



Dorota Krajewska*

The diachrony of the Basque marker *bait-*: from a manner expression to subordinator

<https://doi.org/10.1515/ling-2021-0198>

Received October 30, 2021; accepted December 28, 2022; published online March 20, 2024

Abstract: This article analyzes the diachrony of the Basque marker *bait-*, which is a verbal prefix in subordinate clauses, but also has other functions: for example, it appears in independent clauses and indefinite pronouns. In subordinate clauses, it is used in two ways. First, it co-occurs with clause-initial conjunctions in reason, manner or result clauses or with pronouns in relative clauses. Secondly, it is used on its own, in relative, reason, result and complement clauses (with a limited group of verbs, such as emotive factive predicates or predicates of happening). The article combines evidence from a corpus study (6822 examples from 16th- to 20th-century texts) and internal reconstruction to (1) determine if and in what way the subordinator *bait-* and the affirmative *bai* ‘yes’ can be diachronically related, and (2) try to establish diachronic relations between the functions of *bait-*. It is proposed that the missing link between the subordinator and the affirmative particle might be a manner expression *bai* which had anaphoric functions. The marker *bait-* emerged as a reanalyzed form of the manner expression, which then gradually and through various pathways spread to different types of subordinate clauses and was reanalyzed as a subordinator.

Keywords: adverbial clauses; Basque; complement clauses; diachrony; relative clauses

1 Introduction

According to Lafon (1999 [1966]: 667), the particle *bait-* is one of the most remarkable aspects of the Basque grammar. It is used primarily as a subordinator but is also found in independent clauses as well as in a series of indefinite pronouns (e.g., *zerbait* ‘something’). It is also one of the very few grammatical prefixes in Basque.

One of the issues related to the marker’s diachrony is whether and how it is related to the affirmative adverb *bai* ‘yes’. While this idea is commonly accepted by

*Corresponding author: Dorota Krajewska, Linguistics and Basque Studies, University of the Basque Country (UPV/EHU), Vitoria-Gasteiz, Spain, E-mail: dorota.krajewska@ehu.eus. <https://orcid.org/0000-0002-6909-985X>

scholars (e.g., Lafon 1999 [1966]; Oyharçabal 1987), so far no satisfactory explanation has been proposed for the nature of this connection. In this article, I take up this issue to answer the following questions:

- What could be the connection between *bai* ‘yes’ and the subordinator *bait-*?
- Is it possible to establish diachronic relations between the different functions of *bait-*? In particular, what is the place of complement clauses with *bait-*?

I will argue that the missing link between the subordinator and the affirmative adverb might be a manner expression that had anaphoric functions. The subordinator could have grammaticalized from this manner expression and extended to various subordinate structures in a stepwise way. Even though many of the changes occurred before the first written records of the language, this article attempts to gather textual evidence of those processes and reconstruct the main stages of the grammaticalization processes involved. This is possible due to the phenomenon of persistence (Hopper 1991), i.e., the tendency of the traces of the original lexical meaning to remain in the grammaticalized item and to influence its grammatical distribution.

As regards the methodology, the study combines corpus analysis, internal reconstruction and insights from cross-linguistically common processes in the diachrony of subordinate structures. The main corpus, used especially to analyze the functions of *bait-*, consists of 37 16th- to 20th-century texts (listed in the Appendix), which cover the eastern and northern regions of the Basque Country: Navarre, Labourd, Lower Navarre and Soule (i.e., the areas where the prefix is extensively used). 6822 tokens of *bait-* were extracted from the corpus. R statistical tools (R Core Team 2022) were used for the quantitative part of the study.

The article is structured as follows. Section 2 presents the functions of *bait-*. Section 3 describes changes observed in the history of the language. Section 4 argues that the affirmative adverb *bai* can be seen as a manner expression and presents several hypotheses on how the subordinator developed from it and attempts to reconstruct how the marker extended to different types of clauses.

2 Functions of *bait-*

2.1 Overview

The morpheme *bait-* is usually described as a subordinator (see e.g., Artiagoitia et al. 2003: 711; Euskaltzaindia 1999: 12; Oyharçabal 1987). The main arguments in favor of the subordinate status of clauses with *bait-* are the following (Oyharçabal 1987: 263): (a) it is incompatible with other markers of subordination (relative *-en* or completive

-la) or with the conditional prefix *ba-*, (b) the verbal forms with *bait-* cannot have allocutive forms (which are only possible in independent clauses). However, the degree of integration of the subordinate clause varies, as will be shown. Moreover, *bait-* is occasionally found in independent clauses.

Nowadays, it is used in eastern and northern dialects of Basque. However, old western and central texts contain a few examples of *bait-*, which suggests that the morpheme was once common to all varieties (Azkue 1923: Section 529; Lafon 1999 [1966]: 681; Lakarra 1986; Lakarra 1996: 191; Mitxelena 2011 [1981]: 530–531).

The marker behaves like a proclitic (Oyharçabal 1987: 257). No other morpheme can go between *bait-* and the verb. If a verbal particle is used, it has to be placed before it (Oyharçabal 1987: 259). The negative particle *ez* also precedes *bait-* and, since it ends in a voiceless sibilant, it causes devoicing of the initial consonant of the marker, which is sometimes reflected in writing, e.g., *ezpaikara* = *ez* ‘no’ + *bait-* + *gara* ‘we are’. It can attach to any independent finite verb: it cannot appear on imperatives, allocutive forms, verbs carrying the interrogative suffix or hypothetical verb forms (Oyharçabal 1987: 257). Several phonological changes occur between the final occlusive of the marker and the first consonant of the verb: *-t + d-/g- = -t-/k-* (*baitira*, *baikara*), *-t + z- = -tz-* (*baitzen*), *-t + n-/l-/h- = -n-/l-/h-* (*bainaiz*, *bailira*, *baihaiz*). Additionally, Souletin has the form *beit-*, sometimes reduced to *be-* before *n*, *h* and *l*.

In subordinate clauses the marker is used as relativizer, complementizer and adverbializer. Its major uses can be divided into two classes (Figure 1). The classification I present here is roughly based on Oyharçabal (1987: 248).

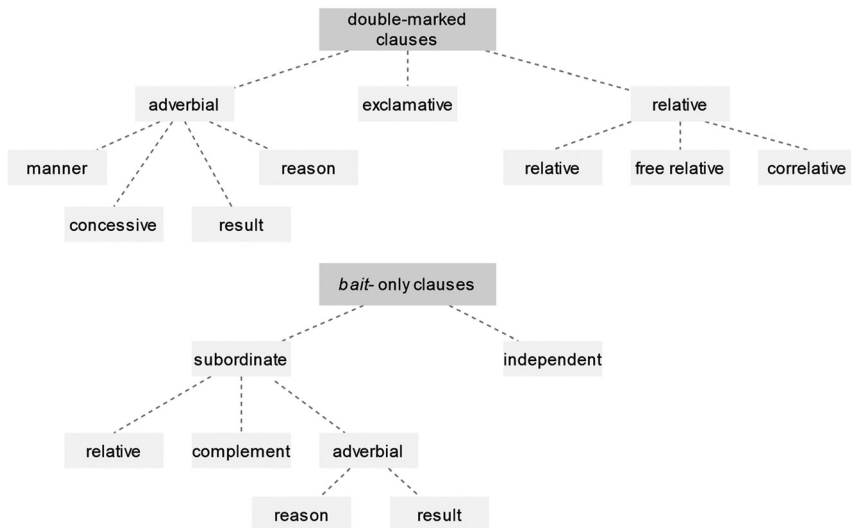


Figure 1: Functions of *bait-*.

First, *bait-* is employed together with several clause-initial elements:

- conjunctions in adverbial clauses (reason (1), manner (2), concessive and result clauses are the main types)
- interrogative pronouns in headed relatives (3), correlatives (4) and free relatives (5)

- (1) *Erra-ten da mortale-a zeren ill-zen baitu arima*
 say-IPFV AUX.3SG mortal-DEF because kill-IPFV bait.AUX.3SG>3SG soul
 ‘It is called mortal because it kills the soul.’
 (BeriainDotrina, 1626)¹
- (2) *Eta kita ietzaguk gure zorr-ak, nola gu-k ere*
 and forgive AUX.IMP.2SG>3PL<1PL our debt-DEF.PL how we-ERG also
gure zordun-ei kita-tzen baitrauegu.
 our debtor-DAT.PL forgive-IPFV bait.AUX.1PL>2SG<3PL
 ‘And forgive us our debts, as we forgive our debtors.’
 (LeizarragaTest, Matt 6: 12, 1571)²
- (3) *aitafamilia bat, zein-ek landa baitzezan mahasti bat*
 landowner one which-ERG plant.RAD bait.AUX.AOR.3SG>3SG vineyard one
 ‘a landowner who planted a vineyard’
 (LeizarragaTest, Matt 21: 33, 1571)³
- (4) *Nor-i ere pot egin-en baitraukat, hura da*
 who-DAT PTCL kiss make-FUT bait.AUX.1SG>3SG<3SG that be.3SG
 ‘Whoever I will kiss, he is the one.’
 (LeizarragaTest, Matt 26: 48, 1571)⁴
- (5) *Eta eman-en daraie-la [zer bai-tute merezi]*
 and give-FUT AUX.3SG>3SG<3PL-COMP what bait-AUX.3PL>3SG deserve
 ‘And that he will give them what they deserve.’
 (Etxepare, 1545)

¹ Unless stated otherwise, the texts come from the *Euskal Klasikoen Corpusa* corpus (Euskara Institutua 2013). Note that the Basque examples, depending on the source, are mostly given in modernized orthography.

² Since Leizarraga translated the Bible from French, it might be interesting to compare the Basque examples to the original (Le Nouveau Testament 1563). For (2), the French text reads: “*Et nous remet nos dettes, comme aussi nous les remettons à nos detteurs*”.

³ The French source: “*It y avoit un pere de famille, qui planta une vigne*”.

⁴ The French source: “*Celuy que ie baiseray, c’est luy*”.

These kind of *bait-* clauses can be called ‘doubly-marked’ subordinate clauses, as there are two subordinating elements: *bait-* on the verb and a clause-initial conjunction or pronoun.

As regards the syntactic status of doubly-marked *bait-* clauses,⁵ except for free relatives, they are adjoined to the main clause, i.e., they are not embedded and cannot function as a constituent in the main clause. They occupy a marginal position with regard to the main clause (they can appear before or after it, depending on the clause type). Relative clauses can be extraposed (they do not have to appear immediately to the right of the head noun) and they can also appear as parentheticals in the middle of the main clause. In correlatives the degree of integration with the main clause might be considered higher (though they are not embedded), as the clause with *bait-* (which always precedes the main clause) is not easily omittable, and in the second clause there is usually a pronoun referring back to the clause with *bait-*. Finally, free relatives are embedded and function as a constituent in the main clause.

The prefix *bait-* is also used together with the conjunction *ala* in exclamatives (see Rebuschi [2008] for an analysis of their syntax):

- (6) *Ala ni dohacabe handi-a bainaiz!*
 ala I unlucky big-DEF bait.be.1SG
 ‘I am so unlucky!’
 (Materria, 1623)

In the second class (*bait-* only clauses), there is no conjunction or any other element that determines the meaning. Thus, the interpretation depends on the context and often more than one reading is possible. Based on the function that the *bait-* clause fulfills, we can distinguish relative clauses (7)–(8),⁶ reason clauses (9)–(10), result clauses (which sometimes can be interpreted as expressing manner (11)) and complement clauses (discussed in Section 2.2). When the clause with *bait-* is placed before the main clause, the marker expresses reason, background information or topic (12).

- (7) *Anaya, igori-co darauat goutum-bat, [jaun-a-c eman*
 brother send-FUT AUX.1SG>3SG<2SG letter-one lord-DEF-ERG give
baytaraut galcerdy ceta-z-ko bat-en barnean].
 bait.AUX.3SG>3SG<1SG sock silk-INS-RM one-GEN inside
 ‘Brother, I will send you a letter which the lord gave to me inside a silk sock.’
 (15th cent., Mitxelena 2011 [1964]: 3.2.8)

⁵ I will follow Lehmann’s (1988) distinctions when describing the syntax of subordination, especially with respect to the level of integration of the main and subordinate clauses.

⁶ A prenominal construction (typical of head-final languages) is the least marked relative construction in Basque since the oldest texts, even though several postnominal constructions have also emerged at different stages of the language.

- (8) *Ethorri da ordu-a [gizon-a-ren Seme-a glorifikatu-ren baita].*
 come AUX.3SG hour-DEF man-DEF-GEN son-DEF glorify-FUT
baita].
 bait.AUX.3SG
 ‘The hour has come that the son of the man will be covered in glory.’
 (LeizarragaTest, John 12: 23, 1571)⁷
- (9) *Erran-en dautzuet, gero ere jakin-en baituzue.*
 say-FUT AUX.1SG>3SG<2PL later also know-FUT bait.AUX.2PL>3SG
 ‘I will tell you, because you will know later anyway.’
 (Larzabal, 20th c.)
- (10) *Ordea nolatan ez-tuzue egi-ten? Ezin baitaidikezue.*
 so why NEG-AUX.2PL>3SG do-IPFV cannot bait.AUX.POT.2PL>3SG
 ‘So why you don’t do that? Because you can’t.’
 (Axular, 1643)
- (11) *Gau-a jin da betbetan, bi urrats-etan ez baitzen fits-ik ageri.*
 night-DEF come AUX.3SG suddenly two step-INES.PL NEG bait.be.PST.3SG
 nothing-PART visible
 ‘The night came suddenly, so that one couldn’t see anything two steps ahead.’
 (Etxepare Buruxkak, 1910)
- (12) *Eta erran baituçu Iesu Christo-ren fede-a*
 and say bait.AUX.2SG>3SG Jesus Christ-GEN faith-DEF
du-en-a de-la guiristino, cer-tan dago
 have.3SG>3SG-SUB-DEF be-COMP Christian what-INDEF.INES be.3SG
principalqui Iesu Christo-ren fede-a?
 mainly Jesus Christ-GEN faith-DEF
 ‘So you have said that the one who has the faith of Jesus Christ is Christian,
 what does the faith of Jesus Christ mainly consist of?’
 (Materra, 1623)

Leaving aside complement clauses, the clauses with *bait-* illustrated above are adjoined and not embedded, similar to their doubly-marked equivalents. This also applies to examples interpreted as relative: the head noun and the relative clause do not form a single constituent and the *bait-* clause does not necessarily appear immediately after the noun it modifies. The only syntactic difference between relative and other *bait-* only clauses is that in the former there must be a participant

⁷ The French source: “L’heure est venue, que le Fils de l’homme doit estre glorifié”.

shared between the main and the subordinate clause, while in the latter there is no such restriction.

Moreover, *bait-* can have a discourse function, expressing shades of consequence, result, contrast or emphasis (13)–(14). In the oldest texts, the verb marked with *bait-* can appear in the last clause of a sequence, and the clause with *bait-* can be linked with the preceding one using the conjunction *eta* ‘and’, as in (13).

- (13) *baina erra-k solament hitz-a, eta sendatu-ren baita*
 but say-IMP.2SG just word-DEF and heal-FUT bait.AUX.3SG
ene muthill-a.
 my boy-DEF
 ‘but say just a word and my boy will be healed’
 (LeizarragaTest, Matt 8: 8, 1571)⁸

- (14) *Zoin laster eta aise juja-tzen duzu-n, zu-k,*
 how fast and easily judge-IPFV AUX.2SG>3SG-SUB you-ERG
jende-a! Baititut ene arrazoin-ak!
 people-DEF bait.have.1SG>3PL my reason-DEF.PL
 ‘– How fast and easily you judge people! – I (bait-) have my reasons!’
 (Larzabal, 20th c.)

Finally, the marker shows up in a few other areas of grammar: in place names, possibly in the animate locative postposition *baita-*, in the archaic prefix *albait-*, in indefinite pronouns and as a temporal conjunction.

Toponyms with *bait-* are found in some northern regions of the Basque Country, for example, *Espela bayta* ‘(the place) where the box tree is’ (1591) (Salaberri Zaratiegi 1996: 226). According to Salaberri Zaratiegi (1996: 226), they are Satznamen, which developed from *bait-* relative clauses. We can reconstruct (15) as the source construction with the subsequent ellipsis of the head noun.

- (15) *(leku-a) [espel-a baita]*
 place-DEF box-DEF bait.be.3SG
 ‘(the place) where the box tree is’

The locative postposition *baita-* used with animate referents should also be mentioned here.

- (16) *ni-re baita-n*
 I-GEN ANIM-INES
 ‘in me’

⁸ The French source: “*mais seulement di le mot, et mon garson sera guari*”.

- (17) *anaia baita-ra*
 brother ANIM-ADL
 ‘to the brother’s place’

Mitxelena (2011 [1970]: 264) and Creissels and Mounole (2011: 179) mention the possibility that *baita-* could have developed from a subordinate clause (most probably a relative clause) with the verbal prefix *bait-*. According to Creissels and Mounole (2011) a change from ‘at the place where N is’ to ‘at N’s place’ could have taken place. From there it could develop into the locative postposition (see also Krajewska 2022), cf. ‘house/home’ > locative grammaticalization path (Kuteva et al. 2019: 233, 235–236).⁹

Albait-, called *prescriptif* by Lafon (1980 [1944]), is an archaic verbal prefix no longer employed in Basque, found only in 16th- and 17th-century northeastern texts (Oyharçabal 1997: 61). As for *al-*, Lafon (1999 [1966]: 692) proposed that it is the particle of possibility *ahal*, as in *ahal izan* ‘can’. *Albait-* appears in second and third-person hypothetical verb forms and expresses an imperative that is to be executed in the future or if a condition is met (18) (Lafon 1980 [1944]: 491).

- (18) *Eta orduan baldin nehor-c ba-darraçue, Huna hemen*
 and so if someone-ERG COND-say.3SG>3SG<2PL Here here
Christ edo Hara han: ez-albeitzineçate sinhets.
 Christ or there there NEG-albait.AUX.HYP.2PL>3SG believe
 ‘And if someone tells you ‘Christ is here!’ or ‘There!’, do not believe.’
 (LeizarragaTest, Matt 24: 23, 1571)¹⁰

Furthermore, we find *-bait* as a suffix on a series of indefinite pronouns: for example, *norbait* ‘somebody’, *zerbait* ‘something’ or *nonbait* ‘somewhere’. Azkue (1923: 357) was first to realize that the subordinator and the suffix that attaches to interrogative pronouns are one and the same and that the particular variant used in subordinate clauses is also found in pronouns. Thus, in Soule the verbal marker became *beit-* (the first examples appear in the 17th century, but it became frequent in the 18th [Padilla-Moyano 2017: 703]), and the same form appears in pronouns too. In Biscay, the variant *baist-* is found as a prefix in a handful of examples in the anonymous collection of proverbs known as *Refranes y sentencias* (1596) and also occasionally in pronouns.

⁹ Creissels and Mounole (2011) conclude that this explanation is problematic, because, if it were true, we would not expect genitive marking on the noun (as in (16)). However, this might not be a problem since, as explained in Krajewska (2022: 249), in the earliest texts the absolutive (17) actually predominates and it is possible that the genitive was added later in analogy to other nominal forms which are built on the genitive.

¹⁰ The French source: “Lors si quelqu’un vous dit, Voici le Christ, ici ou là, ne le croyez pas”.

Finally, in some areas, especially in Soule, *bait* can be used as a temporal conjunction in non-finite clauses with the meaning ‘by the time’ (19). The affirmative adverb *bai* or forms built upon it – *baiko* or *baikoz* – are employed in some other dialects (Lafon 1999 [1973]: 44).

- (19) *zū jin bait, egin-ik dūket*
 you come bait do-PART have.FUT.1SG>3SG
 ‘I will have done this before by the time you have arrived.’
 (Gèze 2010)

2.2 Complement clauses with *bait-*

The most common complementizer in Basque is *-ela* (Artiagoitia 2003; Artiagoitia and Elordieta 2016), which is used in indicative complements as well as in some subjunctive complements (*-ela* contrasts with *-en*, employed in indirect questions and some subjunctive complements). The range of use of *bait-* is more limited than that of *-ela* as only few verbs take complements with *bait-* apart from the more usual *-ela* complements.

The first class of verbs are factive predicates. As put by Kiparsky and Kiparsky (1970: 147), with factive predicates “[t]he speaker presupposes that the embedded clause expresses a true proposition, and makes some assertion about that proposition”. For instance, whichever sentence in (2) the speaker utters, he or she presupposes what is said in (21).

- (20) a. *Bill regrets that Sheila is no longer young.*
 b. *Bill doesn’t regret that Sheila is no longer young.*
 c. *Does Bill regret that Sheila is no longer young?*
- (21) *Sheila is no longer young.*
 (Karttunen 1971b: 55)

Predicates which can take *bait-* complements are mostly affective factive verbal and non-verbal predicates, e.g., *pena/domaia da* ‘it’s pity’ (22), *harritzeko da* ‘it’s surprising’, *ez da dudarik* ‘there is no doubt’, *kontent izan* ‘be happy’, *xantza izan* ‘be lucky’ (23). Epistemic factive verbs like ‘know’, ‘see’ or ‘realize’ never take *bait-* complements. Neither verbs of saying nor verbs of desire are attested with *bait-*. The marker also often surfaces with adverbs such as *beharrik* ‘luckily’, though examples without the subordinator on the verb are also possible (OEH,¹¹ s.v. *beharrik*).

11 OEH: *Orotariko euskal hiztegia – General Basque dictionary* (Euskaltzaindia 2019).

- (22) *Pena da egiazko hobendun-ek ez baitute hain laster*
 pity be.3SG true culprit-ERG.PL NEG bait.AUX.3PL>3SG so fast
aitor-tzen.
 confess-IPFV
 ‘It’s pity that real culprits do not confess so fast.’
 (Larzabal, 20th c.)
- (23) *Xantza duzue gizon-a itsaso-an baitut.*
 luck have.2PL>3SG man-DEF sea-INES bait.have.1SG>3SG
 ‘You’re lucky that my man is at sea.’
 (Larzabal, 20th c.)
- (24) *Plazer dut erran baituzu ongi.*
 pleasure have.1SG>3G say bait.AUX.2SG>3SG well
 ‘I’m pleased because/that you have said it well.’
 (BerianDotrina, 1626)

The predicates mentioned above are not uncommon in modern texts (22)–(23), but they are scarce in older sources. Before the 18th century, only a few examples are found, and they are usually ambiguous between reason and complement as in (24). This difference might be partially due to the type of sources available for earlier stages of the language (i.e., mostly religious texts), which probably did not favor the use of such predicates. Nevertheless, it seems that factive predicates tend to appear with non-finite complements in older texts (something still possible nowadays), as in (25).

- (25) *Dolu dizit eta damu zure kontra egin-a-z.*
 pain have.1SG>2SG.ALLOC>3SG and remorse you.GEN against do-DEF-INS
 ‘I feel pain and remorse that I have acted against you.’
 (Etxepare, 1545)

Another class, quite different from factive predicates, consists of predicates such as *gutitarik egin du* or *doi doiak egin du* that expresses that something was very close to happening but eventually did not (26).

- (26) *Gutitarik egin zuen, ez baitzioten*
 nearly do AUX.PST.3SG>3SG NEG bait.AUX.PST.3PL>3SG<3SG
bertze-a-ri zango-a moztu izan behar.
 other-DEF-DAT leg-DEF cut AUX must
 ‘They almost had to cut his leg.’
 (Hiriart-Urruti, 1891–1914)

Furthermore, in some texts, the *bait-* clause appears to be the complement of a noun, e.g., of *kausa* ‘reason’ (27).

- (27) *baina bekatu-a izan da kausa, gizon-a bothere hunez*
 but sin-DEF be AUX.3SG reason man-DEF power DEM.INS
gabetu izan baita
 deprive AUX bait.AUX.3SG
 ‘But the sin was the reason that the man was deprived of this power’
 (LeizarragaABC, 1571)

This appears to be a special case of a more general structure briefly described by Lafitte (1991 [1944]: Section 770), who suggests that any verb can take complement with *bait-* and the marker adds the sense of ‘the fact that’. In such cases, the subordinate clause is sometimes anticipated in the main clause by a demonstrative. In (28), for example, we have the demonstrative *haur* ‘this’. In (29), on the other hand, the main clause in the answer is elided. Such complements are also factive: as argued by Kiparsky and Kiparsky (1970), factive predicates can be paraphrased with *the fact that* (e.g., *It’s surprising that ...* → *The fact that ... is surprising*).

- (28) *Zer erhokeria da haur, uste baitu gizon hun-ek,*
 what madness be.3SG this think bait.AUX.3SG>3SG man this-ERG
berretu-z karga arindu-ko zeikala?
 increase-INS load decrease-FUT AUX.COMP.3SG<3SG
 ‘What is this madness, (the fact) that this man thinks that the load will become lighter by increasing it?’
 (Axular, 1643)
- (29) *Hon-en sortzi-a-k zer dü bereberik? Grazia-n*
 DEM-GEN birth-DEF-ERG what AUX.3SG>3SG special grace.DEF-INES
sorthü baita.
 be.born bait.AUX.PST.3SG
 ‘What is special about his birth? (The fact) that he was born in grace.’
 (Belapeire, 1696)

The final class are predicates of happening, such as *gertatu* or *heldu* ‘happen’. Contrary to emotive verbs, they appear already in the oldest texts (30). From the point of view of semantics, when such predicates are used the formally subordinate clause contains the foreground information (‘focal clause’ in terms of Dixon and Aikhenvald [2009]) and the main clause only emphasizes the statement. They can be classified as implicative verbs in the sense of Karttunen (1971a), that is, they imply the factivity of the subordinate clause. Factive predicates also have this property: the difference is that negated implicative verbs imply the negation of the complement (e.g., *It*

happened that the poor one died → *he died*, but *It did not happen that the poor one died* → *he did not die*).

- (30) *Gertha zedin bada hil baitzedin paubre-a*
 happen.RAD AUX.AOR.3SG so die bait.AUX.PST.3SG poor-DEF
 ‘So it happened that the poor one died’
 (LeizarragaTest, Luke 16: 22, 1571)¹²

Thus, complement clauses with *bait-* tend to express facts or attitude towards them (Euskaltzaindia 1999: 1.3.6). This concurs well with other uses of the marker. For example, reason and result clauses are factually oriented, i.e., they typically express something that has happened, as opposed to purpose clauses which express unrealized events (Hetterle 2015: 50–53; Thompson et al. 2007: 250–251). With regards to reason clauses, Pérez Saldanya (2020: 591) observes that *bait-* is used in clauses which are assertive or presupposed. Also, exclamatives can be seen as factive (see Grimshaw 1979), as they can be paraphrased with a factive predicate (“It’s amazing that ...”, “I’m surprised that”).

Crosslinguistically, complement clauses tend to be embedded in the main clause and thus occupy a grammatical slot in it. In Basque, it is the case with complements with *-ela*. The situation is more complicated with *bait-* complements. Clearly, *bait-* complements are not easily omissible, which suggests tighter integration with the main clause (as compared to, for example, reason clauses). The order of clauses is fixed: *-ela* complements can precede or follow the main clause, but complements with *bait-* can only occupy the second position (Oyharçabal 1987: 249). In many cases, the clause with *bait-* is at the margin of the main clause: in (28) the link between the clauses is established with the demonstrative placed in the main clause which refers to the content of the subordinate clause. However, in examples such as (22) the subordinate clause might be analyzed as occupying the subject slot. Modern data (gathered in Norantz corpus [Oyharçabal et al. 2009]) add interesting details to the issue: some speakers put both clauses in single intonational phrase, but others place a break between the two clauses (31), suggesting that *bait-* complements are not always embedded even nowadays.

- (31) *Domaia da/ euri i-ten baitu.*
 pity is.3SG rain do-IPFV bait.AUX.3SG>3SG
 ‘It’s a pity that it is raining.’
 (Norantz, C161-XLEAHA)

¹² The French source: “*Or advint que le povre mourut*”.

3 Diachronic changes in the use of *bait-*

Changes in the use of *bait-* were analyzed in the corpus described in Section 1. The corpus was divided into four periods: (i) 1500–1600, (ii) 1600–1750, (iii) 1750–1900 and (iv) 1900–1970.¹³

Since the earliest texts, i.e., the 16th century, the patterns of use of *bait-* have changed: in the oldest texts *bait-* clauses are typically doubly marked, and *bait-* only clauses are infrequent. In modern texts, the proportions are the opposite (Figure 2).¹⁴

At first sight, this could appear as a major change in subordination strategy. However, it rather represents the preference of early Basque writers – and even more so of translators – for the more explicit and more Romance-like constructions (regardless of whether they were indeed borrowed from Romance, as argued by, e.g.,

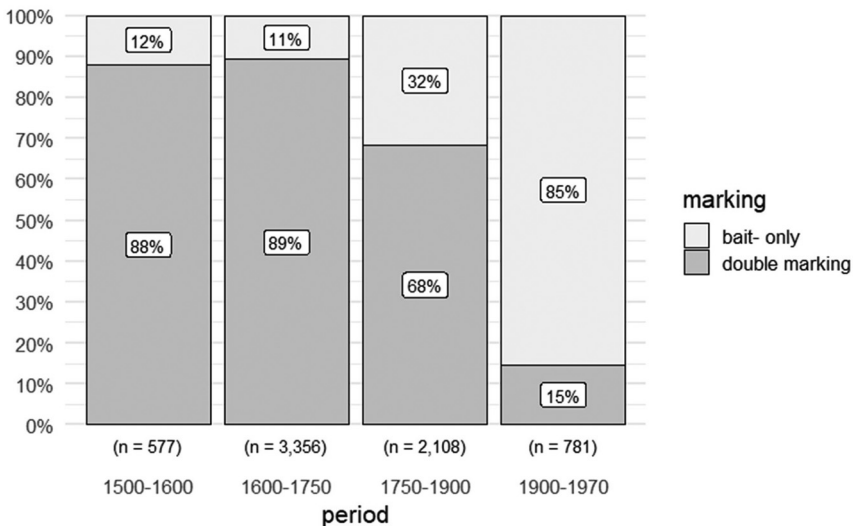


Figure 2: The proportions of the two kinds of clauses with *bait-* in the corpus.

¹³ This classification follows the most widely accepted periodization of the language (see, e.g., Lakarra et al. 2019), which is based on criteria related to language-internal factors and the most influential landmarks in the history of the Basque literature.

¹⁴ A logistic regression analysis was performed with the kind of marking as the dependent variable and the period as the independent variable (and text as random effect). The effect of period is significant ($\chi^2 = 32$, $df = 3$, $p < 0.001$). Pairwise comparisons between the periods confirm that the differences between the 2nd and 3rd and 3rd and 4th periods are significant. The details of all statistical analyses can be found in the supplementary materials.

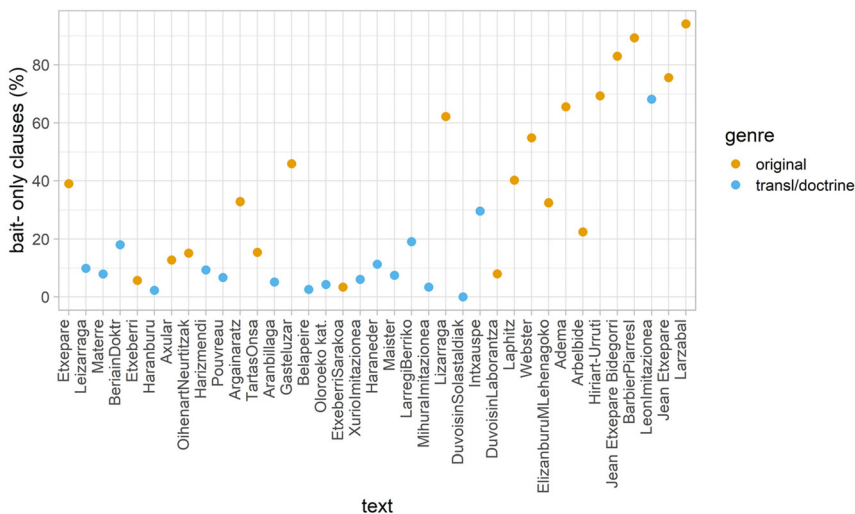


Figure 3: The proportion of *bait*-only clauses in the texts in the corpus (the texts are ordered chronologically, from the oldest on the left).

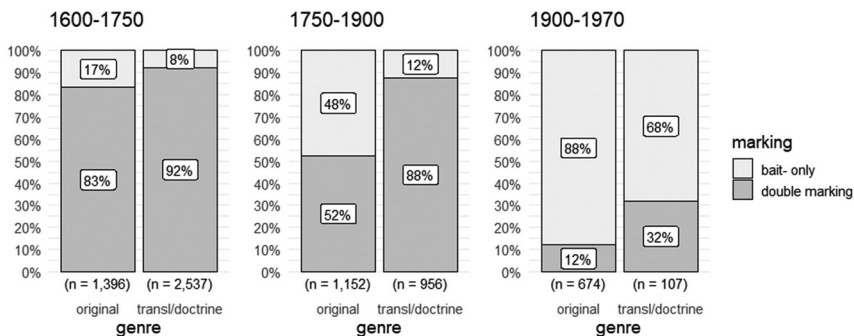


Figure 4: Clauses with *bait*- in original versus non-original texts.

Lafon 1999 [1973]). In fact, in the few early texts which are not translations, the proportion of *bait*-only clauses tends to be higher, as shown in Figure 3, which plots the proportion of *bait*-only clauses for all texts in the corpus. In general (Figure 4), the proportion of clauses with a conjunction is systematically higher in translations and doctrines than in original texts in all periods.¹⁵

¹⁵ A logistic regression analysis was performed with the kind of marking as the dependent variable and the period and text genre as the independent variables (and text as random effect). The effect of genre is significant ($\chi^2 = 16.7$, $df = 1$, $p < 0.001$).

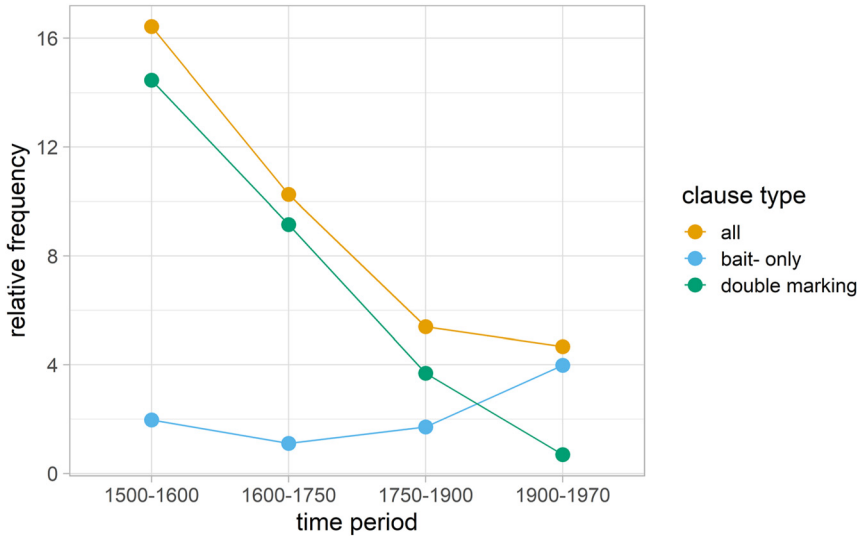


Figure 5: The relative frequency (per 1,000 words) of all types of *bait-* clauses together, and *bait-* only clauses and those with a conjunction separately.

Figure 5 plots relative frequencies (per 1,000 words) of all examples with *bait-* and separately for the two types of clauses, *bait-* only and *bait-* with an additional element. It can be seen that the relative frequency of the marker is much lower in modern texts. This is caused by a sharp decrease in the use of clauses where *bait-* appears with a conjunction or a relative pronoun. However, the relative frequency of *bait-* only clauses increased in the late 19th century.¹⁶

As regards *bait-* only clauses, nowadays the most common type is reason. However, in historical sources the situation is different. Table 1 presents the proportions of the following types of clauses: complement, relative, result, reason, clauses ambiguous between relative and reason and all other clauses (i.e., *bait-* in independent clauses). The number of examples for the 16th century is low, but complement clauses in this period are very frequent. However, this is caused by the high incidence of ‘happen’ verbs with *bait-* complements in the translation of the New Testament, which is the most extensive source for this period. In the second period, for which we have much more data, the most prominent functions of *bait-* are relative and reason, but other functions are also quite common (except for the result,

¹⁶ Three negative binomial regression models were fitted with the period as the independent variable and the count of (a) all *bait-* clause, (b) *bait-* only clauses and (c) clauses with conjunction or pronoun (controlling for the length of the text in each case). The results show that the effect of period is significant in all models.

Table 1: Proportions of the various types of *bait*- only clauses.

Period	Tokens	Comp.	Rel.	Result	Reason	Rel./reason	Other
1500–1600	69	23.2	24.6	1.4	13.0	2.9	34.8
1600–1750	361	9.1	29.6	2.2	34.6	12.7	11.6
1750–1900	668	7.9	18.1	4.2	52.2	7.9	9.6
1900–1970	666	4.8	5.7	6.6	64.4	3.6	14.9

Table 2: Frequencies of different types of predicates taking *bait*- complements (in percentages).

Period	Tokens	'happen'	Noun compl.	Emotive
1500–1600	16	81.2	18.8	0.0
1600–1750	33	9.1	75.8	15.2
1750–1900	53	32.1	15.1	52.8
1900–1970	32	6.2	9.4	84.4

which is, generally speaking, infrequent). An interesting detail is that clauses ambiguous between relative and reason are particularly frequent in this period. The reason reading starts to be dominant in the third period (with around half of the examples expressing reason). This tendency becomes even more pronounced in the most recent texts, where 64 % of all *bait*- only clauses are of that type.

Turning to complement clauses, they are not very frequent overall. Nevertheless, as regards the types of predicates, we can tentatively say that a change in preference occurred between the earliest and the newest sources. Table 2 provides the proportions of verbs of happening, complements of nouns (together with 'the fact that' constructions) and factive emotive verbs. The latter type is not attested with *bait*- in the earliest sources, but they predominate since the 19th century.

4 The hypothetical link between *bait*- and *bai* 'yes'

Lafon (1999 [1966]) mentions several times that there is a relationship between *bai* and *bait* and he suggests that the initial function of *bait*- was to emphasize affirmation (1999 [1966]: 667, 683). However, the exact nature of this connection is never clearly spelled out. The question is if and how an affirmative adverb can turn into a marker of subordination. Cross-linguistically, some classes of words tend to grammaticalize as markers of subordination (e.g., demonstratives or pronouns), but affirmative adverbs are not among them (Givón 2009: chap. 5; Heine and Kuteva 2002:

113–114, 335; Heine and Kuteva 2007: sec. 5.3.1; Hendery 2012: 2.2; Kuteva et al. 2019: 476–477), though Hendery (2012: 66–71) discusses several discourse markers which were source of a relative clause marker (including the Basque *bait-*). An interesting parallel is that of South Slavic languages which use the complementizer *da*, which is also the affirmative particle ‘yes’ in many Slavic languages (Wiemer 2019). The particle *da* was used first with optative or hortative function (Wiemer 2019: 120), and in Old Church Slavonic it was already employed as a clause-initial particle in a wide range of clauses, and “its function was associated with unrealized states of affairs” (Wiemer 2019: 121). Wiemer (2019) argues that in the first stage of the development of the complementizer there were two juxtaposed clauses, with *da* in the second. In this context, the function of the particle was reanalyzed as a subordinator, and the relation between the clauses became asymmetrical. The reconstruction of developments in Basque presented here differs in semantic details but implies a similar process of increasing clausal integration. More generally, as shown by Hopper and Traugott (2003: Ch. 7), many subordinate structures grammaticalize along the path from parataxis (independent clauses) to hypotaxis (interdependency) to subordination (understood as embedding).

The process of which the endpoint is the subordinator *bait-* as used in the modern language can be seen as an example of grammaticalization. First, the affirmative *bai* emerged from its lexical source and acquired various grammatical functions. Then the marker *bait-* developed from *bai* and underwent further changes in which it acquired even more grammatical function (‘secondary grammaticalization’). Since the oldest texts, *bait-* functions as a subordinator with a wide range of functions, but it has been proposed for various languages that this situation is the endpoint of a series of extensions from one clause type to another (see examples in Hendery [2012: Section 2.2.8]).

The goal of the remainder of this article is to try to reconstruct the details of the process. The discussion is going to be speculative in part, as most of the processes in question took place before the earliest texts. Nevertheless, a careful analysis of textual evidence is crucial. One of the most prominent features of grammaticalization is ‘persistence’ (or ‘source determination’ in Bybee et al. [1994: 15]), or the adherence of the original lexical meaning, which might determine the grammatical distribution of the grammaticalized element (Hopper 1991). As put by Bybee et al. (1994: 18), “we find that multiple uses and the retention of lexical specificities can be employed as diagnostics of the earlier history of grammatical material, even in languages for which historical attestation is sparse or nonexistent”.

The view presented here is not completely different from Lafon’s (1999 [1966]) explanations of the various uses of *bait-*. The main difference is that Lafon tried to link all of the functions of the prefix to the affirmative meaning of *bai*. What I propose here is that the subordinator cannot go back to an affirmative adverb, but rather to

the anaphoric function of *bai*. Thus, *bai* ‘yes’ and *bait-* are related to each other, not because the latter developed from the former, but because both share a common source. Based on textual evidence, I will show that *bai* can be considered a manner expression with anaphoric functions.¹⁷ There are two arguments in favor of this idea, which I discuss in Section 4.1: (1) the etymology of *bai* is compatible with this proposal, and (2) *bai* has or had functions typical of grammaticalization paths of manner deictics. Then, in Section 4.2, I will focus on the development of *bait-* and the further stages of its grammaticalization: introduction in subordinate clauses, extension and generalization to more types of clauses and appearance of more tightly integrated structures.

4.1 Etymology and functions of *bai*

According to Trask (1997: 209), “The word *bai* ‘yes’ is doubtless related to the affirmative morph *ba-*, as in *badator* ‘he/she’s coming’, formerly ‘he/she is coming’; *bai* may well be an ancient verb-form, along the lines of ‘it is so’, possibly involving the root *-di-* of the archaic verb **edin*”. Lakarra (2018: 158) elaborated on this idea suggesting that the underlying verb form is **badadi*, a form of the participle **edin* ‘become’ and copular verb (subjunctive of ‘be’ in the modern language). It would mean something like ‘it is (so)’ or ‘let it be’. The prefix *ba-* could have functions similar to those it performs in the modern language: reinforcing finite verbs in clause initial positions and emphasizing affirmation. The affirmative function of *bai* might have appeared very early, but it is not the only function the particle has had. It seems possible that it came to be used as a kind of a manner expression, with the meaning of ‘so, as, thus, this way’.

This explanation has parallels in other languages: items with discourse functions frequently derive from verb forms. For instance, Bourdin (2008) analyzes grammaticalization of ‘go’ and ‘come’ into textual connectives and, in particular, devices expressing consequence of the previous discourse, or to the contrary, expressing something unexpected. Also several emphatic affirmative expressions similar to the

¹⁷ More recently, the idea that *bait-* is grammaticalized from a manner deictic is presented in Pérez Saldanya (2020), who, as regards the development of *bait-*, accepts the ideas discussed in Krajewska (2017). Pérez Saldanya focuses on causal clauses with *bait-* and shows interesting parallels with grammaticalization processes that occurred in Romance and Latin, which provide additional support for the hypothesis linking *bait-* with the manner expression. Additionally, Pérez Saldanya (2020) points out to the fact that the grammaticalization of *si* proceeded further in Occitan than in Spanish. Thus, as he suggests, the differences between Basque dialects might be related to different contact situations: Eastern varieties, traditionally in contact with Occitan varieties retained and extended the use of *bait-*, while the Western varieties lost the marker.

reconstructed Basque phrase can be mentioned: Spanish (*que*) *así sea* ‘let it be so’, *así es* ‘it is so’ or French *ainsi soit-il* ‘let it be so’. Pérez Saldanya (2020) provides interesting parallels between *bai*, *bait-* and the prefix *ba-* and Latin and Occitan-Catalan forms. In Latin, the conditional *si* comes from the manner expression *sīc* (which in turn was reinforced as *acsīc*, which gave way to manner deictics in some modern Romance languages, e.g., Spanish *así* ‘this way’).

Moreover, other grammatical words in Basque can be explained through a grammaticalization of the verb **edin*, e.g., *baino* ‘than’ or *baina* ‘but’ (Lakarra 2018: 109). Another particle, in a way more closely related to *bai*, is *bada* ‘so, thus’. The source of this particle could well be *ba-da* with the emphatic-affirmative *ba* and 3sg form of *izan* ‘be’. It would, thus, initially mean something like ‘it is so (indeed)’. If the source of *bai* was indeed a form or the verb **edin*, such as **ba-dadi*, *bai* and *bada* would have very similar sources (though probably different chronologies, *bada* being more recent).

König (2015: 41) examines manner deictic elements (such as *so*, *this way*, *like*) and various grammaticalization paths they typically move along and argues that renewal of manner deictic elements is very common. In many languages, the exophoric function is often lost: in English, for example, *so*, *such* or *thus* are rarely used exophorically, and *this way* or *like this* are employed instead. Partial renewal also took place, for instance, in Italian (*ecco + si* became *così*).

Anaphoric and cataphoric use develop from exophoric use: instead of pointing to the external world, deictic elements start to indicate parts of the discourse. As explained by König (2015: 43), manner deictics may develop into propositional anaphors and have whole clauses as antecedents, e.g., *The meeting has been postponed? I suppose so*. Anaphoric deictics can also turn into connectives expressing different circumstantial relations, e.g., causal (*I didn’t like it, so I wrote to him.*), concessive (*John is very sick. Even so he goes to work*) or conditional. Another well-documented change involving cataphoric demonstratives is the development of quotative markers (Güldemann 2008). Manner deictics are also frequently found in other grammatical markers, which are especially relevant for the discussion of Basque data: affirmative particles, additive markers, comparative markers and subordinators.

As regards affirmative particles, according to König (2015),

Affirmative particles typically derive from manner deictics. In the languages under discussion Italian *si* and French *si* (after negative questions) derive straightforwardly from the Latin manner deictic *sic*. English *yes* is the result of a fusion between *yeah* and *swa* and Polish *tak* means both ‘so’ and ‘yes’. Moreover, narratives can be confirmed by expressions like *So ist es* (German) or *It must have been so* and prayers, wishes and plans for the future by phrases like *So be it* or *Ainsi soit-il* (French) in a wide variety of languages (König 2015: 44).

The Basque *bai* has been used as an affirmative adverb ‘yes’ in answer to a question since the earliest texts. It can also reinforce affirmation (32), emphasize an opposition to a previous negation (33) or function as an adversative conjunction (34) (OEH s.v. *bai*).

- (32) *Eta harekin ioai-ten ziraden gizon-ak, geldi zitezen*
 and he.COM go-IPFV AUX.PST.3PL man-DEF.PL stay-RAD AUX.AOR.3PL
izitu-rik, haren boza bai enzu-ten, baina nehor ikus-ten
 frighten-PART his voice yes hear-IPFV but nobody see-IPFV
e-tzute-la.

NEG-AUX.PST.3PL>3SG-COMP

‘And the men who were going with him got frightened, they did hear his voice, but they saw nobody.’

(LeizarragaTest, Acts 9: 7, 1571)¹⁸

- (33) *Bada baldin adultera-tzen ez-pa-duk, baina*
 So if commit.adultery-IPFV NEG-COND-AUX.2SG>3SG but
hil-tzen bai, Lege-a-ren transgredizale egin aiz.
 kill-IPFV yes law-DEF-GEN transgressor make AUX.2SG
 ‘For, if you do not commit adultery, but you do kill, you have thus become a transgressor of the Law.’

(LeizarragaTest, James 2: 11, 1571)¹⁹

- (34) *zuiek ukhen baituzuie hezaz guti ansia. / Bai erhoki*
 you.ERG have bait.AUX.2PL>3SG DEM.INS little care but madly
konplazitu ene kontra etsai-a
 please my against enemy-DEF

‘you did not care too much for him at all, but instead you madly pleased against me the enemy’

(Etxepare, 1545)

The Basque *bai* can also be used in various coordinating-additive constructions. According to König (2015: 45), “(...) anaphorically used manner demonstratives develop into additive focus markers (Engl. *also*, Fr. *aussi*, Germ. *ebenso*, Swed. *også*, Russ. *takže*) and further into coordinating conjunctions”. Moreover, *bai* also signals ellipsis, usually of the verbal phrase. This is similar to the behavior of the English *so* (*John is writing a book and so is MARY*) or the French *aussi* (König 2015: 45).

18 The French source: “Or les hommes que cheminoyent avec luy, s’arrestèrent estonnez, oyans bien la voix d’iceluy, mais ne voyans personne”.

19 The French source: “Si donc tu n’es point adultere, mais tu tues, tu es transgresseur de la Loy”.

While it is not very common, the particle *bai* on its own can be used as a conjunction meaning ‘as well as, and also’:

- (35) *Cec ecarri çaitu orain decreto-a egui-te-ra, bay*
 what.ERG bring AUX.3SG>2SG now order-DEF make-NMLZ-ADL yes
Vizcaya yl-te-ra?
 Biscay kill-NMLZ-ADL
 ‘What has brought you now to give orders and kill Biscay?’
 (1688, Lakarra 1984)

Another group of constructions involves *eta* ‘and’ and *ere* ‘also’ apart from *bai*. For example, *bai* is used in elliptical affirmative clauses together with the additive particle *ere* ‘also’. *Bai* is not obligatory, but there is a tendency to use it:

- (36) *Zuk badakizu eta ni-k ere (bai)*
 you-ERG know.2SG>3SG and I-ERG also yes
 ‘You know and I do too’ (Hualde 2003: 327)

Importantly, *bai* cannot be used when there is no verbal ellipsis. This suggests that rather than expressing addition, *bai* signals that a part of the verb phrase was elided:

- (37) *joan-go da eta ni ere bai*
 go-FUT AUX.3SG and I also yes
 ‘She will go, and me too’
- (38) **joan-go da eta ni ere bai joan-go naiz*
 go-FUT AUX.3SG and I also yes go-FUT AUX.1SG
 ‘She will go, and I will go too’

Interestingly, verbal ellipsis is also frequent in the emphatic use of *bai* discussed earlier (as in Examples (32), (33) and (34)).

Furthermore, there is *bai eta*, which means ‘and also’ (39). It is frequently contracted to *baita*, but it is attested in its full form (e.g., *baieta ere* in Etxepare). In modern Basque it is usually reinforced with the particle *ere* (OEH s.v. *baita*). However, there are numerous examples in texts without it.

- (39) *Iainko-a bera ere Saindu dei-tzen da baieta haren*
 god-DEF himself also saint call-IPFV AUX.3SG and.also his
Aingeru-ak
 angels-DEF.PL
 ‘God himself is also called saint, and also his angels’
 (LeizarragaTest, Dictionary, 1571)

Finally, in the oldest texts we find *bai* as standard marker in equative constructions (Lafon 1999 [1957]: 651–652). Equatives express equal extent and often consist of the following elements (Haspelmath and Buchholz 1998: 279):

- (40) *My sister is as tall as you*
 CMP PAM PARA STM STAN
 CMP – comparee, PAM – parameter marker, PARA – parameter, STM – standard marker, STAN – standard

In many languages equatives are similar to similitive constructions, which express equal manner, e.g., *He sings like a nightingale* (Haspelmath and Buchholz 1998: 313). According to König (2015: 50), manner (or, more precisely, degree) demonstratives are used frequently as degree markers (German *so gross wie Karl* ‘as tall as Karl’), but sometimes also as standard markers (or both: *as tall as Charles*; *as < alswa* ‘quite so’).

In 16th-century Basque texts, *bai* is occasionally encountered as a standard marker in equative (41) and similitive constructions (42). According to OEH, this use is not attested in any other text afterward.

- (41) *Egundano izan daia ni bai ditxa-tako-rik?*
 till.today be AUX.3SG.QUEST I bai luck-RM-PART
 ‘Has there been anyone as unlucky as I am?’
 (Etxepare, 1545)
- (42) *Elias zen gu bai pasion-en suiektione-tako gizon-a*
 Elias be.PST.3SG we as passion-GEN.PL subject-RM man-DEF
 ‘Elias was a man subject to passions as we are’
 (LeizarragaTest, James 5: 17, 1571)²⁰

To conclude the discussion of the functions of *bai*, we can say that it appears or appeared in several contexts in which manner deictic elements tend to be found according to König (2015): as affirmative particle, additive marker and standard marker. Though *bai* is not morphologically related to demonstratives (*hau* ‘this’, *hori* ‘that’, etc.), taking into account its etymology – **badadi* ‘it is (so)’, ‘let it be (so)’ – it does not seem impossible that in the process of grammaticalization, it acquired anaphoric functions. This could happen because **badadi*, uttered in response to a question, refers back to the content of that utterance, and even though there is no overt deictic element, it contains a zero anaphor.

4.2 From *bai* to *bait-*

In this section, I will put forward several hypotheses on how the marker *bait-* emerged and became a subordinator. The reconstruction proposed here for Basque can be divided in various stages, which are represented in Figure 6. In what follows, I

²⁰ The French source: “*Elie estoit homme suiet à semblables passions que nous*”.

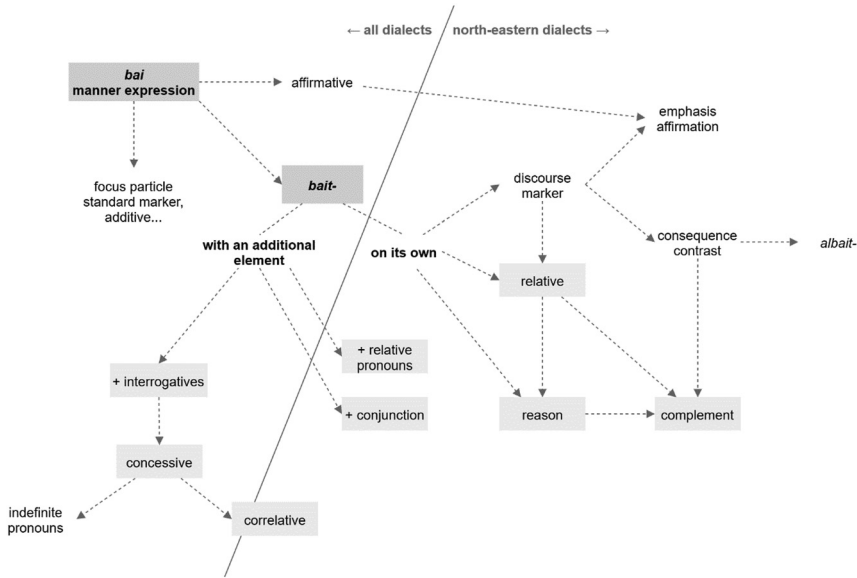


Figure 6: Map of proposed extensions of *bai* and *bait-*. Subordinate clauses are marked with light gray color.

will first focus on the development of the marker from the manner expression. Then I will try to explain how *bait-* started to be used in complex constructions and was eventually reanalyzed as a subordinator.

4.2.1 Development of *bait-*

The first question is how *bait-* developed from *bai*. Concerning the morphology of *bait-*, voiceless stops in final position are secondary in Basque and typically result from the apocope of the final vowel, which then triggers the devoicing of the final consonant if it is voiced (Mitzelena 2011 [1977]: 6.1, 6.4, 12.10). This suggests that the marker had some additional material.

The first possibility could be the following: *bait-* < **baita* < *bai* ‘yes, thus’ + *eta* ‘and’. This possibility is supported by the existence of a few conjunctions with *eta*, for example, *zein eta* ‘which, that, as’ or *zeren eta* ‘because’. The addition of *eta*, at least in historical sources, does not seem to alter the meaning of the conjunction. An argument possibly against this idea is related to the existence of the particle *baita* ‘and also’ (attested as *bai eta* as well), mentioned earlier. As pointed out by Etxepare (2001, 2021) *bait-* and *baita* are different in that *baita* ‘and also’ does not have *baist-* or *beit-* variants. However, as shown by Padilla-Moyano (2017: 703), the Souletin variant *beit-*

is rather recent. The lack of form with the sibilant could mean that the particle *baita* is more recent than the subordination prefix. Nevertheless, given that the renewal of such markers is a common process, this is not necessarily a reason to reject this hypothesis.

Another possibility is that *bait-* comes from *bai* and that *-t-* is simply an effect of being placed frequently before *d-* and *z-* starting verbs. Such an explanation was put forward by Lafitte (1991 [1944]), according to whom the consonant *-t-* can be explained through reinforcement of *-d-* and *-z-* after diphthongs (Lafitte 1991 [1944]: Section 405), but the parallels he gave are rather marginal, and the explanation does not appear very likely. Nevertheless, there are varieties (e.g., the now extinct Roncalese), where the prefix appears to be *bai*, rather than *bait-* (Mitzelena 2011 [1954]: 643), but it is unclear whether it is secondary (reinterpreting the marker as *bai-*) or is an archaism.

Finally, the final consonant might come from the copula: *bait-* < **baida* < *bai* + *da* 3sg of *izan* ‘be’. As mentioned, manner deictics show a tendency for renewal (König 2015: 41–42). Adding another finite copular verb form to *bai* would be an example of partial renewal.

Tentatively we can say that the last explanation appears the least problematic. Pérez Saldanya (2020) mentions the first and the second hypotheses and concludes that both are possible, but the first appears simpler because it allows explaining the development of *bait-* starting directly from *bai*. However, if *bait-* emerged through partial renewal, the explanation can also be based on what we know about *bai*, because we might assume that the newer form would inherit functions and properties of the older form. In any case, the starting point would be *bai*, a manner expression, which functioned as an affirmation marker and as a focus particle (at least as an additive focus particle and standard marker). Then, its reinforced form emerged: *bait*. Initially, it would have functions similar to the older variant, i.e., manner deictic and focus particle.

4.2.2 Indefinite pronouns and *bait-* plus conjunction/pronoun clauses

In a subsequent stage, the particle *bait-* started to appear in complex constructions, initially as a focus particle, not a marker of subordination. It was placed in the second position of the clause, similarly to the particles *ere* and *eta* (as in the already mentioned conjunction *zeren eta* ‘because’). An early construction with *bait* could be the one that brought the emergence of indefinite pronouns in *-bait*. As explained, the subordinating *bait-* is found in the northern and eastern varieties of Basque (with a few exceptions), but the series of pronouns in *-bait* is general. Thus, we can hypothesize that pronouns developed earlier than the subordinator. Etxepare (2001)

proposed that the source of the pronouns was a free relative clause (the “it may be” type in Haspelmath [1997]):

- (43) *nor baita* > *norbaita* > *norbait*
 who *bait.be.3sg*
 ‘who is’ > ‘someone’

Forms with the final *-a*, which provide evidence for this idea, are attested, as noted by Etxepare (2001): e.g., we find *cerbayta* ‘something’ in 1537. Additionally, there are very few instances where *bait-* remains an independent particle, e.g., *nork bait* in the poems by the 16th-century author Bernard Etxepare with the interrogative pronoun in the ergative case (it later became *norbaitek*, externalizing the inflection).

Alternatively, also building on Haspelmath (1997), we could start from the so-called ‘parametric concessive conditional clause’, illustrated in (44). The concessive clause might then be grammaticalized into an indefinite pronoun, (e.g., in French *qui que ce soit* ‘anyone’). Haspelmath (1997: 137) lists the following characteristics of this type of clause: (a) the verb is typically in some kind of subjunctive, (b) focus particles such as ‘also, even’ are often present, (c) there might be a conditional marker, (d) a pleonastic negator, (e) additional general subordinator, or (f) an emphatic particle.

- (44) *You can take something, whatever it may be* (concessive conditional)

The concessive clause which could be the source of indefinite pronouns (45) is attested in texts (47), though typically with the particle *ere*. This scenario also allows the inclusion of the *baista* variant in the argument (46). Etxepare (2001) mentions that the fricative could be a reflex of *ez* ‘no’ (the laminal sibilant, <z>, is likely to become apical, <s>, before /t/),²¹ but does not find an explanation for it. However, as argued by Haspelmath, pleonastic negators are common in that kind of clause.

- (45) *zer bait da* > *zerbait(a)*
 what *bait* is
 ‘whatever is’ > ‘something’
- (46) *zer bait ez da* > *zerbaist(a)*
 what *bait* _{NEG} is
 ‘whatever is (not)’ > ‘something’
- (47) *Zer e[re] baita, eduki-ko-ren dut kontu jaki-te-ra.*
 what _{PTCL} *bait.be.3sg* have-FUT-FUT _{AUX.1SG>3SG} care know-NMLZ-ADL
 ‘Whatever it might be, I will take care to find out.’
 (1595, Orpustan 2010)

21 An alternative solution, proposed by Lafon (1999 [1966]: 683), is to link the sibilant to the instrumental case.

Moreover, if *-s-* in *baist* goes back to the negative marker, it must have been added before *bait-* became a verbal prefix, because it is placed between *bait-* and the verb (the modern order would be negation-*bait-*verb and not *bait-*negation-verb). This order would be possible if *bait* was a particle linked to the interrogative pronoun. Thus, we can propose that initially in these concessive clauses *bait-* was not a fully-fledged subordinator, but rather a particle with a focus function. This is convenient: it is simpler to think that the subordinator did not fully develop in Western varieties than to think it was lost there and was only maintained elsewhere.

It is difficult to decide whether the departure point for the grammaticalization of indefinite pronouns was a relative or a concessive construction. As pointed out by Haspelmath and König (1998: 577), in many languages both exhibit similar structure and change from one to another can occur (the change from concessive to relative implies, among others, a tighter integration: contrary to concessive clauses, free relatives occupy a slot in the main clause).

The use of the focus particle also extended to other kinds of clauses: it appeared after a conjunction or an interrogative pronoun in the second position of the clause. It did not mark subordination, but frequent usage turned it into the unmarked option and, since it often happened to be placed before the finite verb, it cliticized to the verb and was eventually reinterpreted as a subordinator. This process proceeded only to a limited extent in the western dialects, where *bait-/baist-* is attested only a few times in correlative structures (48) In eastern/northern varieties, we find a broader range of clauses: relative and adverbial. In such clauses, a subordinator (*-en* or *bait-*)²² is obligatory in the modern language, though in the older texts some types of clauses, e.g., those with *zeren* ‘because’ appear sometimes without any subordinator.

- (48) *Zelan-go-a baista ame-a alan-go-a oi da*
 how-RM-DEF bait.be.3SG mother-DEF such-RM-DEF usually be.3SG
alabe-a
 daughter-DEF
 ‘Such mother, such daughter’
 (1596, Lakarra 1996)

4.2.3 Discourse function and *bait-* only clauses

Parallel to these developments, but probably only in the northern/eastern varieties, *bait-* acquired the function of a clause-level discourse particle, along the lines of Fraser’s (1988: 21–22) definition: discourse particles “signal a comment specifying the

²² It is impossible to know which of them, *bait-* or *-en*, is older. If it is *-en* (which is used in all dialects), then *bait-* could replace it in some contexts.

type of sequential discourse relationship that holds between the current utterance – the utterance of which the discourse marker is a part – and the prior discourse”. The particle, which did not necessarily appear in subordinate clauses, would express meanings related to consequence, contrast and emphasis. It was placed before the finite verb, as it is the typical position for particles in Basque, such as the modal particles *omen* ‘hearsay information’ or *bide* ‘apparently’ (Hualde and Ortiz de Urbina 2003: 316). The use of *bait-* in independent clauses is attested since the oldest texts (see Examples (13) and (14)) and its reflex is also found in the archaic *al-bait-* prefix (see (18)). The presence of *bait-* in exclamatives might also be related to this function.²³

Finally, we have *bait-* only subordinate clauses (relative, reason, result or complement). The presence of *bait-* in these constructions can be linked to the manner deictic, but the discourse marker function of *bait-* could also influence the use of *bait-* in such clauses.

In the oldest texts, as shown earlier, *bait-* only clauses perform already very diverse functions, but relatives appear to be the most frequent type. Furthermore, relative clauses are the basis of old place names with *bait-* (Salaberri Zaratiegi 1996), which suggests that the use of the marker in such clauses is quite old, and it could be a candidate for the initial stage of the extension of the marker. Deictic elements, such as demonstratives, are suitable as markers of relative constructions because of their capability to refer back to a referent in the previous discourse (Heine and Kuteva 2002, 2007; Hendery 2012; Kuteva et al. 2019). Also, manner deictics can appear in relative clauses, as shown by König (2015: 52). The particle *bai* has anaphoric functions, and, thus, if *bait-* developed from it, its use in relative constructions is not surprising.

Turning to adverbial clauses, cross-linguistically, manner words can be found in consequence clauses or reason clauses. According to Kortmann (1998) in European languages, the most common source construction for the expression of cause is temporal followed by manner. König (2015; 2020) lists various adverbial relations as possible extensions of manner deictics.

As regards Basque, reason clauses appear already in the oldest texts, but they became the most prominent function of the marker only later. The presence of *bait-* in these clauses can be explained in two ways and both possibly contributed to the increase of their frequency.

²³ As noted by a reviewer, independent uses of *bait-* could emerge via insubordination. This explanation cannot be ruled out completely as both independent and subordinate uses of the marker are attested already in the oldest texts. However, the insubordination analysis appears problematic because it is not clear what kind of construction(s) could be the source. For exclamatives, factive complements could be suitable, but texts suggest that such complements are rather recent.

The first one starts from accord/comment clauses like (12), with *bait-* in the first clause, which refers to the background information or, more generally, links what is to be said in the main clause to the previous discourse (compare English *as* in e.g., *As I have mentioned ...*). It is possible to strengthen this meaning to reason: from something like ‘so you have said that ...’ to ‘since you have said ...’. The grammaticalization of manner expressions (such as the Romance conjunctions derived from the Latin *quōmodo* or the English *as*) into causal conjunctions tends to start precisely in such a context, where the clause containing the manner expression appears as the topic at the margin of the main clause (Pérez Saldanya 2020: 594; Pérez-Saldanya and Hualde 2017).

The second route goes from relative clauses to reason.²⁴ Many examples with *bait-* are ambiguous between these two meanings, especially in the 17th–18th centuries, just before the sharp increase in the proportion of reason clauses. The ambiguity happens when a participant appears both in the subordinate and the main clause, as in (49). Taking into account relative and reason clauses (including ambiguous ones), 26 out of 27 clauses in the 16th century and 83 % in the second period have co-referential participants in the main and subordinate clauses. This proportion is 77 % in the third period and 60 % in the most recent data.²⁵ Thus, the ambiguous clauses may provide a bridging context for the change from relative to reason.

- (49) *E-tzaye eman behar aza osto-ric, ema-iten*
 NEG-AUX.3SG<3PL give must cabbage leaf-PART give-IPFV
bei-teyo aragui-a-ri khiño gaisto bat.
 bait-AUX.3SG<3SG meat-DEF-DAT smell bad one
 ‘They should not be given cabbage leaves because they / which give the meat bad smell.’
 (IntxauspeSolast, 1857)

Finally, we have complement clauses. The use of *bait-* in them be explained differently depending on the subtype of complements: verbs of happening, nominal complements and factive complements.

In the oldest texts *bait-* complements appear with verbs of happening and in complements of nouns. Manner deictics can be grammaticalized into quotation

²⁴ Such process was proposed for several languages (see examples in Disterheft and Viti 2010: 244–245; Hendery 2012: sec. 2.2.6), though the opposite direction, i.e., from adverbial to relative is also possible.

²⁵ Using a subset of the data (*bait-* only clauses that can be interpreted as relative or reason), a regression analysis was performed with the co-referentiality as the dependent variable and the period as the independent variable (and text as a random effect). The effect of the period is significant ($\chi^2 = 13$, $df = 3$, $p = 0.004$). See details in the supplementary materials.

markers and from there into complement clause markers (from ‘he said so: “ ... ” to ‘he said that ... ’ [Güldemann 2008; König 2015]). Nevertheless, *bai* or *bait* are not attested as cataphors. Alternatively, the particle could initially have discourse function (50) (cf. Example (30)). This scenario explains well the presence of *bait-* in complements of ‘happen’ verbs.

- (50) *gertha zedin hil bait zen*
 happen AUX die bait AUX
 ‘it happened: he (indeed) died’
 (LeizarragaTest, Luke 16: 22, 1571)

Another possible route would go from relative to complement clauses. As mentioned, nouns take *bait-* complements in early texts. Relative clauses are typical modifiers of nouns, and because of that, their use could be extended to other types of nominal structures. This grammaticalization path is attested in various languages (relative > complementizer grammaticalization path in Kuteva et al. [2019: 367]).

Finally, *bait-* complements of factive verbs could emerge via reanalysis of reason clauses: from something like *I am surprised because ...* it is easy to arrive at *I am surprised that ...*, and in Basque texts there are examples – though not numerous – ambiguous between reason and complement. This may explain the use of *bait-* with factive emotive verbs: it is possible to express a reason for an attitude towards a fact (*I am surprised because ...*), but not with many other kinds of complement-taking verbs (e.g., *say*). Also, this path has crosslinguistic parallels. Kortmann (1997: 64) shows that markers of adverbial clauses can be placed in the middle of the complementizer – adverbial subordinator – relativizer – relativizer continuum. López-Couso and Méndez-Naya (2015) analyze various ‘minor complementizers’ in English, which are grammaticalized from adverbial clause markers such as *but*, *as if*, *if* or *like*. They argue that the adverbial subordinators in complement clauses show a higher degree of integration: the subordinate clause becomes an argument in the main clause. The grammaticalization process also results in partial loss of the original adverbial meaning, even though its traces continue to constrain the range of predicates with which such complements are possible. The Basque case happened in a different domain of adverbial subordination (the various English markers do not include causal conjunctions), but the grammaticalization operated similarly. The subordinate clause became increasingly integrated into the main clause (cf. hypotaxis > subordination cline as described by Hopper and Traugott [2003]), even though in the modern language the two clauses need not (yet) be in a single intonational phrase. The meaning of cause undergoes bleaching, but it continues to influence the use of *bait-* complements by restricting it to predicates with semantics compatible with it.

5 Conclusions

The main goal of this article was to explain in what way the Basque subordinator *bait-* might be related to the affirmative adverb *bai* ‘yes’. I have argued that the missing link could be a manner expression from which both the affirmative adverb and the subordinator developed. Earliest texts provide some support for the idea, as *bai* is used there in a range of functions typical of grammaticalization pathways of manner deictic expressions. The etymology proposed by Trask (1997) and Lakarra (2018) for the adverb (**badadi* ‘it is (so)’) is also compatible with the idea. The marker *bait-* could come from *bai* plus either the conjunction *eta* ‘and’ or the verb form *da* ‘is’. In either case, given the tendency of manner deictics to renew, such developments are not surprising.

The exact stages of the development of the subordinator are difficult to reconstruct, as practically all of its functions are already attested in the earliest Basque texts. However, it appears plausible that initially, the particle had focus or discourse functions. A reflex of this is found in different areas of Basque grammar (e.g., in indefinite pronouns). Due to the frequent use in subordinate clauses, it was eventually recategorized as a subordinator. It seems that the extension of the marker *bait-* to the different types of subordinate clauses (and to other areas of grammar) happened in a piecemeal fashion, from one construction to another. In the modern language *bait-* is a subordinator with a wide range of functions, but it retains certain semantic features related to the earlier stages.

As shown throughout the article, the changes that brought the emergence of the subordinator as well as the subsequent developments can be described as a grammaticalization process. In particular, the following features of grammaticalization are prominent:

- (a) Phonetic erosion: in *bait-* the *-t* must be secondary and is probably the result of the apocope of the final vowel. Additionally, the particle lost its independence and cliticized to the verb.
- (b) Decategorialization: a change in the categorical status occurred when a particle was reinterpreted as a subordinator.
- (c) Bleaching: the marker partially lost its original lexical meaning, especially in contexts where it is used with another element which determines the type of clause.
- (d) Persistence: at the same time, however, the influence of the older lexical meaning can also be seen, e.g., in the range of predicates which take *bait-* complements.
- (e) Layering: the newer uses of the marker add to the older ones.

- (f) Extension and obligatorification: throughout the process, the subordinator extended to new contexts: via the emergence of new constructions (such as [conjunction + *bait*-verb]) and also through the reanalysis of certain structures (e.g., reason to complement). In certain constructions the subordinator became obligatory.

Additionally, increased clausal integration can also be mentioned as a feature of grammaticalization processes across clauses (Hopper and Traugott 2003): in some clause types a change towards embedding has happened (or is still underway). This is especially clear in complement clauses, which, according to the arguments presented here, are a late addition to the family of *bait*-clauses and emerged from more loosely connected structures (e.g., reason clauses).

Abbreviations

ADL	adlative
AFF	affirmative
ANIM	animate
AOR	aorist
AUX	auxiliary
COND	conditional
COM	comitative
COMP	complementation
DAT	dative
DEF	definitive
DEM	demonstrative
DEST	destinative
ERG	ergative
FUT	future
GEN	genitive
HYP	hypothetical
IMP	imperative
INDEF	indefinite
INES	inessive
INS	instrumental
IPFV	imperfective
NEG	negation
NMLZ	nominalization
PART	partitive
PL	plural
POT	potential
PST	past
PTCL	particle

QUEST	question particle
RAD	radical
RM	relational marker
SG	singular
SUB	subordinator

In glosses of finite verbs the sign “>” distinguishes ergative and absolutive arguments and “<” distinguishes dative ones.

Research funding: The research for the article was made possible by the grants from the Spanish Ministry of Science and Innovation (PGC2018-098995-B-I00, PID2021-124769NB-I0 and PID2020-118445GB-I00). Support given by the Diachronic Linguistics, Typology, and the History of Basque Research Group (IT1534-22) funded by the Basque Government is also acknowledged.

Data availability: Supplementary materials for this article (data and code used for the quantitative part of the study) can be accessed at <https://osf.io/uqcf5/>.

Appendix: Corpora used in the study

The quantitative part of the article is based on a corpus consisting of texts listed in the following table. Unless stated otherwise, the texts come from the *Euskal Klasikoena Corpusa* corpus (Euskara Institutua 2013). The spelling of the texts included in this corpus was modernized.

Abbreviation	Author and title
Period 1 (1500–1600)	
Ettxepare	Ettxepare, <i>Linguae Vasconum Primitiae</i> (1545)
LeizarragaAbc	Leizarraga, <i>ABC edo Kristinoen instrukzioea</i> (1571)
LeizarragaKat	Leizarraga, <i>Katexisma</i> (1571)
LeizarragaTest, Matt	Leizarraga, <i>Iesus Krist Gure Jaunaren Testamentu Berria</i> (1571), The Gospel of Matthew
Period 2 (1600–1750)	
Materra	Materra, <i>Dotrina kristiana</i> (1623)
BeriainDotrina	Beriain, <i>Dotrina kristioarena euskaras</i> (1626)
EttxeberriZiburukoa	Ettxeberri Ziburukoa, <i>Manual debozionezkoa</i> (1627, 1669), <i>Noelak</i> (1630, 1645)
Haranburu	Haranburu, <i>Debozino eskuarra</i> (1635)
Axular	Axular, <i>Gero</i> (1643), chapters I–XV
Oihenart	Oihenart, <i>O.ten gaztaroa neurtizetan</i> (1657)
Harizmendi	Harizmendi, <i>Ama birjinaren ofizioa</i> (1658)
Pouvreau	Pouvreau, <i>Filotea</i> (1664), chapters 1–14 and <i>Iesusen imitazioea</i> (1669), chapters 1–13
Argainaratz	Argainaratz, <i>Deboten brebiarioa</i> (1665)

(continued)

Abbreviation	Author and title
TartasOnsa	Tartas, <i>Onsa hilzeko bidia</i> (1666)
Aranbillaga	Aranbillaga, <i>Jesu Kristoren Imitazionea</i> (1684)
Gazteluzar	Gasteluzar, <i>Egia katolikak</i> (1686)
Belapeire	Belapeire, <i>Katexima labürra</i> (1696)
Oloroeko kat.	<i>Oloroeko katixima</i> (1706) (Padilla-Moyano 2015)
EtxeberriSarakoa	Etxeberri Sarakoa, <i>Eskual-Herriko gazteriari</i> (c. 1718) and <i>Lau-Urduri gomen-diozko karta, edo guthuna</i> (1718)
Xurio	Xurio, <i>Jesu-Kristoren imitazionea</i> (1720)
Haraneder	Haraneder, <i>Jesu Kristoren ebanjelio saindua: San Mateo</i> (1740)
Period 3 (1750–1900)	
Maister	Maister, <i>Jesü-Kristen imitaziona</i> (1757), the 1st and 3rd books
Larregi	Larregi, <i>Testamen berriko historia</i> (1777)
Lizarraga	Lizarraga, <i>Zenbait sanduen biziak asteaz datozinak</i> (1793–1813), first 35 chapters
Mihura	Mihura, <i>Andredena Mariaren imitazionea</i> (1778)
IntxauspeApokalipsia	Intxauspe, <i>Apokalipsia</i> (1857) (Pagola et al. 1997)
IntxauspeSolast	Intxauspe, <i>Iturriagaren Solastaldiak</i> (1857) (Pagola et al. 1997)
Duvoisin	Duvoisin, <i>Laborantzako liburua</i> (1858) and <i>Solastaldiak</i> (1857)
Laphitz	Laphitz, <i>Bi Saindu Heskualdunen Bizia</i> (1867)
Webster	Webster, <i>Euskal ipuinak</i> (1877)
Adema	Adema, <i>Eskualdun peleginaren bidaltzailea</i> (1877)
Arbelbide	Arbelbide, <i>Igandea edo Jaunaren eguna</i> (1895)
Hiriart-Urruti	Hiriart-Urruti, <i>Gontzetarik jalgiaraziak</i> (1891–1914)
Elizanburu	Elizanburu, <i>Lehenagoko Eskualdunak zer ziren</i> (1889)
Period 4 (1900–1970)	
EtxepareBuruxkak	Etxepare Bidegorri, <i>Buruxkak</i> (1910)
Barbier	Barbier, <i>Piarres I</i> (1926)
LeonImitazionea	Leon, <i>Jesu-Kristoren imitazionea</i> (1929)
EtxepareMendekoste	Etxepare Landerretxe, <i>Mendekoste gereziak</i> (1962)
Larzabal	Larzabal, <i>Roxali</i> (1970), <i>Matalas</i> (1968), <i>Antzerki laburrak</i> (1930–1970)

References

- Artiagoitia, Xabier. 2003. Complementation (noun clauses). In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 634–709. Berlin & New York: Mouton de Gruyter.
- Artiagoitia, Xabier & Arantzazu Elordieta. 2016. On the semantic function and selection of Basque finite complementizers. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages*, 379–411. Berlin & Boston: De Gruyter Mouton.

- Artiagoitia, Xabier, Bernard Oyharçabal, José Ignacio Hualde & Jon Ortiz de Urbina. 2003. Subordination. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 632–762. Berlin & New York: Mouton de Gruyter.
- Azkue, Resurrección María. 1923. *Morfología vasca*. Bilbao: Editorial Vasca.
- Bourdin, Philippe. 2008. On the grammaticalization of ‘come’ and ‘go’ into markers of textual connectivity. In María José López-Couso & Elena Seoane (eds.), *Rethinking grammaticalization: New perspectives*, 37–59. Amsterdam & Philadelphia: John Benjamins.
- Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago, IL: University of Chicago Press.
- Creissels, Denis & Céline Mounole. 2011. Animacy and spatial cases: Typological tendencies, and the case of Basque. In Seppo Kittilä, Katja Västi & Jussi Ylikoski (eds.), *Case, animacy and semantic roles*, 157–182. Amsterdam & Philadelphia: John Benjamins.
- Disterheft, Dorothy & Carlotta Viti. 2010. Subordination. In Silvia Luraghi & Vit Bubenik (eds.), *Continuum companion to historical linguistics*, 230–249. New York: Continuum.
- Dixon, Robert M. W. & Alexandra Y. Aikhenvald. 2009. *The semantics of clause linking: A cross-linguistic typology*. Oxford: Oxford University Press.
- Etxepare, Ricardo. 2001. *Nz-bait* izenordain zehaztugabean jatorriaz [On the source of *wh-bait* indefinite pronouns]. *Lapurdum* 6. 93–117.
- Etxepare, Ricardo. 2021. From free relatives to indefinite pronouns in Basque, or how can anyone end up being someone. Available at: <https://hal.archives-ouvertes.fr/hal-03505652/>.
- Euskaltzaindia. 1999. *Euskal gramatika: Lehen urratsak* [Basque grammar: First steps], vol. 5–1. Bilbao: Euskaltzaindia.
- Euskaltzaindia. 2019. *Orotariko Euskal Hiztegia* [General Basque Dictionary], 6th edn. Bilbao: Euskaltzaindia.
- Euskara Institutua. 2013. Euskal Klasikoen Corpusa (EKC) [Corpus of Basque classics (EKC)]. UPV/EHU. Available at: <https://www.ehu.eus/ehg/kc/>.
- Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguistica Hungarica* 38(1–4). 19–33.
- Gèze, Louis. 2010. *Éléments de grammaire basque: Dialecte souletin, suivis d'un vocabulaire basque-français & français-basque*. Munich: LINCOM Europa.
- Givón, Talmy. 2009. *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam & Philadelphia: John Benjamins.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10(2). 279–326.
- Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin & New York: Mouton De Gruyter.
- Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.
- Haspelmath, Martin & Oda Buchholz. 1998. Equative and similative constructions in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 277–334. Berlin & New York: Walter de Gruyter.
- Haspelmath, Martin & Ekkehard König. 1998. Concessive conditionals in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 563–640. Berlin & New York: Mouton de Gruyter.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
- Hendery, Rachel. 2012. *Relative clauses in time and space: A case study in the methods of diachronic typology*. Amsterdam & Philadelphia: John Benjamins.

- Hetterle, Katja. 2015. *Adverbial clauses in cross-linguistic perspective*. Berlin & Boston: De Gruyter Mouton.
- Hopper, Paul. 1991. On some principles of grammaticalization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 1, 17–35. Amsterdam & Philadelphia: John Benjamins.
- Hopper, Paul & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Hualde, José Ignacio. 2003. Conjunctions and connectors. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 323–327. Berlin & New York: Mouton de Gruyter.
- Hualde, José Ignacio & Jon Ortiz de Urbina (eds.). 2003. *A grammar of Basque*. Berlin & New York: Mouton de Gruyter.
- Karttunen, Lauri. 1971a. Implicative verbs. *Language* 47(2). 340–358.
- Karttunen, Lauri. 1971b. Some observations on factivity. *Paper in Linguistics* 4(1). 55–69.
- Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 143–173. Berlin: Walter de Gruyter.
- König, Ekkehard. 2015. Manner deixis as source of grammatical markers in Indo-European languages. In Carlotta Viti (ed.), *Perspectives on historical syntax*, 35–60. Amsterdam & Philadelphia: John Benjamins.
- König, Ekkehard. 2020. Establishing transphrastic relations: On the grammaticalization of demonstratives. In Janet Zhiqun Xing (ed.), *A typological approach to grammaticalization and lexicalization: East meets West*, 55–74. Berlin & Boston: De Gruyter Mouton.
- Kortmann, Bernd. 1997. *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin & New York: Mouton de Gruyter.
- Kortmann, Bernd. 1998. Adverbial subordinators in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 457–562. Berlin New York: Mouton de Gruyter.
- Krajewska, Dorota. 2017. *Euskararen sintaxi diakronikorantz: Egitura konplexuak/towards a historical syntax of Basque: Complex constructions*. Vitoria-Gasteiz: UPV/EHU dissertation. Available at: <http://hdl.handle.net/10810/26877>.
- Krajewska, Dorota. 2022. The marking of spatial relations on animate nouns in Basque: A diachronic quantitative corpus study. *Journal of Historical Linguistics* 12(2). 241–281.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.
- Lafitte, Pierre. 1991. *Grammaire basque: Navarre-labourdin littéraire*. Donostia: Elkar.
- Lafon, René. 1980. *Le système du verbe basque au XVIe siècle*. Donostia: Elkar.
- Lafon, René. 1999 [1957]. L'expression de la comparaison en basque. In Jean Haritschelhar & Pierre Charritton (eds.), *Vasconiana*, 649–666. Bilbao: Euskaltzaindia.
- Lafon, René. 1999 [1966]. La particule *bait* en basque : Ses emplois morphologiques et syntactiques. In Jean Haritschelhar & Pierre Charritton (eds.), *Vasconiana*, 667–687. Bilbao: Euskaltzaindia.
- Lafon, René. 1999 [1973]. La langue basque. In Jean Haritschelhar & Pierre Charritton (eds.), *Vasconiana*, 3–55. Bilbao: Euskaltzaindia.
- Lakarra, Joseba A. 1984. Bertso Bizkaitarrak (1688) [The Biscayan verses (1688)]. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 18(2). 89–183.
- Lakarra, Joseba A. 1986. Bizkaiera zaharra euskalkien artean [Old Biscayan among the Basque dialects]. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 20(3). 639–681.
- Lakarra, Joseba A. 1996. *Refranes y sentencias (1596): Ikerketak eta edizioa* [Refranes y sentencias (1596): Analysis and edition]. Bilbao: Euskaltzaindia.
- Lakarra, Joseba A. 2018. Euskararen historiaurrea [The prehistory of Basque]. In Joaquín Gorrochategui, Iván Igartua & Joseba A. Lakarra (eds.), *Euskararen historia* [The history of Basque], 23–229. Vitoria-Gasteiz: Eusko Jaurilaritza.

- Lakarra, Joseba A., Julen Manterola & Iñaki Seguro. 2019. *Euskal hiztegi historiko-etimologikoa (EHHE-200)* [The Basque historical and etymological dictionary]. Bilbao: Euskaltzaindia. Available at: https://www.euskaltzaindia.eus/index.php?option=com_ehhe&task=home&Itemid=1193&lang=eu.
- Le Nouveau Testament*. 1563. *Le Nouveau Testament, c'est à dire, la Nouvelle Alliance de Nostre Seigneur Jesus Christ*. Genève: Jean Bonne-Foy. ark:/12148/bpt6k314602j.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam & Philadelphia: John Benjamins.
- López-Couso, María José & Belén Méndez-Naya. 2015. Secondary grammaticalization in clause combining: From adverbial subordination to complementation in English. *Language Sciences* 47. 188–198.
- Mitxelena, Koldo. 2011 [1954]. La posición fonética del dialecto vasco del Roncal. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. VII, 611–647. Bilbao: UPV/EHU.
- Mitxelena, Koldo. 2011 [1964]. Textos arcaicos vascos. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. XII, 1–171. Bilbao: UPV/EHU.
- Mitxelena, Koldo. 2011 [1970]. Nombre y verbo en la etimología vasca. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. VII, 235–267. Bilbao: UPV/EHU.
- Mitxelena, Koldo. 2011 [1977]. Fonética histórica vasca. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. VI. Bilbao: UPV/EHU.
- Mitxelena, Koldo. 2011 [1981]. Lengua común y dialectos vascos. In Joseba A. Lakarra & Iñigo Ruiz Arzalluz (eds.), *Obras completas*, vol. VII, 517–344. Bilbao: UPV/EHU.
- Orpustan, Jean-Baptiste. 2010. Correspondance basque à la fin du XVIIe siècle (1595–1598): 20 lettres de renseignements sur la politique de Henri IV et la fin des guerres contre la Ligue. *Lapurdum* 14. 137–162.
- Oyharçabal, Beñat. 1987. *Etude descriptive de constructions complexes en basque, propositions relatives, temporelles, conditionnelles et consessives*. Paris: Université de Paris VII dissertation.
- Oyharçabal, Beñat. 1997. Hiru aditz aurizki zahar XVI. mendeko testuetan [Three old verbal prefixes in 16th century texts]. *Lapurdum* 2. 45–62.
- Oyharçabal, Beñat, Irantzu Epelde & Jasone Salaberria. 2009. *Norantz proiektua*. Available at: <http://www.norantz.org/>.
- Padilla-Moyano, Manuel. 2015. *Oloroeko katixima: Edizioa eta azterketa* [The catechisms from Oloroe: edition and analysis]. Unpublished manuscript.
- Padilla-Moyano, Manuel. 2017. *Zuberoako euskararen azterketa diakronikoa: XVI–XIX. mendeak/Analyse diachronique du dialecte souletin: XVI–XIX siècles*. Bordeaux: UPV/EHU & Université Bordeaux Montaigne dissertation. Available at: <http://hdl.handle.net/10810/26535>.
- Pagola, Rosa Miren, Juan Jose Iribar & Itziar Iribar (eds.). 1997. *Bonaparte ondareko eskuizkribuak: Zuberera* [Manuscripts from the Bonaparte's archive: Souletin Basque]. Bilbao: Universidad de Deusto.
- Pérez Saldanya, Manuel. 2020. El prefijo subordinante *baít-* y sus paralelos en el dominio occitano-catalán. In Iñaki Camino, Xabier Artiagoitia, Irantzu Epelde & Koldo Ulibarri (eds.), *Eibartik Zuberoara euskalkietan barrera*. Koldo Zuazori gorazarre, 581–599. Bilbao: UPV/EHU.
- Pérez-Saldanya, Manuel & José Ignacio Hualde. 2017. From theme to rheme: The evolution of causal conjunctions of temporal origin in Catalan. *Studies in Hispanic and Lusophone Linguistics* 10(2). 319–348.
- R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Rebuschi, Georges. 2008. On older northern Basque exclamatives in *ala*. In Xabier Artiagoitia & Joseba A. Lakarra (eds.), *Gramatika jaietan: Patxi Goenagaren omenez*, 691–708. Donostia: Gipuzkoako Foru Aldundia and UPV/EHU.

- Salaberri Zaratiegi, Patxi. 1996. Toponimia eta dialektologiaren arteko harremanen inguruan: Artaxonako jabego genitiboaz eta bestez [On the relation between the toponymy and dialectology: The genitive in Artaxona and other issues]. *Fontes Linguae Vasconum* 72. 223–234.
- Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2, 237–269. Cambridge: Cambridge University Press.
- Trask, Robert L. 1997. *The history of Basque*. London & New York: Routledge.
- Wiemer, Björn. 2019. On illusory insubordination and semiinsubordination in Slavic: Independent infinitives, clause-initial particles and predicatives put to the test. In Karin Beijering, Gunther Kaltenböck & María Sol Sansiñena (eds.), *Insubordination*, 107–166. Berlin & Boston: De Gruyter Mouton.