

# TYPOLOGICAL PASSPORT OF BASQUE AS A MINORITY LANGUAGE<sup>1</sup>

Nataliya A. Lavrova

MGIMO University

**Abstract.** The aim of this paper is to make an inventory of the most salient typological traits of Basque as an extant minority language, whose origin and typological status have been debated ever since the 16th century. Despite the well-meaning attempts to compile a unified standard of Basque, the majority of its speakers communicate using one of its numerous dialects. Apart from that, most of the contemporary coursebooks of Basque convey contradictory messages about its main grammatical features. This lack of uniformity acts as an impediment on the way to learning Basque or forming a well-defined concept of its typology. To enhance and spread knowledge about Basque typology, the present research adopts the methodology of contrastive-typological studies, namely, the method of typological passport developed by Professor V. D. Arakin. The empirical data for the research was drawn from the course- and textbooks recommended by *The Royal Academy of the Basque Language*. The research results revealed a number of universal typological traits shared by Basque with well-studied languages of the world, such as the presence of articles, cases, assimilation, post-modifying adjectives, etc. However, from the perspective of contrastive typology, it is the unique typological features that are most typologically and conceptually salient, such as voiced and devoiced aspiration, variations in the pronunciation of diaphonic sounds, non-systematic (non-pleonastic) marking of number, conceptual salience of specific numerals (2, 4, and 11), neutral and emphatic world-building patterns as well the absence of some models altogether (blending, shortening, back-formation). Research results show that the typological passport of Basque can be defined as an agglutinative language with suffixes, syntactic agreement and fusions at morphemic junctures. One of the avenues for further research is a multifactorial analysis with a view to obtaining statistical data about typologically significant categories of Basque.

**Keywords:** minority language, Basque, contrastive typology, typological passport

---

<sup>1</sup> English translation from the Russian text: Lavrova N. A. 2023. Tipologicheskij pasport minoritarnogo baskskogo yazyka [Typological passport of Basque as a minority language]. *Filologicheskie nauki v MGIMO [Linguistics & Polyglot Studies]*. 9(4). P. 26–41. DOI: <https://doi.org/10.24833/2410-2423-2023-4-37-26-41>

The Basque language is one of the minority languages of the Iberian Peninsula and includes nine dialects that differ significantly from one another (Amorrtu 1969; Ostrowski). The central Basque dialects of Gipuzkoan, Upper Navarrese and Labourdin are more open to linguistic innovation than Roncalese, which is spoken in the peripheral Basque areas (Trask 1996). Basque dialects are traditionally divided into northern (French) and southern (Spanish) (Bengtson 2011). What is known as Standard Basque (*euskara batua*) today is the result of the efforts of the Royal Academy of the Basque Language since 1960 to develop an official version of the language, which incorporates forms from many dialects (Libarona 2008). *Euskara batua* is based on the Central Gipuzkoan dialect, which certain features of Labourdin and Lower Navarrese also included.

The **aim** of this paper is to make an inventory of the main typological features of Standard Basque at all linguistic levels. The **problem** with this is that, despite the attempts to create a unified standard for the Basque language, most of its speakers speak one of the local variants.<sup>2</sup> Furthermore, a study of the practical and theoretical literature shows that many textbooks and research papers use both standardized and marginal features of the Basque language as examples, which makes it extremely difficult for learners to get a feel for the typical features of Standard Basque, and this is a significant roadblock to mastering the language.

### A Little History of Research into the Basque Language

Of particular note among the scholars who have studied the Basque language and played a pivotal role in establishing its genetic relationship and typological classification are Wilhelm von Humboldt (Humboldt 1959), Hugo Schuchardt (Schuchardt 1925), Ferdinand de Saussure (Saussure 1959), Christianus Cornelius Uhlenbeck (Uhlenbeck 1947), and others. However, the most significant works were produced by the Basque-French researcher Luis Michelena (Michelena 1977) and the Anglo-American Basque scholar Robert Lawrence Trask (Trask 1996).

Wilhelm von Humboldt visited the Basque country twice during his lifetime – once in 1799, and again in 1802. While his works contain some valuable observations on the typology of Basque, it for the most part suffers from a certain bias. The lack of a clear methodology and the fact that he worked exclusively with written texts led Humboldt to such erroneous conclusions that Basque was related to Iberian and Greek. His neglect of Latin borrowings and inaccurate analysis of the structure of the Basque sentence led to incorrect conclusions about the origins and workings of the Basque language

<sup>2</sup> The problem of distinguishing between the concepts of “local variant” and “dialect” in the Basque language is not the subject of this study.

The next milestone in the study of Basque was the work of Napoleon Bonaparte's nephew, Prince Louis Lucien Bonaparte, who published a work entitled *Le verbe basque en tableaux* (Bonaparte 2018) in 1869 comparing verb forms in various Basque dialects.

The late 19th century saw a surge of interest in the Basque language in Germany. The Basque Linguistic Society was created in 1886, publishing the journal *Euskara* from 1886 to 1889. Between 1893 and 1923, the famous novelist and creolist Hugo Schuchardt published a number of works on the Basque language (Schuchardt 1925). Robert Lawrence Trask was critical of Schuchardt, however, for his sloppy presentation and the overall unreadability of his studies. But his main mistake, according to Trask, was his incorrect analysis of the structure of the Basque sentence as passive. Although this does not diminish the significant contribution that Schuchardt clearly made to the study of Basque, namely, his description of the language's normative and verbal morphology and the fact that he identified traces of the Romance languages in its lexical composition, which Trask duly notes: "It is not too much to assert that Schuchardt found Basque an amateur pastime and left it a scholarly discipline" (Trask 1996: 60).

Taking his lead from Schuchardt, the Dutch researcher Christianus Cornelius Uhlenbeck also argued that Basque sentences were passive in nature, while simultaneously recognizing their ergative structure (Uhlenbeck 1947). Most of the articles he published concern the phonology and morphology of the Basque language. Uhlenbeck is known among Basque scholars as the originator of the theory of the origin of the prefix \*be-, which, according to him, is a morphological indicator of the names of body parts. Like some other researchers, Uhlenbeck insisted that Basque was genetically related to the Caucasian languages.

The most significant contribution to the study of the Basque language was made by Luis Michelena, who systematically applied the methodology of structural linguistics to the study of Basque. Historical phonology, verbal system, toponymy, anthropony, etymology, syntax, Romance superstrate, accentology – this is but an incomplete list of areas that Michelena investigated in his works. His greatest contribution is his research establishing systematic phonological correspondences in various Basque dialects and the study of Romance (including Latin) borrowings.

Modern research into the Basque language focuses on the problem of standardizing Basque and developing it into a unified language.

For example, the work of Concepcion M. Valadez et al (Valadez etl. 2014) is devoted to the issue of Basque acquiring the status of an unmarked language. The researchers note that teachers in educational institutions in the Basque Country expect immigrants to speak Spanish first and foremost, and then their native language, while knowledge of Basque is not only not mandatory, but is also perceived negatively if immigrants do not speak Spanish.

José Ignacio Hualde and Koldo Zuazo (Hualde, Zuazo 2007) note that both centrifugal and centripetal forces are involved in the process of standardization of the Basque language. The first attempts to describe the typology of Basque were made

as early as the 16th century. But most scholars described the typological features of the dialect they themselves spoke. The emergence of the Basque Nationalist Party did nothing to further the development of a standard Basque language due to the strong purist tendencies that dominated at the end of the 19th century.

An ethnographic analysis carried out by Anna Augustyniak (Augustyniak 2021) demonstrated that issues of linguistic self-determination are particularly acute for immigrants in the Basque Country – most make little attempt to master Standard Basque, using separation as the leading strategy of acculturation.

Placing the issue of standardization of Basque in the context of the revival of minority languages, Eduardo Apodaka, Asier Basurto and Auxkin Galarraaga (Apodaka etl. 2022) note that the early 21st century was marked by open antagonism towards the development of a single version of the Basque language, yet by the end of the 2000s the issue was being discussed at the state level. As a result, the government started to allocate grants to researchers who wanted to describe the typology of Basque with a view to developing a single written standard.

S. Dronova (Dronova 2020) carried out a content analysis, which allowed her to identify the most frequently used Basque words and concepts in newspaper political discourse and draw the conclusion that most of the political terms used help strengthen the image of the Basque Country. This serves one of the political goals of the Basque Country – the recognition of the Basque people as a nation that is equal in standing with the Spanish nation, as well as the establishment of special inter-ethnic partner relations between the region and Spain.

Using onomastic and linguacultural techniques, O. Chesnokova and L. Dzhishkariani (Chesnokova, Dzhishkariani 2017) studied modern Basque toponymy, defining its space and boundaries through the prism of linguistic contact with the Spanish language. Their analysis of toponymic material revealed that “an important feature of modern Basque toponymy in Spain is the existence of dual names (a Basque name and a Spanish name), reflecting the centuries of contacts between the Basque and Spanish languages” (Chesnokova, Dzhishkariani 2017: 1074). The main conclusion put forward by the researchers is that “the linguistic consciousness of modern Basque preserves the toponyms of original origin as a national treasure, although in some cases preference is given to the Spanish name (Bilbao, not Bilbo, for example” (Ibid.).

Our analysis of both early and modern studies of the Basque language thus demonstrates that the problem of developing a typological classification of a single variant is particularly challenging. Most works discuss issues of national identity and self-determination of the Basque people (Gorter, Cenoz 2011; Haddican 2007; Cots, Martin-Rubió 2008; Laka etl. 2012; Urrutia, Irujo 2008; Ray, Lete 2010), while the linguistic characteristics of a unified Basque language and an analysis of its main typological features are, with some exceptions (see the works of Michelena and Trask), not a priority area of these studies.

## Methodology

This paper uses the typological passport method developed by the famous Soviet and Russian scholar V. Arakin (Arakin 1989) to describe the typology of five differently structured languages. We employed the concept of a typological passport to analyse the main typological features of the Basque language in partial comparison with Russian and English. For empirical material, we used textbooks and teaching aids recommended by the Royal Academy of the Basque Language (see the Appendix).

## Results

### Phonetics and Morphology

The most significant discrepancy between ancient and modern Basque texts is revealed in the distribution of the grapheme *h* and the corresponding sound used in French Basque to represent the aspirated plosives *p*, *t*, and *k* (*ph*, *th*, and *kh*).

Aspiration in Basque is different from aspiration in English in the following respects: first, Basque aspiration can be both voiceless and voiced; second, the Zuberoan dialect features a nasalized, laryngeal aspiration that is characteristic of the intervocalic position (Lavrova 2019). In modern Basque, aspiration is not contrastive. In the northern dialects, the plosives [*p*], [*t*], [*k*] are traditionally aspirated, which is reflected by the phoneme *h*. The southern dialects have lost aspiration: cf. *ethorri* vs. *etorri* (“to come”). When used at the beginning of a word, the aspirated *h* represents a separate phoneme that differentiates the meanings of words: *hura* (“that one”) vs. *ura* (“water”). In the intervocalic position, the aspirated *h* is used to eliminate gaps, without distinguishing the meaning of words: *behar* (“to need something”), *aho* (“month”), *zahar* (“old”).

Approximately half of the native Basque nouns and adjectives begin with a vowel sound. Several hypotheses have been offered in the scientific literature regarding the origin and extensive use of the vocalic initial. According to the “null hypothesis,” the frequent use of the vocalic initial is explained by the phonotactics of Old Basque (Trask 1996). Another hypothesis states that the initial vowel is a prosthetic sound added to a syllable beginning with a consonant [*ibid.*]. According to Schuchardt (Schuchardt 1925), these vowels represent traces of ancient grammatical prefixes such as the definite or indefinite article. John Bengtson (Bengtson 1991) suggests that the vocal initial is an ancient classifier added to a noun depending on its lexical and semantic category. A third theory states that the initial vowels remained after the loss of the preceding consonant. The Vennemann (Vennemann 1994) refers to Ferdinand de Saussure’s theory of laryngeal consonants in the Proto-Indo-European language by calling these consonants “laryngeal,” which include voiceless alveolar and voiceless dental

fricatives. Following F. de Saussure, the author of the theory of laryngeal consonants in the Proto-Indo-European language, T. Vanneman (Vennemann 1994) calls these consonants “laryngeal”, which he includes among them non-hissing and non-whistle voiceless fricatives.

Unlike English, in which compound words are formed without changes in sounds at the juncture of two stems, Basque is characterized by **amorphic morphological seams** caused by changes in the quality or quantity of both vowels and consonants. Only a small number of compound words are formed by agglutination: cf. *ke* + *zulo* = *kezulo* (“smoke” + “hole” = “chimney”); *jo* + *aldi* = *joaldi* (“to hit/strike” + “occasion, reason” = “strike”). If the first stem consists of three syllables, then the final vowel of the third syllable is dropped: cf. *itsaso* (“sea”) + *gizon* (“person”) = *itsasgizon* (“sailor”). In a disyllabic first stem ending in -a, -o, -e, the final vowel turns into -a: cf. *etxe* (“house”) + *gain* (“summit”) = *etxagain* (“roof”); *baso* (“forest”) + *jaun* (“head”) = *basajaun* (“old man of the forest”). In disyllabic stems ending in the combination “voice plosive + vowel i,” the vowel is dropped and the voiced plosive is transformed into the sound [t]: *begi* (“eye”) + *lie* (“hair”) = *betile* (“eyelash”); *gurdi* (“cart”) + *abere* (“animal”) = *gurta-bere* (“draft animal”). With combinations of two plosives at the juncture of stems, the cluster of consonants is simplified to a single sound, becoming deafened and acquiring the place of formation of the second sound: cf. *begi* (“eye”) + *buru* (“head”) = *bepuru* (“eyebrow”). Thus, despite the agglutinative morphology of the Basque language, in which different grammatical meanings are conveyed by different postpositive affixes, sound changes at the juncture of the root and the agglutinative morpheme are, in a number of cases (mainly in finite and auxiliary verbs), of a fusional nature, as a result of which the addition of a postpositive article to the root allows for variable pronunciation and spelling: cf. *zaldi*- + -a = *zaldi(j)a* [saldije], [saldize], [saldifa].

## Morphology

The morphological indicator of the category of determinacy is the definite article -a, which possesses an agglutinative character and is added to the end of a noun, or, more frequently, a noun group (i.e. a phrase of a combination of a word with a preposition). Just like in English, the definite article in Basque harkens back to the demonstrative pronoun. Unlike English, however, where the article can trace its roots to a demonstrative pronoun of near distance, definite articles in Basque go back to a demonstrative pronoun of far distance.

The use of the indefinite article (sing. *bat*, pl. *batzu(k)*) is not as widespread as in English, since its function in Basque is limited. *Bat* is mostly used to mean “one among several or many.” For example, the English sentence “I have a son” can be translated into Basque in two ways, depending on how many sons the speaker has. If he or she has one son, then the indefinite article is not used: *Semea dutldaukat*; if he or she has more than one son, then the definite article must be used: cf. *Seme bat dutl daukat*.



Personal pronouns in Basque are characterized by two distinguishing features: first is the absence of a third-person pronoun in the standard language; and second is the existence of the so-called **familiar** second-person singular pronoun: *ni* (“I”), *hi* (“you” – familiar), *zu* (“you” – neutral), *gu* (“we”), *zuek* (“you all”).

Unlike English, which distinguishes between pronouns of near and far distance, Basque distinguishes between pronouns of **near**, **middle**, and **far** distance: *hau* (combining the forms *hon-*, *haue-* and meaning “this” and “these”) *hori* (combining the forms *horr-*, *horie-* and meaning “that” and “those,” referring to objects near the speaker), and *hura* (combining the forms *har-*, *haie-* and meaning “that” and “those,” referring to objects far away the speaker). The objective proximity/remoteness of the addressee does not play a decisive role in the choice of pronoun; rather, the choice of pronoun is largely determined by the speaker’s attitude to what they are talking about.

Unlike Russian, in which a typical preposition differs as a part of speech from a typical noun both morphologically and semantically (cf. *на* (*na* – “surface”), *в* (*v* – “internal”), *около* (*okolo* – “proximity,” and so on)), postpositions in the Basque language are case forms of a noun, which means that a postpositive preposition can be used in parallel to a noun in a nominative group: cf. *gizonaren atzean* (“behind the man”; *atze* means “backside”); *eliza aldean* (“by the church”; *alde* means “side”); *mendien erdira* (“to the centre of the mountain”; *erdi* means “centre”). Trask calls these “postpositive nouns” [26], although in terms of their grammatical function these elements are closer to prepositions, since they serve to express relationships between words.

These and certain other word-building elements resemble English **semi-affixes** insofar as they can function of independent lexemes and as post-positive elements of words that have a generalized meaning and are partially desemantized. Sometimes, in terms of phonetics, semi-affixes in Basque differ from their meaning-bearing correlates by the presence of a voiceless plosive at the beginning of a syllable: cf. *ume* (“child”) vs. *kume* (“kid”) (*katakume* – “kitten”; *arkume* – “lamb”; *emakume* – “woman”); *alde* (“side”) vs. *kalde* (“located nearby”) (*sukalde* – “kitchen,” lit. “located near the fire”). Trask prefers the term “pseudo-suffix” [25] and includes common word-building elements among them: *-eme* (“woman,” “female”), *-ar* (“man,” “male”), *-bide* (“road”), *-buru* (“head, main”), *-bitxi* (“beautiful”) (cf. *asteme* – female donkey; *katar* – “cat”; *aitabitxi* – “godfather”; *galbide* – “danger”; *mahaiburu* – “officer”).

Along with pseudo-suffixes, the Basque language also includes a group of pseudo-prefixes, which are used both as independent words and as the initial elements of compound words in a meaning that is common to all derivatives. The most frequently used elements belong to the “flora and fauna” semantic-thematic field – *basa-* (“wild”), *otso-* (“lupine”), *sasi-* (“lile wild blackberry”): cf. *basalore* (“wild (field) flower”); *basurde* (“wild boar”); *otsolizar* (“rowanberry”); *sasiarrosa* (“wild rose, rose hip”). Curiously, the last pseudo-affix here has acquired a pejorative connotation in modern Basque, meaning “not real,” “fake,” “counterfeit”: *sasipoeta* (“quill-pusher, sorry excuse for a poet”); *sasiezkontza* (“cohabitation”), from the word *ezkontza* (“marriage”).

When viewing an adjective as a part of speech, the relative suffix *-ko*, whose function is not limited to the formation of what in Russian is commonly called a “relative adjective,” is worthy of special attention. The main function of this suffix is to form a compound definition by adding it to an adverb, phrase, prepositional noun, participle, etc. The scope of the suffix *-ko* includes what in English is conveyed by the genitive, subordinate clause and the prepositional phrases:

- *azto* (“yesterday”) – *aztoko* (“from/of yesterday”);
- *hemen* (“here”) – *hemengo*<sup>3</sup> (“local, of this place”);
- *meridian* (“on a/the mountain”) – *meridianko* (“mountainous”);
- *dirurik gabek* (lit. “without money”) – *dirurik gabeko* (“penniless”);
- *trumoiak adituta* (“having heard the sound of thunder”) – *trumoiak aditutako umeak* (“children who heard the thunder”).

Basque verbs are classified based on the perfect participle, or the so-called participle II. Verbs of the first class end in *-i*: *ikusi*, *ekarri*, *ipini*, etc. Verbs of the second class are formed using participle II with the suffix *-tu*: *hartu*, *lortu*, *aberastu*. In verbs of the third class, the form of the perfect participle is the same as the infinitive: *ito*, *bota*, *hil*. The same is in most cases true for verbs of the fourth class, although the final *-n* can disappear in some forms and contexts.

Unlike English, where the functional-semantic field of the actor category is monocentric – that is, where the absolute majority of names of actors are formed using the productive suffix *-er* – in Basque, the corresponding meanings are conveyed by two productive affixes: *-le* and *-(t)zaile*. The first and oldest is added to verbs with stems ending in *-i* and *-n*: cf. *ikusi* (“to see”) – *ikusle* (“observer”); *erosi* (“to buy”) – *erosle* (“buyer”). The suffix *-(t)zaile* is added to other types of verbs: cf. *hil* (“to kill”) *hiltzaile* (“killer”). Like English, agent affixes in Basque are polysemantic and can form both agent names and instrument names, although they are not productive in this second function. The agent suffixes should be distinguished from the “professional” (i.e. denoting a type of occupation) suffix *-ari*, which marks a representative of a given profession: *merkatari* – “merchant” (from *merkatu* – “market”); *ostalari* – “hotel owner” (from *ostatus* – “hotel”); *musicari* – “musician” (from *musika* – “music”). Unlike the agent suffix, *-ari* is added to the stems of nouns. The suffix *-ari* forms the name of traditional Basque performers known as *bertsolari*, singers who improvise songs on topics suggested by the audience. Competitions are often held among *bertsolari*, with the winner (the *txapeldun*) awarded the traditional Basque beret, the *txapela*. Over time, the definition of the word *txapeldun* has expanded by metonymic transfer to mean “winner,” quite unlike the original meaning of “bearer/owner of a beret.”

**Compounding** and **derivation** are productive word-building methods used in Basque. Most words consist of two or three stems. Among compound words, nouns and verbs predominate. As a rule, the second stem is the main component of a com-

<sup>3</sup> In this case, assimilation takes place.



pound word. There are rare cases where a compound noun is formed by combining two verb stems: *janedan* – “food and drink, the pleasures of the table” (from *jan* – “to eat” + *edan* – “to drink”); *joanetorri* – “a round trip” (from *joan* – “to go” + *etorri* (“to arrive”).

Something that is extremely important in the Basque language is **sound symbolism**, which is realized through the addition of stems that are fully or partially reduplicated. A distinctive feature of this group of words is in the marked phonotactics: atypical sound combinations and uncharacteristic sound positions are widely used. A large group of words is represented by reduplicated adverbs denoting the nature of a given action, specifically a repetitive, unnecessary, excessive or extravagant action: *mara-mara* (“constantly, without interruption”); *zirti-zarta* (decisively); *triki-traka* (“to walk unsteadily”); *pirri-parra* (“abundantly, extravagantly”); *ikusi-makusi* (“hide-and-seek”). The initial *m-*, which is not typical for Basque, is found in sound-symbolic words denoting the presence of a physical or moral defect: *matzar* (“crooked,” “bent”); *matxar* (“deformed”); *mazkaro* (“dirty”); *moxkor* (“drunk”); *mixkiri* (“envious”); *mukur* (“clumsy”); *mutxin* (“angry”); *malmutx* (“fat,” “cunning”). In this case, Basque exhibits what is known as **diagrammatic iconicity**, whereby a marked exponent is used to express marked content.

Unlike in Russian and English, compound verbs are used widely in Basque, and their spellings vary depending on when they first appeared in the language. The earliest verbs are written together, and many of them have been de-etymologized; verbs that appeared later are written separately and thus preserve their transparent morphological composition: cf. *eskaini* – “to offer” (*esku* – “hand” + *ipini* “to put”); *iguriki* – “to wait” (*egun* – “day” + *eduki* – “to hold”). Compound verbs are formed by combining an adjective or noun with verbs that have multiple meanings – *izan* (“to be”), *edun*, *egin* (“to do”), *eman* (“to give”). For example: *bizi* (“alive”) – *bizi izan* (“to live”); *behar* (“necessity”) – *behar edun* (“to need”); *nigar egin* (“to cry”); *hegaz egin* (“to fly”); *musu eman* (“to kiss”); *sue eman* (“to ignite”). This model can be compared with syntactic word composition in English, the difference being that in English nouns are formed from verbs in a similar manner, while in Basque verbs are formed through nouns. One extremely rare example is where a noun is formed from two clauses: cf. *eztabaida* – “argument” (*Ez da! Bai da!* – “No, that’s wrong! Yes, that’s right!”).

## Vocabulary

Most new words in Basque come from the **Romance** languages: primarily Latin in ancient times and the Middle Ages; and now Spanish and French, as well as English (a Germanic language). Nouns from Latin are generally borrowed into Basque in the accusative form, and the same applies to adjectives borrowed in the masculine and neuter. The Basque coordinating conjunction *eta* is generally believed to come from the Latin *et*. The origin of the conjunctions *edo* and *ala* (“or”) remains unclear. The original Basque elements (vocabulary and affixes) are represented by the following groups:

- numerals up to 100
- most pronouns and link words
- inflectional and some derivational affixes
- kinship terms
- the names of some geographical objects
- the names of local flora and fauna
- the names of many instruments (“wheel,” “knife,” “pickaxe,” etc.)
- the adjectives “big,” “red,” “black,” “hot,” “cold”; the verbs “to be,” “to do,” “to have,” “to sleep,” “to take,” “to carry,” “to go,” “to know,” etc.

Borrowings from Spanish are adapted in Basque in the same way as other lexemes. For example, the Spanish participle ending -do becomes -tu in Basque, and the Spanish -ón is replaced by -oi. According to the rules of Basque phonotactics, the sound [d] cannot be in the final position of a word, which explains why the Spanish suffixes -tad and -dad are transformed into -tade and -dade, respectively, in the neutral and colloquial register, and into -tate in the elevated style. The last suffix remains faithful to the Latin original: cf. *fakultate* (literary) vs. *fakultade* (colloquial); *unibersitate* (literary) vs. *unibersidade* (colloquial).

Despite the fact that contacts between speakers of Basque and speakers of Celtic and other Indo-European languages developed for centuries before the arrival of the Romans in Spain and Gaul, the number of Indo-European loanwords in Basque is small; in fact, not a single such word can be identified with a high degree of probability. The most likely candidates are: *adar* (“branch”), *andere* (“Ms.”), *aran* (“plum”), *argi* (“light, light-coloured”), *gori* (“fiery”), *hartz* (“bear”), *landa* (“plain”), *maite* (“beloved”), *mando* (“mule”), *mendi* (“mountain”). Among the most notable borrowings from German are *ain hund*, rendered *ehun* (one hundred) in Basque, with the restored pronunciation [enhund]. The Muslim conquest of Spain led to some Arabic words being assimilated into Basque with certain transformations. Similar to Arabic borrowings in other languages, Arabisms in Basque are recognizable by the initial al- and its variants, which correspond to the definite article in Arabic, and by the harmony of vowels in single words: *alkandora* (“shirt”), *azafrai* (“saffron”), *gutun* (“letter; book”) (cf. Arab. *kutub*).

The etymology of colour adjectives remains unclear, although many researchers agree that most of them originated from the corresponding names of objects in the natural world whose colours to a greater or lesser degree resemble those of the attribute. Resurrección María de Azkue [9] traced the adjective *zuri* (“white”) to the noun *zur* (“tree”); the adjective *hori* (“yellow”) to the noun *hor* (“dog”); and the adjective *gorri* (“red”) to the restored \**gor* (“flesh”). He also noted that the adjective *urdin* (“blue”) consists of the stems -ur (“water”) and -din (“reminiscent of”) and carries the general meaning of “similar to water, the colour of water”). However, we can see that

the semantic bases of some derivatives remain unclear. The word for green (*berde*) is an obvious borrowing from the Latin *perde*, which correlates to the Spanish *verde*, French *vert*, and Romanian *verde*.

The “kinship terms” thematic group is represented by approximately the same number of names as in other Western European languages, with the exception of some specific features. First, Basque uses different lexemes for the words “sister” and “brother” depending on whether the person to whom they are related is male or female: cf. *ahizpa* (“woman’s sister”), *arreba* (man’s sister), *neba* (woman’s brother), *anaia* (man’s brother). Second, a single lexeme is used in Basque to denote what in English is designated by the words “nephew,” “niece,” “grandson” and “granddaughter” – *iloba*, with dialectal variants *lloba*, *loba* and *lioba*. A common denominator between Basque and English is that they both have a lexeme that denotes a brother or sister but does not specify gender: cf. Eng. *sibling*, Basque *haurride*. When addressing one’s mother or father, the compound honorifics *aitajaun* (“father” + “lord”) and *amandere* (“mother” + “lady”) are used in some places. To indicate a more distant degree of kinship, the suffix -so is often used: cf. *ilobaso* (“grandniece”); *arbaso* (“great-grandfather” or “ancestor”). Godparents and godchildren are referred to by means of the semi-affix -bitxi, which is also used as an independent lexeme meaning “attractive, nice”: *alababitxi* (“goddaughter”) *aitabitxi* (“godfather”).

An odd feature of the numeral system in the Basque language is that the numeral 11 (*hamaika*) is used to mean an indefinitely large quantity, similar to “a thousand times” in English. The language also features a peculiarity in the designation of time periods, namely, the greater lexicalization of certain concepts such as “in two days,” “in four days,” “two days ago”: *etzidamu*, *etzikaramu*, *laurdenegun*.

The geographical position of the Basque Country has evidently influenced its **absolute system of orientation**, in which exogenous landmarks are represented by various natural objects – mountains, bodies of water, etc. Exogenous reference points are contrasted with endogenous reference points, for example, the human body and its parts (cf. “to the right/left of me”). The English verb *to go*, in combination with the postpositive particles *in*, *out*, *up* and *down* corresponds to separate Basque verbs: *sartu* (to enter), *jalgi* (“to exit”), *igan* (“to go up”), *jaitsi* (“to go down”). In this respect, Basque differs from both English and Russian, languages that have so-called verbal satellites. The term “verbal satellite” has gained traction in foreign linguistics and used to describe the expression of the direction of movement by an element that is not part of the verb stem and is expressed either by a separate particle, as is the case in English (cf. “to go out,” “to drink up”), or by a prefix or suffix, as is the case in Russian (cf. *выход* (*vykhod* – “exit”) *допить* (*dopit* – “to drink up”)). Languages with verbal satellites are distinct from languages with “frame” verbs, which are named as such because the direction of movement is expressed in the verb stem. The first group includes Germanic, Slavic and some other languages, while the second includes mostly Romance languages.

As is the case with most languages, Basque is subject to the phenomenon of **false etymology**, especially in the field of toponymy. The name of the popular resort of Biarritz (pronounced “Miarritz” by the Basques) is traditionally derived from two roots – *bi* (“two”) and *harri* (“stone”); the element -tz is believed to be a variant of the collective suffix -tza. However, this etymology may be subject to scrutiny, given that in the 12th century the name of the city was written without the plosive t: cf. Bearriz, Beiarritz (Trask 1996). This means that the name Biarritz most likely goes back to the proper name \*Beiar, the attested form of which (Biar) is found in written artefacts of the 15th century; the element -iz is a patronymic suffix added to nouns. This etymology is indirectly confirmed by the presence of a large number of toponyms in a similar way. Another example of folk etymology is the toponym *Ondarribia*, which is formed from the Basque roots *ondar* (“sand”) and *ibi* (“ford”). In this case, the original Basque name was reinterpreted by the Spanish as *Fuenterrabía* (from the Latin *fonte*, meaning “fast stream”), which is a more common name than the Basque version.

Anthroponyms in Basque are formed according to strict rules. Most of them are compound names consisting of a noun stem and an adjective, resulting in a compound noun that is not actually used as a common noun (cf. Russ. *Rodnaya*, *Yamskaya*, *Belaya*; Eng. *Smith*, *White*). Basque surnames can be used with or without the definite article, although the speaker is not free to choose which variant to use, as it is a matter of tradition. While Russian and English have the surnames *Ivanov* and *Smith*, respectively, Basque has *Etxeberria* (lit. “new house”). A small number of surnames are represented by adjectives: *Gorria* (“red”), *Zabala* (“wide”). Semantically, most Basque surnames refer to the geographical location or place of origin of the ancestors of the bearer of the surname. It is not common for Basque surnames to reflect a person’s profession or physical characteristics (such as size, colour, etc.). *García* is a widespread surname in the region, and it originates in the proper name *Gartzia*, which started to become popular in the Middle Ages. The etymology of this anthroponym is a point of contention among Basque scholars: most agree that it is a native Basque anthroponym borrowed by the Spanish and Portuguese; however, some believe that the name comes from the Basque *hartz* (“bear”), which is of Indo-European origin. Semantically, personal names show greater diversity, including such semantic fields as animal names (*Usoa* – “dove,” *Ainara* – “swallow”), colours (*Urdina* – “blue”), religious holidays (*Gabon* – “Christmas”), kinship terms (*Umea* – “child,” *Emazteona* – “good wife”). In the Middle Ages, the method of forming proper names using the masculine diminutive affixes -co and -tx became widespread: *Enneco*, *Bellaco*, *Vitaco*, *Txana* (Ana), *Txelena* (Elena), *Txatalin* (Katalin), *Txekolas* (Nicolás). In the 20th century, there was a surge in popularity of native Basque names following the end of Franco’s dictatorship (1975); meanwhile, French Basques were only allowed to give their children native Basque names in 1993.

The almost-forgotten practice of naming houses is still common in the Basque Country and adjacent areas of France and Spain. Interestingly, the name of the house can be used as an unofficial postal address alongside the official address, which has to include the street name and house number. Typical informal names refer to local weather patterns (Aixea – “wind,” Aidesokua – “windy corner”), a location in a city or village (Arregi – “mountainous area,” Arzubialde – “near the stone bridge,” Elizalde – “near the church,” Ugalde – “near the water”), or the nickname of the house owner (Patxi Errege – “King Francis”).

## Syntax

The Basque language is characterized by a relatively stable word order, with most sentences being of the SOV type. Predicates almost never come at the start of a sentence. Like English, Basque uses an “empty” verb for emphasis – *egin*, which functionally corresponds to the English verb *to do*. Unlike English, the Basque verb is in postposition to the semantic verb: cf. Jonek jo *egin* zuen Kepa.

Like some Romance languages (Spanish, Portuguese and Romanian), Basque does not require subject pronouns to be explicitly expressed, since the form of the verb indicates the subject of the sentence. In this sense, Basque is characterized by “zero subjectness.”

As a rule, the adjective stands in postposition to the noun (cf. *gizon handi hauek* – “those big men”), with the following exceptions: adjectives formed using the suffix -*dun* in the meaning of “having something,” “possessing a person” (cf. *haurdun* – “having a child,” “with a child”); adjectives formed by means of the ethnonymic suffix -(*t*)*ar* (cf. *donostiar ikasleak* – “students from St. Sebastian”); relative adjectives formed using the relative suffix -*ko* (cf. *gaurko egunkaria* – “today’s newspaper”).

Demonstrative pronouns are typically placed after the noun stem, although in Biscayan there are pleonastic formations in which the demonstrative pronoun forms a frame that simultaneously precedes and follows the noun: cf. *au gizon au, ori gizon ori, a gizon a*. Trask (Trask 1996) draws attention to the fact that neither adjectives nor adverbs receive affixal endings outside of a phrase, the typical structure of which has the form complex definition + determiner 1 + noun + adj. + determiner 2 + number + case.

The only thing that indicates the number of participants is a noun combined with the definite article: cf. *gizona* (“this person”) vs. *gizonak* (“these people”), *gizon hau* (“this person”) vs. *gizon hauek* (“these people”). Since indefinite determiners are indifferent to number, the question *Zein gizon?* is ambivalent: cf. “What person / which people? Given the postposition of the adjective to the noun, the order of several adjectives in Basque is, on the one hand, directly opposite to the order of English adjectives. On the other hand, if we take the distance from the noun into account, then the order is, curiously, identical to that of English: cf. Eng. *a pretty little white house* vs. *exte zuri txiki polit bat* (lit. “house white little pretty a”).

Most transitive verbs in Basque can be used in the mediopassive (reflexive-middle) voice, which functionally corresponds to the English passive without an explicitly expressed agent. Unlike the passive voice, which requires the expression of the doer, the expression of the agent noun is impossible in the case of the mediopassive.

The main verbal copulas are *izan* (“to be”) and *egon* (“to wait, to remain”). The former predominates in the northern dialects, while the latter is more common in the southern dialects. It should be noted that verbal copulas are not always expressed in the surface structure of a sentence. Thus, the use of the copula *izan* in northern dialects is permissible under the following conditions: with an animate subject and an object in the locative or comitative cases (cf. *Aita etxean da* vs. *Aita etxean dago* – “Father in the house”). Verbs that carry a meaning of possession are represented by the defective paradigm of the stem *-du-*: cf. *Boitura berria dut* (“I’ve got a new car”). Another verb that functions as a verb of possession has the lexical meaning of “to hold” or “to contain” – that is, to have something enclosed inside. For this reason, its use is limited to situations in which one object is located inside another object, or in close proximity to the speaker: the Basque sentence *Liburua daukat* means “I’m holding a book in my hand” (i.e. “I’ve got a book”).

Based on all the above, we can outline the main features of the typological passport of the Basque language (Table 1).

**Table 1. Main Typological Features of Basque at Four Linguistic Levels**

Agglutinative-suffixal language with agreement and the presence of fusion at the juncture of morphemes	
<b>Phonology</b>	<ul style="list-style-type: none"> <li>• Presence of voiced and voiceless aspirated consonants</li> <li>• Opposition of voicelessness and sonority</li> <li>• Assimilation of nasals by place of formation of the following consonant sound</li> <li>• Prosthesis in foreign words</li> <li>• Variability of pronunciation of the diaphone j</li> <li>• Dynamic quantitative stress</li> <li>• Use of atypical sound combinations for emphasis</li> <li>• Synharmonism of hissing-whistling sounds</li> <li>• Vocalic initial</li> <li>• Palatalization as a means of expression</li> <li>• Absence of initial p-, t-, d-, k-, r- in native Basque words</li> <li>• Fusion alternations</li> </ul>
<b>Morphology</b>	<ul style="list-style-type: none"> <li>• Postpositive definite and indefinite article</li> <li>• Developed case system</li> <li>• Second-person singular familiar pronoun</li> <li>• Pronouns of near, medium and far distance</li> <li>• Postpositional prepositions (postpositions)</li> <li>• <b>Unmarked models:</b> compounding, conversion, derivation, substantivation</li> <li>• <b>Emphatic models:</b> <ul style="list-style-type: none"> <li>• reduplication</li> <li>• echoism</li> </ul> </li> <li>• <b>Absence:</b> blending, back-formation, shortening</li> </ul>



<b>Vocabulary</b>	<ul style="list-style-type: none"> <li>• Romance superstrate</li> <li>• Insignificant amount of Indo-European words</li> <li>• Presence of honorific forms for addressing close relatives</li> <li>• Absolute orientation system and, consequently, the presence of certain lexical gaps</li> <li>• Frame nature of verbal semantics</li> <li>• Semantics of place or origin in surnames</li> <li>• Lexicalization of concepts related to periods of time (two and four days ahead or in the past)</li> </ul>
<b>Syntax</b>	<ul style="list-style-type: none"> <li>• SOV</li> <li>• Pronominal null subject</li> <li>• Postposition of definition</li> <li>• Mediopassive voice</li> <li>• Preposition of subordinate clause in relation to the main clause</li> <li>• Variants of forming interrogative sentences</li> </ul>

## Discussion of Results

In typological terms, the Basque language contains a number of grammatical and lexical phenomena whose status remains debatable (Ivanov 2009). This concerns, first and foremost, the plural form of the postpositive indefinite article *batzuk*, which is traditionally translated using the English quantitative pronouns *a few or several*. Is this form a plural article or an indefinite pronoun? If we compare this form with the forms of the indefinite article in other language, for example, in Bulgarian and Romanian, we discover typological parallels: the indefinite articles in these languages also have a plural form (cf. Bulgarian *едни*, Romanian *niște*). However, they are not mandatory, and the emphasis here is strongly on the noun. In addition, these languages have homonymous forms of indefinite pronouns, which makes distinguishing between the plural form of the indefinite article and indefinite pronouns particularly tricky.

Another polemical aspect of Basque morphology is the status of postpositional prepositions, which are used in parallel as forms of nouns. The functional ambivalence of words like *atze* (“backside”), *alde* (“side”), and *erdi* (“centre”) is only removed in the context of their use, which leaves the question of the grammatical status of these words and the words governing their use open.

In terms of vocabulary, the issue of the conceptual significance of the numerals 2, 4, and 11 and the lexicalization of the corresponding concepts remains unresolved.

Determining the typological significance of the number category and the “singularity-plurality” morphological opposition is also crucial for the typological passport of the Basque language, since the form of the noun does not change when there is a numeral to express the meaning of “more than one.”

An interesting finding of the study was the similar order of noun attributions in Basque and English. Is this word order a universal linguistic feature or simply a coincidence? What cognitive mechanisms and objective characteristics of objects and phenomena are responsible for the arrangement of attributes next to a noun?

On the whole, our results do not contradict the data obtained by Basque scholars in the second half of the 20th century. However, the less systematic and at times contradictory nature of previous studies should be noted. In particular, no one has attempted to distinguish between neutral and emphatic (marked) word-building models, the issue of fusional alternations (which brings the Basque language to inflectional languages) has not been addressed, the framework nature of verbal semantics has not been studied, and a comparative analysis of the typology of the Basque language with English and Russian has never been carried out.

What the present study lacks is empirical data that could be obtained from native Basque speakers. But selecting participants for such a study is problematic: What age and level of education are optimal for conducting a survey? What role does the attitude of the Basque people towards their language and culture play? Speakers of which dialect(s) should take part in such a survey? With this in mind, a multifactorial analysis that takes data obtained directly from native Basque speakers into account appears promising for studying the typology of the Basque language. So does a statistical analysis of the proportion of various typological features, as well as a determination of the typological direction in which modern Basque is developing.

## Conclusion

Despite the attempts of some public organizations to promote the revival of the Basque language, Basque speakers make up approximately 20% of the 3 million population of the Basque Country, part of which are Spanish immigrants living mainly in the industrialized provinces of Biscay and Gipuzkoa. Given the centrifugal nature of the development of minority languages, we can assume that the number of native Basque speakers will, unfortunately, continue to decrease, despite the fact that the Basque language was given official status alongside Spanish in four Spanish provinces following the creation of the Basque Autonomous Community – Gipuzkoa, Biscay, Navarre and Álava.

There is a well-known Basque proverb that translates into English as “As long as we are remembered, someone will need us” (“Gogoratzten gaituzten bitartean, norbaitek behar gaituen bitartean”). We hope that the unveiling of the typological passport of the Basque language will speed up the process of its standardization, raise its status on the linguistic map of the world, and help preserve the unique culture of the region, one of the manifestations of which is this mysterious isolated language.

**About the Author:**

**Natalya A. Lavrova** – Doctor of Philology, Associate Professor in Department of English No. 3 at MGIMO University. Research interests: comparative phraseology, cognitive linguistics, areal and contrastive typology, English lexicology. E-mail: lavruscha@gmail.com  
ORCID: <https://orcid.org/0000-0002-6403-781X>

**Conflict of interests.**

The authors declare no conflicts of interest.

**References:**

- Amorrortu E. 1969. *Basque sociolinguistics: language, society and culture*. Center for Basque Studies, University of Nevada.
- Apodaka E., Basurto A., Galarraga A., Morales-i-Gras J. 2022. The governance of Basque language revitalisation in the Basque Autonomous Community: From confrontation toward collaboration. *Journal of Multilingual and Multicultural Development*. P. 1–14.
- Arakin V. D. 1989. *Tipologiya iazykov i problema metodicheskogo prognozirovaniia [Language typology and the problem of methodological prognosis]*. Vysshaia shkola. (In Russian)
- Augustyniak A. 2021. Migrant Learners of Basque as New Speakers: Language Authenticity and Belonging. *Languages*. 6(3). P. 1–19.
- Azkue R. M. 1935. Evolución de la lengua vasca. *Euskera*. 16. P. 57–133.
- Bengtson J. D. 1991. Macro-Caucasian: A historical linguistic hypothesis. *Dene-Sino-Caucasian languages*. P. 162–170.
- Bengtson J. D. 2011. The Basque Language: History and Origin. *International Journal of Modern Anthropology*. 1(4). P. 43–59.
- Bonaparte L. L. 2018. *Le verbe basque en tableaux*. Hachette Livre BnF.
- Chesnokova O.S., Dzhishkariani L.M. 2017. Baskskii iazyk i baskskaia toponimiia: mezhkul'turnye kontakty i standartizatsiia [Basque language and Basque toponymy: Intercultural contacts and standardisation]. *Vestnik Rossiiskogo universiteta družby narodov. Seriya: Teoriia iazyka. Semiotika. Semantika [RUDN Journal of Language Studies, Semiotics and Semantics]*. 8(4). P. 1074–1083. (In Russian)
- Cots J. M., Martin-Rubió X. 2008. Linguistic Identities in a Basque-Speaking School: Teachers' and Students' Discourses of Multilingualism. *Innovation in Language Learning and Teaching*. 2(3). P. 224–237.
- Dronova S. Y. 2020. Baskskii iazyk v ispanskom gazetnom diskurse kak instrument podderzhaniia politicheskoi identichnosti [Basque in the Spanish newspaper discourse as an instrument for maintaining political identity]. *Gumanitarnye Nauki. Vestnik Finasovogo Universiteta [Humanities and Social Sciences. Bulletin of the Financial University]*. P. 118–124. (In Russian)
- Gorter D., Cenoz J. 2011. Multilingual education for European minority languages: The Basque Country and Friesland. *International Review of Education*. 57. P. 651–666.
- Haddican B. 2007. Suburbanization and language change in Basque. *Language in Society*. 36(5). P. 677–706.
- Hualde J. I., Zuazo K. 2007. The standardization of the Basque language. *Language Problems & Language Planning*. 31(2). P. 142–168.
- Humboldt V. von 1959. *Primitivos pobladores de España y lengua vasca*. Translated from the German by Francisco Echebarria.
- Ivanov V. V. 2009. *Izbrannye trudy po semiotike i istorii kul'tury. Tom V. Mifologiya i fol'klor [Selected works on the semiotics and history of culture. Volume V. Mythology and folklore]*. Znaki. (In Russian)

- Laka I., Santesteban M., Erdocia K., Zawiszeński A. 2012. The Basque language in the minds of native and non-native bilinguals. In *The Challenge of a Bilingual Society in the Basque Country*. University of Nevada. P. 157–172.
- Lavrova N. A. 2019. K voprosu o fonologicheskikh osobennostiakh baskskogo iazyka [Towards the phonological system of modern Basque]. *Topical Issues in English Linguistics and Linguodidactics: Traditions and Innovations*. P. 33–40. (In Russian)
- Michelena L. 1977. Notas sobre compuestos verbales vascos. *Revista de Dialectología y Tradiciones Populares XXXIII*. P. 245–271.
- Ostrowski M. History of the Basque Language. *Basque Linguistics. Basque History*. Academia.edu. (accessed: 01.10.2022)
- Ray N. M., Lete N. 2010. Basque studies: commerce, Heritage, and a language less commonly taught, but whole-heartedly celebrated. *Global Business Languages*. 12. P. 117–131.
- Saussure F. de 1959. *Course in General Linguistics*. Translated from the French by W. Baskin. Philosophical Library.
- Schuchardt H. 1925. *Das Baskisch und die Sprachwissenschaft*. Hölder-Pichler-Tempsky.
- Trask R. L. 1996. *The history of Basque*. Routledge.
- Uhlenbeck C. C. 1947. La langue basque et la linguistique générale. *Lingua*. 1. P. 59–67.
- Urrutia I., Irujo X. 2008. The Basque Language in the Basque Autonomous Community (BAC). In *The Legal Status of the Basque Language Today: One Language, Three Administrations, Seven Different Geographies and a Diaspora*. P. 165–195.
- Valadez C., Etxeberria F., Intxausti N. 2014. Language revitalization and the normalization of Basque: A study of teacher perceptions and expectations in the Basque Country. *Current Issues in Language Planning*. 16(1–2). P. 60–79.
- Vennemann T. 1994. Linguistic reconstruction in the context of European prehistory. *Transactions of the Philological Society*. 92(2). P. 215–184.

## Appendix

Some textbooks recommended by the Royal Academy of the Basque Language for learning Basque

1. Hualde J. I., de Urbina J. O. 1973. *A Grammar of Basque*. Walter de Gruyter.
2. Patrick J. D., Ibarrondo I. Z. 2001. *A Student Grammar of Euskara*. LINCOM Europa.
3. Peillen T. 1995. *Parlons euskara: La langue des Basques*. L'Harmattan.
4. Maia J., Kepa L., Muxika L. 2008. *Euskara batua eta euskalkiak haur eta lehen hezkuntzan*. Universidad del País Vasco.
5. Rudolf P. G. de Rijk. 2007. *Standard Basque: A Progressive Grammar*. MIT Press.
6. Enparantza J. L. A. 2019. *Txillardegi Euskara batua*. Jakin.
7. Urla J. 2012. *Reclaiming Basque: Language, Nation and Cultural Activism*. University of Nevada Press.
8. Sarasua P. U. 2013. *Gramática de la Lengua Vasca*. Editorial UNED.
9. Jansen W. 2007. *Beginner's Basque*. Hippocrene Books.