*, etc.:

maa
country.SG.NOM
ma-i-ssa
country-PL-INE
'in (the) countries'
tie
road.SG.NOM,
te-i-llä
road-PL-ADESS
'on (the) roads'
lasi
glass.SG.NOM
lase-i-ssa
glass-PL-INE
'in (the) glasses'
pila
joke.SG.NOM
pilo-i-ssa
joke-PL-INE
'in (the) jokes'

Many word forms simultaneously display several alternations, e.g., *virka* job.SG.NOM, *viro-i-ssa* job-PL-INE 'in (the) jobs' (gradation and mutation). The nominal inflectional class 3 is particularly complex. On top of gradation and mutation (deletion of *Be*) there is assibilation of stem-internal *BtB* in SG.NOM and deletion of *Be* also in SG.PART: *käsi* hand.SG.NOM, *käte-en* hand-SG.ILL 'into (a/the) hand', *käde-llä* hand-SG.ADESS 'with (a/the) hand', *kät-tä* hand-SG.PART 'hand' (e.g., as direct object in negated clauses), *käs-i-ä* hand-PL-PART '(some indefinite) hands'.

Morphophonological Alternations in Suffixes

Suffixes might lose their final consonant in front of a consonant starting the next ending, or assimilate their first consonant to the last consonant of the preceding stem. Consonants are deleted in front of possessive suffixes: *maa-han* country-SG.ILL 'into (a/the) country', *maa-ha-mme* country-SG.ILL-POSS.1.PL 'into our country', *talo-n* house-SG.GEN, *talo-t* house-PL.NOM, *talo-mme* house-SG.NOM.POSS.1.PL ~ *house-SG.GEN.POSS.1.PL* ~ *house-PL.NOM.POSS.1.PL*. Note the three-way ambiguity arising due to consonant deletion in a form like *talo-mme* expressing nominative singular, genitive singular, and nominative plural.

The illative case has an extreme number of allomorphs. It has four basic suppletively related allomorphs, 'quasimorphemes,' occurring in different phonologically determined contexts: *BVu*, *BbVu*, *Bseen*, *Bsiin*. The morphophoneme *V* is realized by reduplication as a copy of the preceding vowel. In

Table 5 Inflectional classes of verbs

Class	INF.NOM	PRES.INDIC.1.SG	PAST.3.SG	PAST.PART.SG.NOM
(1)	sano-a say	sano-n	sano-i	sano-nut
	anta-a give	anna-n	anto-i	anta-nut
(2)	hala-ta embrace	halaa-n	hala-si	halan-nut
	kara-ta escape	karkaa-n	karka-si	karan-nut
(3)	saa-da get	saa-n	sa-i	saa-nut
	haravoi-da harvest	haravoi-n	haravo-i	haravoi-nut
(4)	nous-ta rise	nouse-n	nous-i	nous-sut
(5)	lämme-tä warm up	lämpene-n	lämpen-i	lämmen-nyt

Table 6 Consonant gradation

kaup.pa shop.SG.NOM	kau.pa-n shop-SG.GEN	kau.pa-s.sa shop-SG.INE	kauppa-a shop-SG.PART
sil.ta bridge.SG.NOM	sil.la-n ridge-SG.GEN	sil.la-i.la ge-SG.ADESS	sil.ta-an ge-SG.ILL
ker.to-o tell-PRES.INDIC.3.SG	ker.ro-n tell-PRES.INDIC.1.SG	ker.ro-i-m.me tell-PAST-1.PL	ker.to-.vat tell-PRES.INDIC.3.PL

Syllable boundaries are indicated by periods.

addition, the final consonant may be deleted before possessives. Therefore the illative has no fewer than 36 allomorphs:

talo-on
house-SG.ILL
'into (a/the) house'
talo-o-mme
house-SG.ILL-POSS.1.PL
lasi-in
glass-SG.ILL
lasi-i-mme
glass-SG.ILL-POSS.1.PL
maa-han
country-SG.ILL
maa-ha-mme
country-SG.ILL-POSS.1.PL
puu-hun
tree-SG.ILL
puu-hu-mme
tree-SG.ILL-POSS.1.PL, etc.

Many suffixes have several allomorphs, each linked to specific stem selection criteria and/or inflectional classes. For example, the partitive allomorph *BtA*

only goes with nominal inflection class 5, *BtA* with classes 3, 6, 7, 8, and *BA* with classes 1, 2, and 4.

Fusion and Polyfunctional Suffixes

There are several polyfunctional suffixes where the grammatical functions cannot be segmented, e.g., *Bt* (nominative plural), *Bten* (genitive plural, the segmentable genitive plural has the plural marker *Bi*), *Bine* (comitative singular or plural), *-ttU* (indefinite fourth person perfect participle), *BkAA* (imperative 2.PL).

In colloquial Finnish, there is a tendency to drop *Bi* in unstressed diphthongs. When this *Bi* is the past tense marker, the mutated stem vowel becomes the only marker of the past tense function: *anta-a* give-PRES.INDIC.3.SG, *anto-i* give-PAST.3.SG, (colloquial) *anto* give-PAST.3.SG.

Conclusion

Karlsson (1983) distinguished 45 different morphophonological alternations. They create massive

allomorphy in both stems and endings. A two-way dependency exists between many stems and endings: certain stems take certain endings only, and certain endings go only with certain stems. This mutual boundedness implies that Finnish word forms are highly cohesive, a property that is amplified by vowel harmony stretching over the whole (uncompounded) word, and further amplified also by the fixed initial stress.

Thus, Finnish departs from the theoretically ideal agglutinative type in some respects, as regards the occurrence of nominal and verbal inflectional classes, allomorphy among the affixes, morphophonological alternations, endings expressing composite grammatical functions, and fusion of certain grammatical elements.

See also: Arabic as an Introflecting Language; Chinese as an Isolating Language; Finnish; Morphological Typology.

Bibliography

- Hakulinen A, Vilkuna M, Korhonen R, Koivisto V, Heinonen T & Alho I (2004). *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Karlsson F (1983). *Suomen kielen äänne- ja muotorakenne*. Juva: WSOY.
- Karlsson F (2003). *Finnish: an essential grammar*. London/ New York: Routledge.
- Paunonen H (1976). 'Allomorfiin dynamiikkaa.' *Virittäjä* 80, 102–117.
- Wiik K (1967). *Suomen kielen morfofonemiikkaa*. Turku: Turun yliopiston fonetiikan laitoksen julkaisuja 3.

Finnish Lexicography

K Varantola, University of Tampere, Tampere, Finland

© 2006 Elsevier Ltd. All rights reserved.

Finnish lexicography can be discussed both as a language policy issue and as a commercial enterprise. The Research Institute for the Languages of Finland, KOTUS, is the national organization responsible for the languages spoken in Finland: Finnish, Swedish, and Saami. Five million people speak Finnish, 300 000 speak Swedish, and about 5000 speak Saami. Finnish belongs to the family of Finno-Ugric languages, of which Estonian is a closely related language and Hungarian, with about 15 million speakers, the biggest.

The dictionary projects undertaken by KOTUS can be described as academic projects of linguistic and cultural importance. Modern monolingual and bilingual dictionaries of Finnish naturally also have commercial value and are produced jointly with commercial partners. Recent achievements include the *Basic dictionary of the Finnish language* (BDFL) in three volumes (from 1994). The production and maintenance methods of this dictionary are traditional. Its predecessor, the comprehensive *Dictionary of modern Finnish*, was published in the years 1951–1961.

An updated version of the BDFL was converted in 1997 to an electronic format. This dictionary has about 100 000 entries. The latest electronic CD-ROM, Internet, and intranet edition of the BDFL was published in 2005 by Kielikone Oy, the major publisher of electronic dictionaries in Finland. No

print updates are planned for the time being. Kielikone publications also include the electronic MOT dictionary library, which consists of a number of monolingual and bilingual dictionaries, with and without Finnish. The rights for the electronic versions have been purchased from various European print dictionary publishers.

The recent academically oriented products of KOTUS include the dictionaries of *Finnish dialects* (seven volumes, A to K, published), *Swedish dialects in Finland* (three volumes, A–K, published), *Old literary Finnish* (two volumes, A–K, published), and dictionaries of Karelian, Vogul (Mansi), and Cheremis (Mari). New dictionaries with more appeal to the general public include the *Finnish-Swedish general dictionary*, the *Basic dictionary of the Finnish Sign Language*, the *Finno-Swedish dictionary* (describing the Swedish spoken in Finland), and a Finnish-Estonian dictionary in two parts. This modern dictionary has more than 90 000 entry words and was compiled and published in 2003 in cooperation with the Institute of the Estonian Language in Tallinn.

Commercial lexicography in Finland is practiced in a challenging environment. Print dictionaries are expensive, because the potential market is small and basically restricted to Finnish-speaking audiences. Two major publishers, WSOY and Gummerus Oy, have accepted the challenge and produce a majority of the dictionaries on the market. WSOY has over the years published a number of large bilingual dictionaries between Finnish and English, Swedish, German, Spanish, and Russian. The update intervals tend to be long and vary from one language pair to another. In