Linguistic Research 32(2), 469-480
DOI: 10.17250/khisli.32.2.201508.008

On variable auxiliary selection in Western Europe*

Cathryn Donohue
(The University of Hong Kong)

Donohue, Cathryn. 2015. On variable auxiliary selection in Western Europe. *Linguistic Research* 32(2), 469-480. A number of Western European languages use two different auxiliary verbs when forming the perfect aspect with monadic verbs, roughly corresponding to *have* and *be*. While such a split is cross-linguistically rare, various studies have shown that there are clear similarities in the patterns of use and factors determining the auxiliary selection. In this paper I focus on the origins of this phenomenon and put forward the proposal that Vasconic may be the source of the dual auxiliary usage, following evidence of language spread through toponyms from Vasconic into the same areas as suggested by Venneman (2003). **(The University of Hong Kong)**

Keywords auxiliary selection, unaccusativity, Basque

1. Perfect auxiliaries

The perfect construction is typically a periphrastic one, consisting of an auxiliary verb and a past participle, as in English *I have eaten*. Different languages use different auxiliaries, but the most common two auxiliary verbs may be roughly translated into English as 'have' and 'be'.[1] A handful of languages, however, make use of both HAVE and BE auxiliaries in monadic predicates.

There has been much work on the nature of the auxiliary split, with many claiming that they correspond to a split between unaccusative (BE) and unergative (HAVE) intransitive verbs (e.g. Sorace 2000; Bentley and Eythórsson 2003; Legendre 2007; McFadden 2007). Indeed, as an example of the kind of careful work that has been carried out investigating this issue, Sorace (2000) proposed the

* Many thanks to Mark Donohue and the audience at the ALS2011 for helpful discussion of earlier versions of this work, as well as to the anonymous reviewers of *Linguistic Research* for their insightful suggestions. All errors remain the sole responsibility of the author.

1 Generic referents to equivalents of 'have' and 'be' are in all capitals (e.g. HAVE, BE).

Auxiliary Selection Hierarchy shown in Figure 1 below. The Auxiliary Selection Hierarchy classifies verbs and orders them in such a way that indicates the differential susceptibility of (monadic) intransitive verbs to variable auxiliary selection. The actual division of the verbs into those taking HAVE and those taking BE when forming the perfect construction in languages which use both is of course unique to each language. However, as suggested by the hierarchy, once a point is chosen, all verbs along the scale above that point will take BE and below that point will take HAVE as their auxiliaries in these split-auxiliary languages.

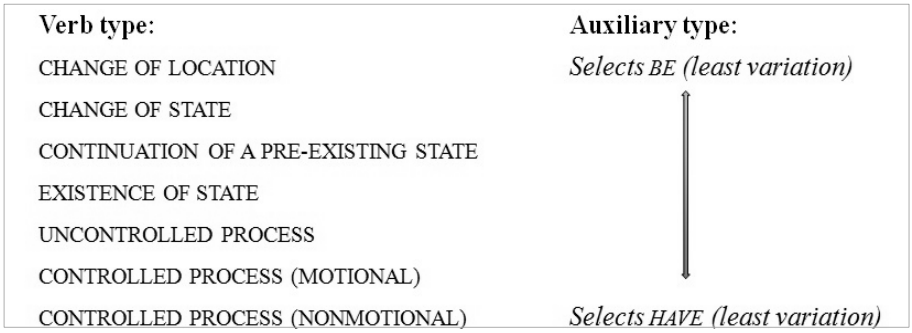

Figure 1. Auxiliary selection hierarchy (Sorace 2000:683)

The following set of examples, taken from Sorace 2000 (pp. 863, 874), illustrates the dual use of auxiliaries in two Romance and two Germanic languages.

(1) **Italian**:

a. Maria **è** venuta alla festa
Maria **is** come to the party
'Maria came to the party.'

b. I colleghi **hanno** chiaccherato tutto il pomeriggio.
the colleagues **have** chatted whole the afternoon
'My colleagues chatted the whole afternoon.'

(2) **French**:

a. Marie **est** arrivée en retard.
Marie **is** arrived late
'Marie arrived late.'

b. Les policiers **ont** travaillé toute la nuit.
the policemen **have** worked whole the night
'The policemen worked all night.'

(3) **Dutch**:

a. De brief **is** met de tweede post gekomen
the letter **is** with the second post arrived
'The letter arrived with the second post.'

b. De trompettist **heeft** met bolle wangen geblazen.
the trumpeter **has** with puffed-out cheeks blown
'The trumpeter blew with puffed-out cheeks.'

(4) **German**:

a. Der Zug **ist** spat angekommen
the train **is** late arrived
'The train arrived late.'

b. Kurt **hat** den ganzen Sonntag gearbeitet.
Kurt **has** the whole Sunday worked
'Kurt worked all day Sunday.'

The (a) examples in (1)–(4) above all use BE for the auxiliary, while the (b) examples use HAVE. Languages vary as to how many predicates take BE and how many take HAVE when expressing the perfect aspect, but Sorace shows that this difference that can be understood according to the Auxiliary Selection Hierarchy, as shown in Figure 1 (Sorace 2000: 683).

Much of the careful examination of languages with dual auxiliary usage has focussed on how to account for the synchronic splits in the auxiliary system, leading

to such advances as the Unaccusativity Hypothesis and related issues of modelling and understanding the contemporary source of the variation (e.g. Perlmutter 1978; Burzio 1986; Levin and Rappaport Hovav 1994, Sorace 2000; Legendre 2007). What is not often discussed at length is that languages that use both HAVE and BE auxiliaries in forming the perfect construction are vanishingly rare, cross-linguistically (e.g. Dahl and Velupillai 2013).[2] They include some (but not all) Romance languages, some (but not all) Germanic languages and Basque.

Table 1. A sample of languages classified by auxiliary selection[3]

Periphrastic perfect constructions in monadic predicates

Split auxiliary: both HAVE & BE		*Single auxiliary used (HAVE or BE)*	
Romance:	**Germanic:**	**Romance:** (HAVE)	**Germanic:** (HAVE)
French	Danish	Portuguese,	English
Italian	Dutch	Sicilian Italian	Icelandic
Romansch	German	Spanish	Swedish
Other:		**Other:** (BE)	
Basque		Bulgarian, Slavic	

A key feature that the languages with the split auxiliary have in common is geographic adjacency: they occupy a small corner of Western Europe as shown in Figure 2.

---

2 Following Bentley and Eythórsson (2003) and others, I leave aside the issue of morphological alternations due to person/number inflections that may confuse the identity of the auxiliary verb.

3 While HAVE is the standardly used auxiliary for Romance/Germanic languages without auxiliary variation, some varieties have been reported to use BE instead (e.g. Terracinese and Shetlandic). However, I leave this aside, as it is not relevant to the focus of this paper.

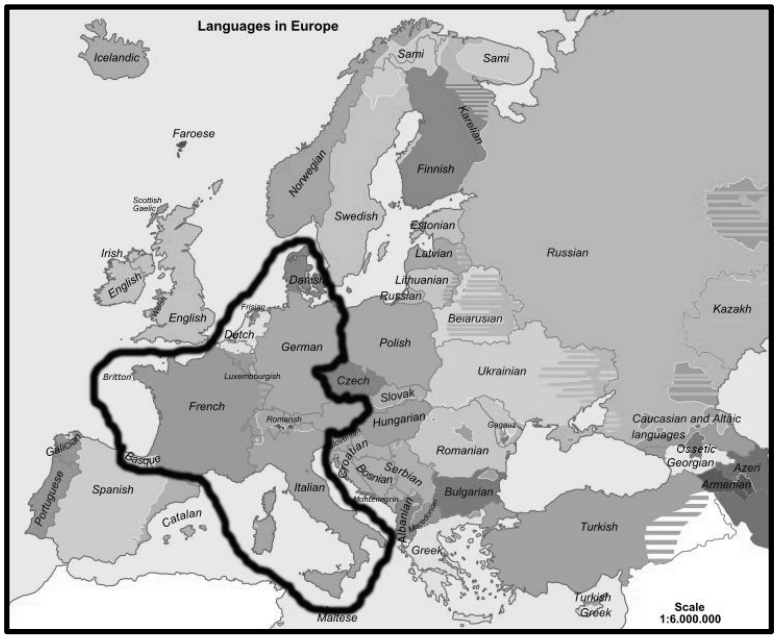

Figure 2. Languages with HAVE/BE in Europe.

Note that Basque is located between France and Spain, on the southwestern edge of the region where HAVE/BE usage is found (Source: original map 'Languages of Europe' from Wikipedia).

The question of how this auxiliary split came to be in the small subset of languages where it is found appears as yet unaddressed. In this paper I propose that Vasconic is a possible source of this typologically rare feature, following Venneman's claims about the linguistic history of Western Europe (e.g. Venneman 2003).

2. HAVE/BE in Basque

Basque is one of the European languages that use two auxiliaries in forming the perfect. It is not typically discussed in the literature on European auxiliary splits, likely because it is not an Indo-European language: it is an ergative isolate with a very different language structure. In a typical ergative system, the subject in a transitive clause bears a case marker (ergative) distinct from the case used to mark the object of a transitive clause and the subject of an intransitive clause (absolutive).

An interesting fact about Basque is that the presence of an ergative case marker correlates exactly with the presence of the HAVE auxiliary (Donohue 2004). Thus, the observed split auxiliary phenomenon is also marked by distinct case arrays: the monadic predicates taking HAVE mark their argument with the ergative case marker; those taking the auxiliary BE mark their argument with the absolutive case. Note that the predicate types conform to Sorace's generalization, and to other semantic-based proposals for understanding the distribution of the different auxiliaries (e.g. van Valin 1990 and others). This split auxiliary usage in Basque monadic predicates is illustrated in the examples given in (5) below (Donohue 2004).

(5) **Basque**:

a. Gizon-a mintzatu **da**.
Man-ABS spoke **BE**
'The man spoke.'

b. Gizon-ak korritu **du**.
Man-ERG ran **HAVE**
'The man ran.'

Basque is known to have calques from Romance languages, such as the verb 'to like' *gustatu* from Spanish *gustar*, which even has the same non-canonical case pattern of marking the subjects with a dative case marker. Perhaps because of this borrowing, it seems to have been generally assumed that the dual auxiliary usage was also a calque.[4] Further, research on Basque split intransitives has tended to focus on explaining the use of the ergative case marker on the sole argument of monadic predicates (e.g. Levin 1983; Donohue 2004). However, with the focus on case marking, the literature has not related Basque to the general question of the HAVE/BE auxiliary alternation elsewhere in Europe.

The current proposal is Vasconic, the modern reflex of which is Basque, is the source of this quirky Western European feature. This is not unreasonable given the work by Aldai (2002) that demonstrates that the perfect in Basque, commonly assumed to have entered the language around the 8th century AD (e.g. Trask 1997

4 Aldai p.c.

and others), was already highly grammaticalized by the 16th century. The primary evidence in support of this proposal comes from two key areas: first, there have been intricate proposals put forward claiming that Vasconic is the source of many other features found in a very similar geographical domain in Western Europe. For this reason, we suggest that the split auxiliary system has had a similar influence on neighbouring languages (section 3). Second, the ergative nature of Basque may support the semantic basis of the auxiliary splits that appear sensitive to notions of agentivity and causation etc., which obviously played a role in auxiliary selection in Basque (section 4).

3. Language spread in Western Europe

Vennemann (2003 and elsewhere) claims that after the last ice age, most of Central and Western Europe was inhabited by speakers of Vasconic languages, the only contemporary survivor of which is Basque. These speakers are said to have formed a substrate to the later-arriving Indo-Europeans. The primary evidence for the presence of Vasconic throughout much of Europe is drawn from the parallelisms in Old European hydronyms and toponyms, many of which were originally identified by Krahe (1964) as Indo-European and reanalyzed by Vennemann as Vasconic. Krahe worked on the reconstruction of the oldest European river-names and place-names and assumed an Indo-European origin. Vennemann's re-analysis of the reconstruction of the data suggests a "non-Indo-European agglutinating and suffix-prevocalizing source language with initial accent and without vowel quantity" (2003:139). Vennemann observes that the structure of the reconstructed language shares substantive similarities with Basque. Basque is thus considered a descendent of the Old European language family, named Vasconic. Without going into the finer details of the phonological reconstructions, let us consider where hydronyms and toponyms meet: places named after the rivers they lie on/near which are based on the underlying root *iz-* 'water' (Basque). Vennemann lists a number of these in different languages (2003: 860–861):

> The French Isère, an old *Isara,* the same name as the river on which Munich is situated, the Isar, earlier also attested as *Isara*. Other names based on the same first element include:

i. Yser (*Isara*) in Belgium
ii. Ijssel (*Isala*) in Holland
iii. Ijsendoorn (*Isandra*) in Holland
iv. Ure (*Isura*) in England
v. Isen (*Isana*) in Bavaria, with the market town Isen
vi. Eisack/Isarco (*Isarcus*) in North Tyrol
vii. Ismaning (*Isamanninga*) on the Isar River
viii. Isny, town in Württemberg, close to Bavaria, with the river *Isine*
ix. Numerous Isebeken, Eisbäche, Eisenbäche all over Germany
x. Isigny (*Isiniacum*) in France
xi. Isella, Isasca, Iseste, Isapis, Isaurus in Italy
xii. Riu Isalle (rivulet in Sardinia)
xiii. Isalli (former village in Sardinia)
xiv. Isallai (landscape in Sardinia)
xv. Izura (village in the Basque country)

Another study, Röder (2000), discusses about 250 such names, noting that they are most densely crowded around the Pyrenees. Vennemann further discusses the root word **is-*, found in the Germanic words for 'ice' (e.g. English *ice*, German *eis*, Icelandic *īs*). Its alleged Indo-European etymology is disputable because of the semantic problem with naming a flowing body of water after its non-flowing state. Vennemann claims it is a prehistoric Vasconic loan-word, still found in some Basque compounds (e.g. Basque *izoz-* 'frost/ice': *izoztela* 'ice', *izotzil* 'January' *izoztruma* 'ice-cone') (Vennemann 2003: 861).

Vennemann also suggests that numerous toponyms, traditionally considered Indo-European, are rather Vasconic words that have been morphophonologically adapted to Indo-European languages. An example of a regional toponym is the Vasconic word **aran*, meaning 'valley' (*haran* in Basque), which is often used to express the same meaning and can be found in many languages and areas such as *Val d'Aran* (Pyrenées), *Arundel* (England), *Arendal* (Norway), *Ahrntal* (Tyrol), *Arntal* (Germany) (Vennemann 2003: 865).

There are other aspects that Venneman claims have their roots in Vasconic, including vigesimal counting (base 20), and certain genetic characteristics found throughout Central and Western Europe. Indeed, Cavalli-Sforza used autosomal DNA

to discover the key components that account for the genetic variation found in Europe. Of these principle components, the fifth (explaining 5% of the genetic variation for 95 classical polymorphisms) corresponds to "the progressive retreat of the boundary of the Basque language" Cavalli-Sforza 1997: 7722), evidenced through the fact that the Basques have retained some of their original genetic characteristics as well as their language (Cavalli-Sforza et al. 1994; Cavalli-Sforza 1997).

However, Venneman's view is not uncontroversial. Scholars have taken issue with the proposal that the Basques formerly populated so much more of Europe or with aspects of the reconstructed Old European vis-à-vis proto-Basque (e.g. Trask 1995, Lakarra 2013), or have criticized Venneman for being overly reliant on the morphology (among other things, e.g. Baldi and Page 2006). Nevertheless, such a proposal allows standard assumptions to be questioned, including the directionality of the borrowing of the auxiliary split. Many of the toponyms that have been claimed to have come from Vasconic are found in Germany and France and other areas of Western Europe, but not in Spain or Portugal, just like the split auxiliary system, despite their proximity and genetic relationship to other languages that have this split. One possible explanation is that the HAVE/BE auxiliary split entered into a handful of geographically adjacent languages, just as the toponyms and hydronyms are most prevalent in these same areas, as it was once an area inhabited by Vasconic people. I make no claim as to whether the auxiliary system was borrowed into just one language (e.g. French) that then spread north and west, or whether the Vasconic people wielded wider influence. The proposal made here is only that Vasconic is a likely source of this feature.

4. Ergativity and HAVE/BE

The hypothesis that the HAVE/BE split originally stems from Vasconic is also supported by the fact that Basque is morphologically ergative, unlike Romance and Germanic languages. One assumes that a language becomes morphologically ergative with the introduction of a case marker for the transitive subject used consistently based on a verb's valency to distinguish the two arguments in a transitive clause. However, many have shown that case marking is not a purely syntactic phenomenon (e.g. Wierzbicka 1981; Donohue and Barðdal 2011 and papers therein). Another

hypothesis for the introduction of case markers in a language is that they come about due to semantic factors. In an ergative language, this could be due to the marking of e.g. volitional or causer subjects (or qualified in different terms that nonetheless refer to properties of the 'subject' arguments in the given predicates, e.g. 'agentivity'). Given the variation in agentivity in different predicates, such case markers would have thus naturally extended to include a subset of intransitive subjects sharing similar semantic properties (occurring with 'unergative' verbs). In fact, the Unaccusative Hypothesis (e.g. Perlmutter 1978; Burzio 1986; Levin and Rappaport Hovav 1994), that subjects of intransitive verbs may be patient-like or agent-like and that these differences have syntactic consequences, is very easily observed in terms of case marking in ergative languages and readily motivates the formation of the observed natural classes for both case and auxiliary splits. Given the correlation in Basque between the presence of an ergative case marker and the use of the HAVE auxiliary, a split auxiliary system is naturally born. Such as system is not so readily explained in the other Western European languages where this syntactic quirk is found, but similar semantic tendencies appear to underlie its usage. This could be due to the underlying principles being fundamentally present in all languages, but it could also be due to the borrowing of a split auxiliary system from adjacent languages from which other significant borrowing has already been established.

5. Summary

The proposal presented here, that Vasconic is the source of the HAVE/BE dual auxiliary usage found in parts of Western Europe, is supported by work on the origin of toponyms and hydronyms in the same region as well as by certain grammatical features of Basque. If we assume, following Vennemann (e.g. 2003), that Vasconic languages were once widespread in Western Europe, as suggested by these geographic terms and other evidence, we can hypothesize that the usage of both HAVE and BE in perfect auxiliaries in a small, geographically contiguous set of Western European languages was in fact derived from Old Basque, rather than the reverse. However, further work is needed to support this hypothesis. Establishing the chronology of the auxiliary split in Basque and its neighboring languages would be particularly helpful in supporting or refuting the claim that Basque is the source of the dual auxiliary usage.

References

Aldai, Gontza. 2002. *The grammaticalization of present and past in Basque*. Ph.D. Dissertation, University of Southern California.

Baldi, Philip, and B. Richard Page. 2006. Review: Europa Vasconica-Europa Semitica: Theo Vennemann gen. Nierfeld. *Trends in linguistics, studies and monographs* 138: 2183-2220.

Bentley, Delia, and Thórhallur Eythórsson. 2003. Auxiliary selection and the semantics of unaccusativity. *Lingua* 114: 447-471.

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach*. Dordrecht: Reidel.

Cavalli-Sforza, Luigi L. 1997. Genes, peoples and languages. *Proceedings of the National Academy of Sciences* 94(15): 7719-7724.

Cavalli-Sforza, Luigi L., Paolo Menozzi, and Alberto Piazza. 1994. *The history and geography of human genes.* Princeton: Princeton University Press.

Östen Dahl, Viveka Velupillai. 2013. The perfect. In Matthew Dryer and Martin Haspelmath (eds.), *The world Atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/68, Accessed on 2015-08-30).

Donohue, Cathryn. 2004. *Morphology matters: Case licensing in Basque.* Ph.D. Dissertation, Stanford University.

Donohue, Cathryn, and Jóhanna Barðdal. 2011. Empirical approaches to morphological case. *Morphology* 21(3): 481-485.

Krahe, Hans. 1964. *Unsere ältesten Flussnamen*. Harrassowitz: Wiesbaden.

Lakarra, Joseba. 2013. On Ancient European and the reconstruction of Proto-Basque. In Jürgen Udolph (ed.), *Europa Vasconica – Europa Semitica? Kristische Beiträge zur Frage nach dem baskischen und semitischen Substrat in Europa*. Beiträge zur Lexikographie und Namenforschung, 65-150. Hamburg: Baar-Verlag.

Languages of Europe. *Wikipedia, The Free Encyclopedia*. Wikimedia Foundation, Inc. 2004. 19 January 2012. http://www.enotes.com/topic/Languages_of_Europe

Legendre, Geraldine. 2007. On the typology of auxiliary selection. *Lingua* 117(9): 1522-1540.

Levin, Beth. 1983. *On the nature of ergativity*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Levin, Beth, and Malka Rappaport-Hovav. 1994. *Unaccusativity: At the syntax-semantics interface*. Cambridge, MA: MIT Press.

McFadden, Thomas. 2007. Auxiliary selection. *Language and linguistics compass* 1(6): 674-708.

Perlmutter, David M. 1978. Impersonal passives and the unaccusative hypothesis. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society (BLS)* 4: 157-189.

Röder, Katrin. 2000. *Struktur und verbreitung der alteuropäischen toponymie: Eine studie am beispiel der wurzelformens* $^{+}$is *und* $^{+}$ur. Ph.D. Dissertation, University of Munich.

Sorace, Antonella. 2000. Gradients in auxiliary selection with intransitive verbs. *Language* 76(4): 859-890.

Trask, Robert Lawrence Larry. 1995. Origin and relatives of the Basque language: Review of the evidence. In José-Ignacio Hualde, Joseba Lakarra, and Robert Lawrence Larry Trask (eds.), *Towards a history of the Basque Language*, 65-77. Current issues in linguistic theory 131. Amsterdam: John Benjamins.

Trask, Robert Lawrence Larry. 1997. *The history of Basque*. London; New York: Routledge.

van Valin, Robert D., Jr. 1990. Semantic parameters of split intransitivity. *Language* 66(2): 221-260.

Vennemann gen. Nierfeld, Theo. 2003. *Europa Vasconica–Europa Semitica*. In Patrizia Noel Aziz Hanna (ed.), *Trends in linguistics, Studies and monographs* 138. Berlin: Mouton de Gruyter.

Wierzbicka, Anna. 1981. Case marking and human nature. *Australian Journal of Linguistics* 1: 43-80.

Cathryn Donohue
Department of Linguistics
The University of Hong Kong
Pokfulam Rd, Hong Kong SAR
E-mail: donohue@hku.hk

Received: 2015. 08. 09.
Revised: 2015. 08. 25.
Accepted: 2015. 08. 25.