

Finnish Annotation Guidelines

Annotational symbols

- - (hyphen) – derivational break, e.g. myy+dä => myy-jä
- + (plus) – inflectional break, e.g. myy+dä ~ myy+n, kirj-a ~ kirj-a+n
- . (one dot) – questionable break, e.g. hank.o ‘pitchfork’ = hanko or hank-o
- / – break between parts of a compound
- **Removed //** – break before a word form grammaticalized as a derivational element
- ' - **NEEDS EXPLANATION**
- !!! – specific cases to be discussed

[Annotational symbols](#)

[Dilemmas and indices](#)

[Derivation](#)

[Verbal inflection](#)

[Nominal inflection](#)

Dilemmas and indices

Derivation

- **A 2-way: options 0,1** questionable derivational break separating an element that may be either an independent derivational suffix or a part of a root / basic stem.

General rules for putting 1 as opposed to hyphen (-), i.e. the non-questionable derivational break, are now formulated as follows.

A derivational morpheme is separated with the non-questionable break (-), if it is present as a derivational suffix in a word that is formed directly from another really existing word as a result of a regular derivational process, with derivative and non-derivative / basic forms being clearly identifiable:

alka+a ‘begin’ => alk-u ‘beginning’

kuiva ‘dry’ => kuiv-u+a ‘become dry, dry out’

teh-dä ‘do’ => tek-o ‘act, deed’

varma ‘certain’ => varm-uus ‘certainty’

kasva+a ‘grow’ => kasva-tt-a+a ‘educate’ => kasva-t-us ‘education’

The verbal markers -a- and -u-, the latter of which regularly/mostly serves to form intransitive/reflexive verbs, are separated only in cases when they appear in a derivative. If there is a pair of verbs differing only by the presence of -a- vs. -u-, both elements are separated with a hyphen, e.g. kaat-a+a ‘fall’ ~ kaat-u+a ‘fell, overturn, spill’. Verb-forming -a- or -u- following a derivational suffix is always separated with a hyphen, no matter if there is an attested pair or not, because in these cases it is clear that the vowel in question is neither a part of the root nor that of the preceding suffix (e.g., the verb-forming suffix -tt- followed by -a- can’t be merged with it into a single morpheme, because this suffix can be followed by -u- as well, cf. hahmo ‘figure, shape’ => hahmott-a+a ‘perceive, understand the outline of something’ and hahmo-tt-u+a ‘take shape’).

A suffixal morpheme is separated with the questionable break (.) with the index 1, if it is present in a word

that obviously has the same basic stem as another one (or in a word that shares the basic stem with a set of other words) and is attested as a derivational suffix, but in the word in question can't be clearly identified in terms of derivational type and/or function:

arv.o 'worth, value', cf. arv-io 'assessment', arv-el-la 'believe, suppose'

(it is not quite clear how to analyze -o in arv.o: there exists a derivational suffix -o, but it mainly serves to form nouns from verbs, not from nouns)

san.a 'word' ~ san.o+a 'say'

a rare type of derivational relationship that doesn't allow (at least synchronically) to establish with certainty, which word is derivative and which is non-derivative.

The questionable break is also used for a morpheme, if it is productive as a word-forming suffix, has a regular derivational function and is clearly related to a certain type of derivational process, being identifiable as representing this type in the word in question and corresponding to this word semantically/syntactically, but, in this word, implies the extraction of a root morpheme that is not attested without this suffix or can not be unambiguously identified with a homonymous root:

odo.tt-a+a 'wait'

suu.tt-u+a 'become angry' (suu- = suu 'mouth'?)

väs.y-ä 'become tired'

The rest of indices for the questionable derivational break are used to mark a morpheme that

- **a)** can be considered either as an inseparable part of a suffix or as a separate derivational morpheme (in particular within a suffixal cluster)
- **b)** may be either a part of a root or a part of a suffix
- **c)** can belong to one of two "neighbouring" suffixes
- **B 4-way: options 2,1** amme 'tub' => ammentaa 'draw, scoop': amme-nt-a+a or amm-ent-a+a,
 - ■ hede 'stamen' => hedelmä 'fruit': hede-l-mä or hed-el-mä
 - murhe 'sorrow' => murehtia:
 - mure-hti+a or
 - mur-ehti+a
- **C 4-way: options 2,1** halpa 'cheap' => halveksia 'contemn': halpa =>
 - halve-ksi+a or
 - halv-eksi+a
- **D 4-way: options 2,1** ylpeä 'proud':
 - ylpe-ä or
 - ylp-eä
- **E 4-way: options 2,1** työskennellä 'work':
 - työ-ske-nn-el+lä or
 - työ-sk-enn-el+lä
- **F 4-way: options 2,1** epäröidä 'hesitate':
 - epärö-i+dä or

- epär-öi+dä
- **G 4-way: options 2,1** kulkija 'traveler, rover': kulki-ja or kulk-ija
 - palvelija 'servant': palv-eli-ja or palv-el-ija
 - vangitsija 'jailer':
 - vang-itsi-ja or
 - vang-its-ija
- **H 2-way: options 0,1** asua 'reside' => asunto 'apartment':
 - (because there are other deverbative noun-forming suffixes consisting of -nt- and a subsequent vowel: -nta, -nti; I think all decisions of that type may be united under one index, for now 8)
 - asu-nto or
 - asu-nt-o
- **I 2-way: options 0,1** suuttua 'be angry' => suuttumus 'anger':
 - (because of the existence of the separate deverbative noun-forming suffix -us/-ys);
 - suu.tt-u-mus or
 - suu.tt-u-m-us
 - antaa => antaminen: ant-a-min.e.n or ant-a-m-i.n.e.n – here the question is whether -minen is one morpheme or a morpheme cluster consisting again of -m- and the element -i.n.e.n that is attested as a separate suffix (for the markup in -i.n.e.n see the following paragraph as well as Nominal inflection, class b)
- **J 2-way: options 0,1** varhainen ~ varhaisen 'early':
 - varha-in.e.n or
 - varha-i-n.e.n,
 - varha-is.e.n or
 - varha-i-s.e.n,
 - suomalainen ~ suomalaisen: suom-a-la-in.e.n or suom-a-la-i-n.e.n, suom-a-la-is.e.n or suom-a-la-i-s.e.n (because we have also the element -nen/-se without preceding -i-, cf. tikkanen ~ tikkasen 'woodpecker'); the same index also for respective derivatives (-is-uus/-is-yy or -i-s-uus/-i-s-yy)
- **K 2-way: options 0,1** ilmoittaa 'announce, inform':
 - ilmo-itt-a+a or
 - ilmo-i-tt-a+a
 - (because we have also verb-forming -tt- without preceding -i-, cf. hahmo 'shape, figure' => hahmottaa 'understand the outline of something'); the same index is used also for another verb-forming suffix demonstrating an analogous dilemma: -ist- or -i-st- (kukka 'flower' => kukoistaa 'bloom', but cf. similar derivation without -i-: raha 'money' => rahastaa 'collect money; cash in' with no -i-); the same index also for respective frequentative forms (-it-el+la or -i-t-el+la)
- **L** (deleted)
- **M 2-way: options 0,1** toinen 'second, another'
 - toi-n.e.n or

- to-i.n.e.n
- **N** ~~4-way, options 0,1,3~~ seems to be 2-way in data juhla 'holiday' => juhlallinen 'festive, solemn':
 - juhl-a-llin.e.n or
 - juhl-a-lli-n.e.n or
 - juhl-a-ll-i.n.e.n
- **O** 2-way: kauppa 'trade' => kauppias 'trader': kaupp-ias or kaupp-i-as (because -as is attested as a separate suffix, e.g. oppi 'study, knowledge' – opas 'guide')
- **P** 4-way, two options: 1 or 2
 - a) consonantal stems: (for words derived from verbal stems that may be considered as ending in a consonant)
 - palvella 'serve' => palveleminen: palv-ele-m.i.n.e.n or palv-el-em.i.n.e.n (cf. Verbal inflection, class C)
 - kuolla 'die' => kuolettaa 'amortize' kuole-tt-a+a or kuol-ett-a+a
 - kuolla 'die' => kuolema 'death' kuole-ma or kuol-ema
 - b) [oppi 'knowledge, study'] => opettaa 'teach' ope-tt-a+a or op-ett-a+a
- **R** 4-way, two options: 1 or 2
 - tapahtua 'happen': tapa-ht-u+a or tap-aht-u+a
 - masentua 'become depressed': mase-nt-u+a or mas-ent-u+a
- **Z**

Verbal inflection

- **t** 4-way, two options: 1 or 2
 - panna 'put' ~ pannessa: pan+ne+ssa or pan+n+essa,
 - poistua 'quit' ~ poistuesssa: pois-tu+e+ssa or pois-tu+essa
- **u** 4-way, two options: 1 or 3
 - antaa 'give' ~ annetaan: ann-e+t+a+an or ann+et+a+an
- **v => P** 4-way, two options: 1 or 2
 - (for verbal stems that may be considered as ending in a consonant):
 - palvella 'serve' ~ palvelen: palv-ele+n or palv-el+en
- **w** 2-way
 - holhota 'patronize' ~ holhoan ~ holhotkaa: holhot+kaa or holho+t+kaa
- **x** 2-way
 - upota 'sink' ~ upposi: uppo+si or uppo+s+i

Nominal inflection

- (a, b, c – for nominal stems that may be considered as ending in a consonant)

- **a** 4-way, two options: 1 or 2
 - rakkaus 'love' ~ rakkautta ~ rakkauden: rakka'-ude+n or rakka'-ud+en
 - suomalainen 'Finn, Finnish' ~ suomalaista ~ suomalaisen: suom-a-la-i.se+n or suom-a-la-i.s+en
 - lanne 'hip, haunch' ~ lannetta ~ lanteen: lant.ee+n or lant.e+en
- **b** 4-way: 0,1,2
 - suomalainen: suom-a-la-i.ne+n or suom-a-la-i.n+en or suom-a-la-i.nen
- **c** 4-way, two options: 1 or 2
 - sopimaton 'unfit, improper' ~ sopimatonta ~ sopimattoman: sop-i-ma-ttoma+n or sop-i-ma-ttom+an;
 - vanhempi 'older' ~ vanhemman: vanh.e.mma+n or vanh.e.mm+an;
 - vanhin 'the oldest' ~ vanhinta ~ vanhimman: vanh+i.mma+n or vanh+i.mm+an
 - arvokas 'precious; dignified; expensive' ~ arvokasta ~ arvokkaan: arv.o-kkaa+n or arv.o-kka+an
- **d** 4-way, two options: 1 or 2
 - laki 'law' ~ lakeja lake+j+a or lak+ej+a
- **e** 4-way, two options: 1 or 2
 - vanha 'old' ~ vanhempi vanhe+mpi or vanh+empi
- **f** 2-way
 - vanhin ~ vanhimman vanh+in, vanh+imm.a.n or vanh+i+n, vanh+i+mm.a.n
- **y** 2-way
 - for -s.en, -s.essa,...
 - (might go back to join class **a** in the future) because it may be reasonable to allow -en, -essa, etc. as allomorphs following **all consonantal** stems.

Copyright © by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? [Send feedback](#)

