

Derivational Networks in Basque

Xabier Artiagoitia
(UPV/EHU)

[March 2018, prepublication version]

1. General notes

Compounding and derivation are the usual word formation strategies in Basque, but the Basque lexicon has historically relied a great deal on borrowing (from neighboring Romance languages, directly from Latin in earlier times); as of today, neoclassical formations in Basque usually take the same form as in Spanish. Derivation in Basque is done mainly by suffixes (Artiagoitia, Hualde and Ortiz de Urbina 2016), hardly ever by prefixes, which are just a few, borrowed from Romance (cf. *des-*; *eman* ‘give’ > *deseman* ‘undo, cancel’) or calqued from it (i.e. *ezaguna* ‘known’ > *ezezaguna* ‘unknown’; *ez* ‘not’); the prefix *ber-* (variant *bir-*) ‘re-’ might be the only autochthonous one (de Rijk 2008: 891); e.g. *eman* ‘give’ > *berreman* ‘give again’. It is also the case that all verb formation or derivation is done through conversion, a process that is highly productive for nouns, adjectives and nouns bearing the allative adposition; cf. *ur_N* ‘water’ > *urtu_V* ‘to melt’, *mehe_A* ‘thin’ > *mehetu_V* ‘to get thin/make something thin’; *begietara_{PP}* ‘to the eyes’ > *begietaratu_V* ‘to come/bring to the eyes’; *-tu* is the participial ending and also the citation form for verbs. Furthermore, conversion may affect comparative adjectives (i.e. it is possible after inflection) or quantifiers; e.g. from *mehe_A* ‘thin’, *meheago_A* ‘thinner’ > *mehegotu_V* ‘to become thinner/make something thinner’; from *gutxi_Q* ‘little’ > *gutxitu_V* ‘to diminish’; from comparative *gutxiago_Q* ‘less’ > *gutxiagotu_V* ‘to become less, lessen’.

Another factor that may be relevant for our research is that of the separation between compounding and derivation: in the Basque grammar tradition the usual decisive criterion for discriminating between a suffix and a word is whether the second member of a compound can stand independently as a word. This sweeps away some potential suffixes which are words in present Basque even though their meaning as independent words has little to do with their use as bound morphemes; e.g. *muga_N* ‘border’, *zain_{N/ADV}* ‘guard’ and ‘waiting’, *mugazain_N* ‘border guard’, with *zain* hardly ever used as ‘guard’ in present Basque. Consequently, I have not considered suffixes second members of compounds that exist as independent words in the dictionary of the Basque Academy (Euskaltzaindia 2016), with the exception of *-tasun* ‘-ness’ (e.g. *zuzen_A* ‘straight’ > *zuzentasun_N* ‘straightness, justice’), a well-known suffix to have been promoted as an independent word in the 20th century with the meaning ‘feature’, and *-tegi* (*su_N*

‘fire’ > *sutegi*_N ‘forge, fireplace’), the meaning of which as a free word (‘barn, storage room’) is dialectally reduced. Conversely, suffixes which were independent words at some stage in the language but do not exist as such have been included, even though the Basque Academy still regards words containing them as compounds (Euskaltzaindia 1994); examples include *-dun* (e.g. *hezur*_N ‘bone’ > *hezurdun*_A ‘vertebrate’), *-gin* (e.g. *harri*_N ‘stone’ > *hargin*_N ‘stonemason’), or *-gile* (e.g. *su*_N ‘fire’ > *sugile*_N ‘firemaker’). The result of these considerations is that, out of the 72 derivational suffixes listed in Villasante (1974: 160), 18 have been eliminated in this project; most of these are what de Rijk (2008) calls parasuffixes.

I have drawn the examples for this study from Sarasola’s (2007) ‘Basque Dictionary’, which is itself the basis for the Basque Academy’s dictionary (Euskaltzaindia 2016), and from Sarasola’s ongoing dictionary, based on a 21st century database (Corpus of Contemporary Basque). All examples are attested words, not possible words.¹

2. Maximum derivational networks in Basque

Remarkably so, Basque never gets beyond the 3rd order of derivation, even the 2nd order in the case of verbs. Derivational networks only seem to be strong for the 1st order.

	1st order	2nd order	3rd order	Σ
Nouns	36	18	2	56
Verbs	36	12	0	48
Adjectives	30	5	5	40
TOTAL	102	35	7	144

Table 1. Maximum derivational networks per order of derivation for all the three word-classes.

Adjectives have the lowest number of derivatives yet they are balanced between 2nd and 3rd order derivatives; however all the 3rd order derivatives correspond to *berri* ‘new’.

3. Saturation values

The mean saturation value for nouns ranges between %30 and %7, as displayed in Table 2. The noun with the lowest saturation value is *zorri* ‘louse’, which also happens to be one of the three

¹ I have avoided using the Basque General Dictionary (Mitzelena & Sarasola 1987-2011) because it provides the history of every single written Basque word, regardless of whether it has been used in the literature or whether it is the sole creation of a dictionary writer. Sarasola’s (2007) dictionary, on the other hand, is based on words actually used by writers of all times and, thus, I find it a more reliable source.

nouns with no 2nd order, together with *hezur* ‘bone’ and *txakur* ‘dog’. The noun with the highest saturation value is *begi* ‘eye’; this and *ur* ‘water’ (%25) are the only ones that get to the 3rd order.

Nouns		Saturation value (%)	1st order (%)	2nd order (%)	3rd order (%)
<i>hezur</i>	‘bone’	17,86	27,78	0	0
<i>begi</i>	‘eye’	30,36	27,78	33,33	50
<i>hartz</i>	‘tooth’	21,43	22,22	22,22	0
<i>egun</i>	‘day’	16,07	16,67	16,67	0
<i>txakur</i>	‘dog’	16,07	25	0	0
<i>zorri</i>	‘louse’	7,14	11,11	0	0
<i>su</i>	‘fire’	17,86	22,22	11,11	0
<i>harri</i>	‘stone’	28,57	33,33	22,22	0
<i>ur</i>	‘water’	25	19,44	33,33	50
<i>izen</i>	‘name’	17,86	13,89	27,78	0

Table 2: Saturation values per order of derivation, nouns

With respect to verbs, no verb reaches the 3rd order of derivation, and three verbs fail to reach the 2nd order. The highest and lowest mean saturation values oscilate between 39% for *jakin* ‘know’ and 8% for *induskatu* ‘dig’; the former has the highest values in 1st and 2nd order.

Verbs		Saturation value (%)	1st order (%)	2nd order (%)
<i>ebaki</i>	‘cut’	16,67	22,22	0
<i>induskatu</i>	‘dig’	8,33	11,11	0
<i>erakarri</i>	‘pull’	20,83	22,22	16,67
<i>bota</i>	‘throw’	10,42	13,89	0
<i>eman</i>	‘give’	22,92	27,78	8,33
<i>eutsi</i>	‘hold’	20,83	22,22	16,67
<i>josi</i>	‘sew’	16,67	13,89	25
<i>erre</i>	‘burn’	20,83	22,22	16,67
<i>edan</i>	‘drink’	35,42	33,33	41,67
<i>jakin</i>	‘know’	39,58	36,11	50

Table 3: Saturation values per order of derivation, verbs

As regards adjectives in Table 4, just four adjectives get to the 2nd order of derivation and only the adjective *berri* ‘new’ reaches the 3rd order; this is precisely the adjective with the highest mean saturation value, i.e. 45%; the lowest value of 10% corresponds to *txar* ‘bad’.

Adjectives		Saturation value (%)	1st order (%)	2nd order (%)	3rd order (%)
<i>estu</i>	‘narrow’	15	20	0	0
<i>zahar</i>	‘old’	27,5	36,67	0	0
<i>zuzen</i>	‘straight’	22,5	26,67	20	0
<i>berri</i>	‘new’	45	33,33	60	100
<i>luze</i>	‘long’	32,5	36,67	40	0
<i>epel</i>	‘warm’	12,5	13,33	20	0
<i>lodi</i>	‘thick’	12,5	16,67	0	0
<i>txar</i>	‘bad’	10	13,33	0	0
<i>mehe</i>	‘thin’	12,5	16,67	0	0
<i>beltz</i>	‘black’	17,5	23,33	0	0

Table 4: Saturation values per order of derivation, adjectives

Table 5 summarizes the average saturation values for the three categories. The average saturation values are very similar in the first order, and the same is true of the 2nd order of derivation. For the 3rd order, the average saturation values are identical for nouns and adjectives.

	1st order	2nd order	3rd order
Nouns	21,94	16,66	10
Verbs	22,50	17,5	0
Adjectives	23,66	14	10

Table 5: Average saturation values per order of derivation of all three word-classes

The conclusion that one draws upon looking into derivation networks in Basque is that derivational networks are not very strong cross-categorially in the language and, furthermore, that they have a balanced behaviour for all three categories.

4. Orders of derivation

As for the average and maximum orders of derivation reached in the Basque data the calculations are shown in Table 6. The results are in line with our comments on previous tables regarding the poverty of derivational networks in Basque.

	Maximum	Average
Nouns	3	1,9
Verbs	2	1,6
Adjectives	3	1,5

Table 6. Maximum and average number of orders of derivation for all three word-classes

5. Derivational capacity

The maximum and average for the three categories is relatively balanced, which reinforces the conclusion in section 2 regarding maximum derivational networks.

	Maximum	Average
Nouns	12	7,9
Verbs	13	8,1
Adjectives	11	7,1

Table 7. Maximum and average derivational capacity for all three word-classes

As table 8 shows, nouns have a higher level of derivatives compared to verbs and adjectives in the 2nd order, even though the number is relatively low in any case. The categories seem to converge again for the 3rd order of derivation. The lower value of adjective derivatives for 2nd order might be due to the existence of alternative (and very productive) processes for adjective formation such as affective palatalization (e.g. *zuzen*_A ‘straight’ > *xuxen*_A ‘rather straight’) and reduplication (e.g. *zuzen-zuzen*_A ‘very straight’).

	1st order	2nd order	3rd order
Nouns	7,9	3	0,2
Verbs	8,1	2,1	0
Adjectives	7,1	0,7	0,5

Table 8 Average number of derivatives per order of derivation for all three word-classes

6. Correlation of semantic categories and orders of derivation

With respect to nouns, QUALITY (value 8) is the most characteristic semantic category in the first order, followed by COLLECTIVE and LOCATIVE (both value 6). In the second order, AGENT (value 4) is most characteristic. In the case of QUALITY, the existence of two synonymous

suffixes like *-tsu* and *-dun* is probably a helping factor; e.g. *hezur_N* ‘bone’ > *hezurdun_A* ‘vertebrate’, *hezurtsu_A* ‘bony’.

When it comes to verbs’ derivational networks, AGENT (value 9) and ACTION (value 8) are the most salient semantic categories in the first order, followed by ABILITY and RELATIONAL (value 6 for both categories) and RESULTATIVE (value 5). In the second order, STATE is by far (value 5) the most characteristic semantic category. The fact that most verbs are action verbs and the existence of two synonymous suffixes (which usually alternate depending on the morphonemics of the verb) help explain the high value of AGENT; e.g. *edan_V* ‘drink’ > *edale_{N/A}* ‘drinker’; *errev* ‘burn, smoke’ > *erretzailen_{N/A}* ‘smoker’. The category STATE has a value of 6 thanks to the suffix *-tasun*, which attaches to just any (deverbal) adjective; e.g. *edan_V* ‘drink’ > *edangarri_A* ‘drinkable’ > *edangarritasun_N* ‘drinkability, state of being drinkable’

Regarding adjectives, STATE (value 10) is the most characteristic semantic category in the first order, followed by DIMINUTIVE, MANNER and PEJORATIVE (all with value 6). This has to do with the fact that the suffix *-tasun* ‘ness’ (e.g. *zahar_A* ‘old’ > *zahartasun_N* ‘oldness’) is again productive for virtually all adjectives; similar considerations apply to the suffixes *-txo* (*zahar_A* ‘old’ > *zahartxo_A* ‘rather old’), *-ki* (*zuzen_A* ‘straight’ > *zuzenki_{ADV}* ‘straightly’) and *-keria* (*zahar_A* ‘old’ > *zaharkeria_N* ‘a despicable old thing’). No clear correlations can be established for 2nd and 3rd order.

7. Semantic categories with blocking effects

It is difficult to figure out any blocking effect for specific semantic categories given that Basque derivational networks hardly ever reach the 3rd order. For adjectives, the categories AUGMENTATIVE, DIMINUTIVE, PEJORATIVE and STATE in the first order block further derivation. A similar claim can be made of the same categories with respect to nouns, with the addition of the category ACTION. In the case of verbs, the category ACTION also blocks further derivation, unlike most of the semantic categories involved.

8. Typical combinations of semantic categories

Basque derivational networks are relatively short and, furthermore, one may hardly establish any typical combination of semantic categories; combinations never happen more than three times (e.g. CAUSATIVE + ABILITY/ACTION/AGENT, ABILITY + STATE, RELATIONAL + STATE, and MANNER + MANNER).

9. Multiple occurrences of semantic categories

There are 3 cases of MANNER occurring twice in a single derivational chain due to the possibility of the adverbial suffixes *-ki* and *-ro* to combine with one another: e.g. *epel*_A ‘warm’ > *epelki*_{ADV} ‘warmly’ > *epelkiro*_{ADV} ‘warmly’; *zuzen*_A ‘straight’ > *zuzenki*_{ADV} ‘straightly’ > *zuzenkiro*_{ADV} ‘straightly’. In the case of *berri*_A ‘new’ > *berriro*_{ADV} ‘newly, again’ > *berriroki*_{ADV} ‘newly, again’, the order of the manner suffixes is reversed.

10. Reversibility of semantic categories

There is no instance of reverse order of semantic categories, given that Basque networks only get to 3rd order derivations in a very limited number of cases.

11. Poor derivational networks in Basque

As explained in the first section, the poverty of Basque derivational networks is partly accounted for by the high productivity of X > V conversion processes. The existence of alternative ways aside from derivational morphemes proper, such as affective palatalization, reduplication, or even compounding, to form nouns and adjectives may help explain the relatively poorness of Basque derivational networks. Allowing conversion and affective palatalization into the calculation of derivational networks would have undoubtedly produced a higher number of 2nd and 3rd order derivatives, specially for nouns and adjectives.

12. Conclusions

Judging from the sample words taken for this study, Basque appears to have a limited strength in producing derivational networks beyond the 2nd order (Table 1). This tendency is consistent for the three classes of words considered with respect to 1st order derivation (similar average saturation values, similar maximum and average derivational capacity as shown in Table 7).

As can be seen in Table 1, nouns exhibit the highest derivational network (56), with the highest number for the 1st order (tied with verbs) and for the 2nd, but they are outscored by adjectives in the 3rd order. The maximum derivational network for verbs is smaller (48) with no 3rd order derivatives, and adjectives come last (40 derivatives) yet they have the highest

number of 3rd order derivatives. The overall saturation values (Table 5) for the three categories are in any case similar for 1st and 2nd order of derivations; these range from 22% to 23% in 1st order and from 14% to 16% in 2nd order. My impression is that reduplication and affective palatalization is most productive with adjectives, which would compensate for their relatively lower derivational capacity.

As a closing remark, Basque derivations cover 30 of the 49 semantic categories provided for this research. In principle, this fact suggests that Basque must resort to other means, either phonological, morphological or syntactic, to convey the remaining semantic categories. The derivational capacity of the Basque language, however, appears quite robust and balanced among the three categories considered, given the number of categories represented in the sample, yet this robustness does not result in the formation of derivational networks beyond the 3rd order.

References

Artiagoitia, Xabier, José Ignacio Hualde and Jon Ortiz de Urbina 2016 Basque In: P. O. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds) *Word formation. An international handbook of languages of Europe*. Volume 5, Berlin, Mouton de Gruyter, 3327-3348.

Euskaltzaindia [Basque Academy] 1994 *Hitz elkarketa /4. Hitz elkartuen osaera eta idazkera* [Compounding /4. The composition and writing of compound words], Bilbao, Euskaltzaindia

Euskaltzaindia [Basque Academy] 2016 *Euskaltzaindiaren Hiztegia* [The Dictionary of the Basque Academy], Bilbao, Euskaltzaindia.

Hualde, José Ignacio, Jon Ortiz de Urbina (eds.) 2003 *A grammar of Basque*, Berlin, Mouton de Gruyter.

De Rijk, Rudolf P. G. 2008 *Standard Basque: A Progressive Grammar*, Boston, MIT Press.

Mitxelena, Luis, Ibon Sarasola 1987-2011 *Diccionario General Vasco - Orotariko Euskal Hiztegia*, 2nd edition, Bilbao, Euskaltzaindia / Desclée de Brouwer.

Villasante, Luis 1974 Palabras vascas compuestas y derivadas, Oñate, Editorial Franciscana de Aránzazu.

Sources

Euskaltzaindiaren Hiztegia [The Dictionary of the Basque Academy] 2016
<http://www.euskaltzaindia.eus>

EH, Sarasola, Ibon 2007 Euskal Hiztegia [Basque Dictionary] San Sebastian, Elkar.

EEH, Sarasola, Ibon 2018 Egungo Euskararen Hiztegia [The Dictionary of Contemporary Basque] <http://www.ehu.eus/eeh/>

EEC, Sarasola, Ibon, Salaburu, Pello, Josu Landa 2018 Egungo Euskararen Corpusa [Corpus of Contemporary Basque] <https://www.ehu.eus/etc/>