

Quantifying aspiration in Zuberoan Basque¹

 Mikhael Hayes

University of Illinois at Urbana-Champaign

 José Ignacio Hualde*

University of Illinois at Urbana-Champaign

ABSTRACT: In modern Basque, aspiration is a feature of Northern dialects and reaches its highest frequency in Zuberoan. Here we undertake a detailed analysis of its distribution in J. Larrasquet's (1939) dictionary of north-eastern Zuberoan. Lafon (1948, 1958) and Mitxelena (1951, 1958) noticed two constraints on aspiration: there can be only one aspirated segment per word, and it must be located on the onset of either the first or the second syllable. Here we examine all apparent exceptions to these constraints and their etymological sources. In addition, we study several other topics related to aspiration, including the relative numeric importance of phenomena such as the lenition of intervocalic /n/, the insertion of glottalization at morpheme boundaries, and the development of aspiration in stress-accented syllables. We also consider the extent to which old patterns of distribution of aspirated and unaspirated voiceless stops have been altered or obscured by lexical change and analogy.

KEYWORDS: Aspiration; constraints on aspiration; Zuberoan Basque; diachronic phonology.

¹ We are grateful to *ASJU* Editorial Board Member Gorka Elordieta and to two anonymous reviewers for comments and suggestions that have helped us improve this paper. We also want to thank the members of the audience at the First International Conference on the History of Basque (Vitoria-Gasteiz, 25-27 November 2025), where a first, oral version of this paper was presented. The corresponding author would also like to acknowledge his participation in the project *Monumenta Linguae Vasconum* 6 (PID2020-118445GB-I00) supported by Spain's Ministerio de Ciencia e Innovación.

* José Ignacio Hualde. Dept. of Spanish and Portuguese & Dept. of Linguistics - University of Illinois at Urbana-Champaign. 4080 FLB, Urbana, IL 61801, USA - jihualde@illinois.edu -  <https://orcid.org/0000-0003-4969-2995>



1. Introduction: Aspiration and its sources in Basque

Since the 16th century, when the first long texts in Basque that have arrived to us were written, a striking dialectal difference has to do with aspiration. In Northern texts, we find a high profusion of /h/ and aspirated voiceless plosives orthographically differentiated from unaspirated ones. In Southern texts, however, there is a total lack of phonological aspiration. Since a phoneme /h/ (whatever its exact articulation, perhaps originally a postvelar fricative [χ]; see Manterola & Hualde 2021; Gonzalez-Eppherre & Oihartzabal 2025) is also represented in medieval Southern toponyms and in the Aquitanian-Vasconic onomastics from Roman times, it is clear that aspiration is a conservative feature in Northern Basque that was lost south of the Pyrenees. However, not all details of the distribution of aspiration in Northern Basque are necessarily reconstructible for Common Basque. Instead, aspiration has also been innovated in Northern Basque in certain contexts. The main sources of aspiration are the following:

- 1) First, as mentioned, aspiration may be etymological, in the sense that it is a phoneme reconstructible for an early common stage. This aspiration may appear word-initially, as in *harri* ‘stone’, intervocally (including after glide), as in *behi* ‘cow’, *oihan* ‘forest’, and after a non-labial sonorant, as in *elhorri* ‘thorn’. This does not mean, however, that all instances of /h/ in these contexts in Northern Basque should be reconstructed for Common Basque; see below.
- 2) Secondly, a nasalized aspiration /*h̃*/ is found as a result of the weakening of an intervocalic alveolar nasal consonant, as in ANATE > *añate* [ãñâte] ‘duck’. A nasalized /*h̃*/ was first proposed as a distinct phoneme in Hualde (1993a), and is included in the phonemic inventory of Zuberoan Basque in Hualde (2003: 18, 25); see also Egurtzegi (2018), whereas Mitxelena (1977) and other sources (such as Igartua 2008, 2015) refer to aspiration flanked by nasalized vowels instead. The weakening of intervocalic /n/ took place prior to dialectal fragmentation, at the Common Basque stage.
- 3) Thirdly, we may find non-etymological aspiration at morpheme boundaries in compounds, as in *asteharte* ‘Tuesday’ from *aste* ‘week’ plus *arte* ‘middle’ (see Mitxelena 1977: 210-211, 216). This aspiration may have developed from reinterpretation of glottalization in this context (Hualde 2007: 298, 2018: 9), e.g.: *aste+arte* > **asteʔarte* > *asteharte* ‘Tuesday’. Whereas glottalization between the members of a compound may be assumed for Common

Basque, to account for changes common to all dialects, the subsequent change [ʔ] > [h] may have been confined to the Northern area (e.g., *begi+azal* > **begʔazal* > *betazal* ~ *bethazal* ‘eyelid’). (We note that this is not a universally accepted historical process and that other explanations may be at play; see Lakarra 2009, 2011).

- 4) Finally, aspiration developed on the onset of the second syllable, likely under stress-accent, as in *solhas* ‘conversation’ < Gasc. *solatz*, *gelhari* ‘chamber maid’ < CELLĀRIU/A. In this position. Word-initially, /h/ has also developed in some cases, whether or not that syllable was stressed, as in *husu* ‘frequently’, *harroka* ‘rock’.

Regarding aspirated plosives, these have likely resulted from a phonemic split in Northern Basque. As proposed by Mitxelena (1951, 1977), in these varieties, an allophonic distribution where the aspirated or unaspirated realization of voiceless stops was predictable from the position of the stress-accent gave rise to a phonemic contrast between /p t k/ vs. /p^h t^h k^h/ after a subsequent shift of the stress to the penultimate syllable of most words.

To assess the distribution in the lexicon of aspiration from different etymological sources in Northern Basque, we have undertaken a study of Larrasquet’s (1939) dictionary of northeastern Zuberoan. We focus on this source because, among modern dialects, Zuberoan is particularly conservative regarding aspiration. In fact, Mitxelena (1958, 1977) notices numerous correspondences between Larrasquet and the 16th c. Bible translator J. Leizarraga and uses primarily these two sources to establish generalizations about the distribution of aspiration. To some extent, the study that we undertake in this paper complements that of Hualde (2018), which primarily focuses on aspiration in P. Axular’s work *Gero* (1643), although here we go beyond that article in several respects, as Larrasquet’s dictionary allows further exploration of generalizations on aspiration.

When we analyze aspiration in Larrasquet’s dictionary, the first remarkable fact is that aspiration has an extraordinarily high frequency. The dictionary contains 6025 entries by our count. After discarding affixes, duplicates, finite verbal forms, interjections, etc., we obtain a total of 5775 valid entries. Of these, 1886 contain some form of aspiration, which is almost a third of all words in the dictionary (32.7%). In the remainder of the article, we consider the distribution of aspiration in Larrasquet’s dictionary.

It has long been noticed that the distribution of aspiration, including both aspirated stops and the laryngeal fricatives /h/, /h̃/, is subject to two constraints: There can be no more than one aspiration per word, and there is no aspiration beyond the second syllable (Lafon 1948, 1958; Mitxelena 1951, 1958, 1977). In section 2, we examine possible exceptions to each of these two constraints in Larrasquet’s dictionary, and we assess the relative productivity of the alternations to which they give rise. In section 3, we consider the impact of the weakening of intervocalic /n/ in the overall frequency of aspiration and the extent to which we can determine whether aspiration has this origin in specific words. Word-internally, aspiration may follow a non-labial sonorant; the frequency and distribution of sonorant+/h/ clusters is assessed in section 4. In section 5, we examine the distribution of aspirated and unaspirated voiceless stops. In section 6, we briefly consider the frequency of word-initial aspiration. Finally, section 7 offers a summary of the main findings of this study.

2. Constraints on aspiration

In spite of the very high frequency of aspiration in Larrasquet’s dictionary, as just mentioned, the distribution of aspiration in Northern Basque is subject to two constraints. First, Lafon (1948, 1958 [1999: 121]) pointed out that there cannot be no more than one aspirated segment per word, including both laryngeal fricatives and aspirated plosives. This constraint, which, as Lafon noticed, is reminiscent of Grassmann’s Law in Greek and Sanskrit, gives rise to morphophonological alternations. A second constraint on aspiration is that there can be no aspiration beyond the second syllable of the word (Mitxelena 1951, 1958, 1977).² In the next three subsections, we consider the evidence for the active application of these constraints in suffixed words included in Larrasquet’s dictionary, before undertaking an analysis of exceptions to each of the two constraints in the dictionary.

2.1. Constraints on aspiration and suffix allomorphy

Mitxelena (1951, 1977: 215-216) remarked on the effects that the constraints on the distribution of aspiration have in the shape of the participial suffix *-thu* ~ *-tu* in Northern dialects, including in

² Both generalizations have some exceptions in Southern medieval toponyms, suggesting a more recent diachrony. However, in a very recent paper, Orduña Aznar (2026) suggests that a tendency to obey these two restrictions may already have been found in Roman times, as attested in the Aquitanian onomastic corpus.

the work of J. Leizarraga (1571) and B. Etxepare (1545), as well as in Larrasquet's dictionary. In Larrasquet's dictionary, we find *sárthü* 'to get/put in', *sóρθü* 'to be born', etc., with an aspirated dental, but *khéntü* 'to take away, remove', *hártü* 'to take', with dissimilatory loss of the aspiration (not ***khenthü*, ***harthü*). The constraint against aspiration beyond the second syllable is obeyed in examples like *agértü* 'to appear', where the stem is bisyllabic (In *gréüthu* 'to become disgusted', from Gasc. *greu*, and in *déithü* 'to call', we obviously have a diphthong and /tʰ/ appears on the onset of the second syllable). Both constraints operate in examples like *khausítü* 'to find', *harrítü* 'to surprise' where the stem is bisyllabic and starts with aspiration or an aspirated stop.

With other suffixes, the constraints on aspiration identified by Lafon and Mitxelena are less systematically obeyed in Larrasquet's dictionary. Thus, with *-khoi* 'fond of' (e.g., *berkhói* 'selfish'), there are exceptions to both constraints. The aspiration is found after the second syllable in *arakhói* 'carnivore' and *ardũkhói* 'fond of wine', as noted by Mitxelena (1977: 419), and we find two aspirations in *hunkhói* 'gourmet, fond of fine things'. These are transparent morphological formations that may be quite recent, which accounts for the relaxation of the constraints. On the other hand, we notice the presence in the dictionary of the dissimilated form *aphürkói* 'trachea'. As Larrasquet himself points out, this word is formed on *áphür* 'bit'. Thus, the trachea is literally a place that may attract small bits of food. The shift in meaning has rendered the morphological composition of this word less transparent, favoring the dissimilation of aspiration in this case.

Beyond the cases of suffixation just mentioned, in compounds, we find few cases of dissimilatory deaspiration in the dictionary. Examples include *ilhérri* 'cemetery' < *hil* 'dead' + *herri* 'county' (but *hílhòtz* 'corpse' < *-hotz* 'cold', also in Etxepare; see Mitxelena 1951), and *iñau* ~ *ñiau* 'thou thyself' < *hi+hau*. Interestingly, the non-etymological aspiration arising at the morpheme boundary is the one that is preserved in examples such as *aurhíde* 'sibling' < **haurhide* (< *haur* 'child' + *ide* 'companion'), and *unhétsi* 'to approve, to deem good' (< *hun* 'good' + *etsi* 'to consider'). Alternatively, some of these examples may involve metathesis of the aspiration, rather than dissimilation, e.g. *honetsi* > *unhetsi* (see Lakarra 2009, 2011: 191; *OEH*, s.v. *onetsi*).

The active application of this dissimilation rule in compounds is thus relatively uncommon. Now we turn directly to the question of how frequently we find words that violate one or the other constraint on the distribution of aspiration in Larrasquet's dictionary.

Before continuing, a note on accentuation may be useful. In Zuberoan Basque there is phonologically contrastive stress-accent, which normally falls on the penultimate syllable of the

word, although there are also words with final and exceptionally antepenultimate stress (see Hualde 1993a and references therein). Larrasquet most often (but not always) indicates the position of the stress-accent in phonetic transcription.³ However, it should be noticed that the stress shifts when inflectional suffixes are added; e.g. *éphér* ‘partridge’, *éphérra* ‘the partridge’. Generally, in this paper, henceforth, we only write the stress with an accent mark when it deviates from the penultimate stress rule and also in compounds, which may have either one or two stressed syllables. When the position of the stress is not indicated, it falls on the penultimate syllable.

2.2. Exceptions to the first constraint: Only one aspiration per word

In Larrasquet’s dictionary, we find 38 apparently exceptional lemmata to the ***h...h* constraint (44 if including words derived from these by suffixation). However, all of these are compounds (and most of them have two stressed syllables in Larrasquet’s transcription). We may classify compounds where each member preserves its aspiration in the following types:⁴

(1) Compounds with two aspirations (n=38)

- a) Compound numbers (n=3): *hamabí-ehün* ‘twelve hundred’, *híru-èhün* ‘three hundred’,⁵ *hama-híru* ‘thirteen’.
- b) Deictic adverbials (n=5): *hánti-harát* ‘from that moment on’, *hara-hunáka* ‘comings and goings’, *hèben-harát* ‘from then on’, *hàn-hében* ‘here and there’, *hàn-hétan* ‘approximately’.

³ It appears that, for typographical reasons, Larrasquet was unable to place more than one diacritic on the same vowel, which prevented him from marking stress on *ü* and vowels provided with a nasalization diacritic. In such cases, sometimes he added a note regarding the position of the stress, but not consistently.

⁴ Except when the origin of vowel nasalization is being directly discussed, throughout the paper, we omit predictable nasalization of vowels in contact with a nasal consonant. Thus, we write *hamabi* ‘twelve’ and not *hāmābi*. When the trigger of nasalization is intervocalic /*h̃*/, we indicate nasality on this consonant. Thus, we write *ah̃atz* ‘forget’, *eh̃i* ‘easy’, instead of *āh̃atz*, *ēh̃i*. However, predictable vowel nasalization is indicated in phonetic transcriptions within brackets.

⁵ Intervocalic *-r-* is indicated as silent in Larrasquet’s phonetic transcriptions. This is because of a recent sound change in Zuberoan. In the orthographic forms in the dictionary this consonant is preserved.

- c) Reduplicated forms (n=6): *handi-handia* ‘very big, haughty’, *ther-ther* ‘little by little’,⁶ *khordóka-khordóka* ‘staggering’, *hàla-hàla* ‘in the same way’, *hala-hùla* ‘anyhow’, *hálaz-hálaz* ‘little by little’.
- d) Quasi-derivational suffixes (n=3): *ah̃atz-é̃hi* ‘easy to forget’, *ház-ẽhi* ‘easy to raise’, *hún-khòì* ‘fond of fine things’.⁷
- e) Other:
- N + N > N (n=11): *aho-mí̃hi(z)* ‘orally (mouth+tongue)’, *eihéra-hárri* ‘millstone’, *harri-kháldu* ‘blow with a stone’, *hàxtur-kháldu* ‘snip of the scissors’, *húin-búlhar* ‘instep (foot+breast)’, *hótz-ikhára* ‘trembling from cold (cold+trembling)’, *húxtü-bélhar* ‘wild oats (whistle+grass)’, *mãhats-mólkho* ‘bunch of grapes’, *híl-mih̃ise* ‘shroud (dead+sheet)’, *ilhar-théka* ‘bean pod’, *ilhar-zúrkhats* ‘stakes for climbing beans’.
- N + Adj > N (n=6): *haro-hún* ‘fine weather’, *haur-hùn* ‘good child’, *híl-hòtz* ‘corpse’, *lepho-hándi* ‘goitrous (neck+big)’, *báhe-xéhe* ‘sieve with a fine mesh’, *has-báhi* ‘asthmatic (breath+seized)’.
- Number + N > N (n=1): *hiru-hórtzeko* ‘three-toothed sheep’.
- N/Adj + V > V (n=2): *hàur(a)-ú̃khen* ‘to give birth’, *hun-hártü* ‘to take in good part’.
- Quant + Adv > Adv (n=1): *hambat-hóbe* ‘even better’.

In addition to the examples just given, we find some cases with two aspirations, but where instead of both aspirations being etymological, in the sense that it is present in each of the members of the compound used as a free word, the first instance of aspiration is from glottalization at the compound boundary, and the etymological aspiration on the second member of the compound has also been preserved. Examples include *bethóhil* ‘with an evasive gaze’ < *begi+ohil* and *erthoháñ* ‘almost for free’ < *erdi+dohain*. This pattern contrasts with the type illustrated by *aurhíde* ‘sibling’ < *haur+ide* type, where aspiration is lost in the first member of the compound (for another possible account, involving metathesis; see Lakarra 2009, 2011).

⁶ In some compounds, Larrasquet does not indicate the accented vowel.

⁷ We classify these as compounds containing quasi-derivational suffixes (or quasi-compounds) because *khòì* ‘fond of’ and *ẽhi* ‘easy’ are used primarily as second members of compounds and only rarely as free morphological elements.

To summarize this subsection, the constraint ***h...h* operates without exceptions in monomorphemic environments and with participles in *-thü*; on the other hand, it is only sporadically applied as active dissimilation in compounds.

As a reviewer points out, the domain within which the constraint applies (or historically applied) must be prosodically defined. In this respect, the lack of dissimilation in most compounds follows from the fact that at the relevant point in time they must have contained two separate accentual domains. For many compounds, Larrasquet in fact indicates two accented syllables, although not in all compounds (see Hualde 1993a: 319-320). At the stage of language that Larrasquet describes, this prosodic condition (i.e. only one aspiration per accentual domain) no longer operated systematically, since for instance, even for a completely transparent reduplicative compound like *handi-handia* ‘very big’, Larrasquet explicitly indicates that it is pronounced “avec un seul accent”. Nevertheless, most compounds in the dictionary do indeed contain two accentual domains.

2.3. Exceptions to the second constraint: No aspiration beyond the second syllable

Mitxelena (1951, 1958, 1977) pointed out that, in general, only the onset of the first or the second syllable in the word may bear aspiration, including aspirated stops. As exceptions in Larrasquet’s dictionary requiring a special explanation, Mitxelena (1977: 419) noticed *arrathü* ‘rat’, *baranthálla* ‘February’ and *borontháte* ‘will’, all three of them borrowings, which exceptionally would have preserved the position of the stress-accent in the source language (although for *baranthálla*, Gonzalez-Eppherre & Oihartzabal [2025: 199] propose an alternative, native etymology as a compound, **berandu+hila* ‘late month’). We have found only one more monomorphemic example with aspiration late in the word (i.e. beyond the second syllable), *marrukhá* ‘wart’, ultimately from VERRŪCA, but whose final stress points to an oxytonic Gascon source, cf. Sp. *verrugón*.

In addition to these four exceptions, we find about 90 words with a single aspiration beyond the second syllable, without including in this number the words with two aspirations discussed in the preceding subsection. Although this is not a small number of exceptions, it should be noticed that all of them are transparent compounds and derived words. Among derived words (or quasi-compounds) with late aspiration, we find examples with *-khoi* ‘fond of’, like *ardükhói* ‘fond of wine’ and *arakhói* ‘carnivore’, already mentioned above, and with *-ehi* ‘easy’, other examples such as *entzunéhi* ‘easy to hear’. The quasi-derivational element *-thegi* ~ *-tegi* ‘place’ does not show a

uniform behavior. We find *urdanthégi* ‘pigsty’, *arthégi* ‘sheep fold’ (*ardi-*) with aspiration, but *ürzotégi* ‘dovecote’ (*ürzo-*), *olláutegi* ‘henhouse’ (*ollo-*) (with a different stress pattern), *aroztegi* ‘forge’ (*arotz-*) (without indication of stress).

In (2), we classify words with a single, but late, aspiration by their morphological structure. The majority of these words are N + N compounds, but other types of compounds are also found, including lexicalized phrases and collocations whose status as compounds may be debatable:

(2) Compounds with a single aspiration beyond the second syllable (total n=46)

- a) N + N compounds (n=28): *ardi-bélhar* ‘a variety of grass (lit. sheep+grass)’, *ardi-míñi* ‘valerian (lit. sheep+tongue)’, *aste-hárte* ‘Tuesday’, *asto-húme* ‘foal of a donkey’, *bégi-kháldü* ‘a glance (eye+blow)’, *begi-thárte* ‘face (eye+between)’, *beso-büllhür* ‘a hobble or fetter for a cow (arm+tie)’, *bizi-náñi* ‘resourceful, sharp (life+wish)’, *bota-hárri* ‘flat stone used in the game of pelote (throw+stone)’, *bürdün-hür* ‘ferruginous water (iron+water)’, *èür-bélhar* ‘purple loosestrife (wood+grass)’, *ègürgur-bíhotz* ‘heartwood’, *ègüz-alhába* ‘goddaughter’, *ègüz-aurhíde* ‘spiritual sibling (godbrother/godsister)’, *gai-añhara* ‘bat (night+swallow)’, *gai-hérdi* ‘midnight’, *gorotz-ürpho* ‘a small dung heap’, *gòse-ahamiát* ‘insatiable (hunger+ Gasc. famished)’, *kuntre-phozũ* ‘antidote (counter-poison)’, *odol-bélhar* ‘pellitory (blood+grass)’, *ogi-bíhi* ‘grain of wheat’, *ogi-bíhotz* ‘crumb of bread (bread+heart)’, *organ-hága* ‘a stake for a cart’, *sar-jálkhi* ‘entry and exit’, *sed-ákhüllü* ‘a muzzle for a calf’, *urdan-thégi* ‘pigsty’, *urx-aphal* ‘turtledove (dove+low)’, *zü-bíhotz* ‘heartwood (wood+heart)’.
- b) Inflected N + X (n=2): *bürüti-behéra* ‘on the face (head.ABL+down.ALL)’, *bürüz-phiko* ‘(falling) on one’s head’.
- c) N + Adj: *andére-hándi* ‘woman in formal attire’, *bürü-á(r)hin* ‘thoughtless (head+light)’, *buru-háutsgarri* ‘(noise) that is nerve-wracking’, *jènte-xéñe* ‘common people’.
- d) V + N (n=2): *egin-náñi* ‘desirous of doing well’, *jín-bèhar* ‘destination’.
- e) N + V (n=2): *meréxi-ükhen* ‘to deserve’.
- f) Adv + N (n=2): *núiz-náñi* ‘anytime’, *núla-náñi* ‘anyhow’.
- g) Adv + V (n=1): *untsa-ikhási* ‘polite, well-mannered’.
- h) Numbers (n=5): *berr-éhün* ‘two hundred’, *biz-pa-hírrur* ‘two or three’, *híru-éhün* ‘three hundred’, *láu-ehün* ‘four hundred’, *zázpi-éhün* ‘seven hundred’.

Summarizing the data presented in this and the previous subsection, in compounds (and quasi-compounds, i.e. polymorphemic words whose second element may alternatively be analyzed as a derivational suffix), we find five aspiration patterns, three of them involving insertion of aspiration at the morpheme boundary. The application of dissimilation can be observed in (3b) and (3e):

(3) Patterns of aspiration in compounds

- a) Aspiration in both members of the compound, with preservation of both instances of aspiration (n=45), e.g.:

$h + h > h-h$ *hil+hotz > hilhòtz*

- b) Aspiration in both members of the compound, with dissimilatory loss of the first aspiration (n=2), e.g.:

$h + h > 0-h$ *hil+herri > ilhérrí*

- c) Insertion of aspiration at morpheme boundary (n=7), e.g.:

$0 + 0 > 0-h$ *aste+arte > astehárte, gai+erdi > gaihérdi*

- d) Insertion of aspiration at morpheme boundary and preservation of aspiration on second member (n=2) e.g.:

$0 + h > h-h$ *begi+ohil > bethóhil*

- e) Insertion of aspiration at morpheme boundary and deletion of aspiration in first member (n=4), e.g.:

$h + 0 > 0-h$ *haur+ide > aurrhíde, hun+etsi > unhétsi*

There seem to be no examples in Larrasquet's dictionary where the first member of the compound has aspiration that is preserved even if a second aspiration is introduced at the morpheme boundary; that is, the type ***haurhide* is not found.⁸ Nevertheless, *haurhide* is documented in the work of a few authors according to the *OEH*, which shows that this pattern is not necessarily excluded in all varieties of the language. A second observation is that aspiration in the second member of the compound is always preserved, unlike what we find with suffixes like *-thü*.

⁸ Larrasquet writes this word orthographically as *(h)aurhide* and provides the transcription [aurhíde].

3. Aspiration from the weakening of intervocalic /n/

In Zuberoan /h/ and /h̃/ are in phonemic contrast. Larrasquet's dictionary is the main source that we have for examining the distribution of these two intervocalic phonemes. In the dictionary, we find minimal pairs such as *ehi* 'finger' (< *erhi*) vs. *ēhĩ* /*ehĩ*/ 'easy'. As is well-known, the main historical source of nasalized aspiration is the weakening of intervocalic /n/, a sound change that is easy to ascertain in loanwords such as ANATE > *añate* 'duck', HONORE > *oñore* > Zub. /*ũh̃üe*/ 'honor', etc. In this section our aim is to determine what the impact of this diachronic phenomenon has been in the overall frequency of aspiration in the lexicon. That is, how many words have intervocalic aspiration as the result of the lenition of /n/ as opposed to other sources?

It should be noted that aspiration from this source is only expected on the onset of the second syllable, since elsewhere only vowel nasalization is preserved, as in **ardano* > *ardãõ* > Zub. *ardũ* 'wine' (i.e. not ***ardaño*), which Mitxelena attributed to the location of the stress-accent at the time intervocalic /n/ lenition took place, with aspiration being preserved on the onset of the stressed syllable (leaving aside a few cases of possible metathesis of the aspiration, e.g. *leõne* > **leoñe* > *leñoe* > Zub. [leñũ], Lakarra 2009).

To begin with, in Larrasquet's dictionary we find 51 lemmata (148 words) with nasalized laryngeal fricatives, all in intervocalic position on the onset of the second syllable. In the dictionary, five of these include explicit notes describing /h/-nasality or whole-word nasality, and the remaining majority are indicated by Larrasquet with nasalization diacritics on the flanking vowels, e.g. *âhâlke* 'shame' = *añâlke* [ãñãlke].

When we consider the etymological source of aspiration in words where Larrasquet also indicates vowel nasalization, we need to keep in mind that all nasal consonants trigger nasalization on flanking vowels, not only /h̃/, cf. e.g. *amodio* [ãmõdíõ] 'love', *animal* [ãñĩmãl] 'huge'. This being the case, we first need to consider separately those words where Larrasquet indicates nasalization either on the *h* itself or on one or both of the adjacent vowels, but where the word also has other possible sources of nasalization; e.g. *behin* [béhĩn] 'once'. This is because, in words of this type, the aspiration may have a source other than the weakening of /-n/; that is, the etymon may have been either **benin* or **beñin*. To decide this matter, it may be relevant that only the second vowel of *behin* is nasalized in Larrasquet's transcription (cf. also Egurtzegi [2023: 288] for spectrographic evidence that /h/ is actually non-nasal in this word). Nevertheless, Larrasquet also

transcribes *lehen* [léhēn] ‘before’ with indication of nasality only on the second vowel, adjacent to /n/, but [lěhēnāgo], suggesting in fact /leñen/, and thus **lenen*, possibly a superlative; see *OEH*. (Possible voiced aspiration is not indicated in Larrasquet’s transcriptions and notes).

Larrasquet’s dictionary includes 32 lemmata of the [béhñ] type, containing intervocalic *h* and where nasalization is indicated on one or both of the adjacent vowels, but where the word also has a nasal stop, which could be the source of vowel nasalization. In some of these words we know with certainty that the aspiration has resulted from the weakening of /n/ because they are borrowings or from inter-dialectal correspondences, as in the examples in (4a). In a second set of words, it is also clear from etymological considerations that aspiration has a different origin than the loss of /n/, as in (4b). Finally, for the majority of words with this pattern, we cannot decide this issue easily, (4c). In (4) we indicate vowel nasalization as it appears in Larrasquet’s dictionary, replacing the circumflex accent that he uses to represent nasalization with the usual tilde:

(4) Nasality from /-n-/ > /ñ/ in words with other nasal consonants? (about 20)

- a) YES: *māhñanka* ‘sleeve, handle’ < Lat. MANICA; *mīhñimen* ‘wicker’ < VIMINE (cf. Sp. *mimbre*); *mīhñi* ‘tongue’ < **bini*, (cf. Gip. *mingain*, Old Bizk. *mīñ*); *mūhñuri* ‘grape berry’ (cf. Baztan *minauri*).
- b) NO: *āhāmēñ* ‘mouthful, food’, cf. *aho* ‘mouth’; *āhñ* ‘plum’, cf. *aran*; *āhñ* ‘light weight, dizzy’, cf. *arin*; *Jōhññē* ‘John’; *(gose)āhñmñāt* ‘insatiable’ < Gasc. *ahamiat* ‘famished’ (**FAMINE*).
- c) PERHAPS: *behñ* ‘once’; *lehñ*, *lēhñago* ‘first’; *ēhñ* ‘hundred’; *ēhñ* ‘to weave’, *māhñ* ‘table’; *māhñats* ‘grape’; *mēhñ* ‘thin, slender’; *mīhñse* ‘sheet, shroud’; *nāhñs* ‘to mix, confuse’, *nāhñ* ‘to want’; *ūhñ* ‘thief’ (Lap. *ohoin*).

Some of the examples in (4) require further discussion. The loanword *māhñanka* [māññanka] /maññanka/ ‘sleeve, handle’ may have had a derivation *maneka* > *mañeka* > *mañanka*, with a nasal stop [ŋ] restored before the velar plosive. In VIMINE > /miñimen/ it appears that there was metathesis of the two nasals, followed by the weakening of /n/ to /ñ/: **bimine* > **binime* > **biñime* > *miñime(n)*. For the native word /miñi/ ‘tongue’, the Gipuzkoan form *mingain* (-gain ‘top’) and the presence of nasalized vowels in Old Bizkaian provide strong evidence for **mini*, from which Mitxelena further reconstructs **bini*.

In the examples in (4b), the aspiration has an uncontroversial source other than the weakening of intervocalic /n/. The word for ‘mouthful’ is a compound of *aho* ‘mouth’, which is listed in the dictionary with non-nasalized aspiration, **aho+men*. In the words for ‘plum’ and ‘light (weight)’, consideration of cognates in other dialects shows that aspiration may have arisen after a rhotic on the onset of the second syllable (see section 4), with subsequent loss of the rhotic consonant: *aran* > *arhan* > *ahan*. In the loanword (*gose*)*ahamiat* ‘famished’ aspiration is the result of /f/ > /h/ in Gascon. This leaves the residue of the words in (4c) for which we can reconstruct intervocalic **/n/* with different levels of confidence.

Leaving now words with both aspiration and a nasal stop aside, then, the words transcribed by Larrasquet with /h/ and nasalized vowels include those in (5), which we classify by the etymological origin of the aspiration:

(5) Nasalized aspiration from /-n-/?

- a) Loanwords: *añate* ‘duck’ (< Lat. ANATE), *uñu(r)e* ‘honor’ (< Lat. HONORE), *lehũ* (< Lat. LEONE), *xaña* ‘to wash’ (< Lat. SANARE). In this set of words, we can be completely certain that the nasalized aspiration derives from an earlier intervocalic /-n-/. Nasalization of the aspiration may have been lost in some additional loanwords where it also derives from /-n-/. Thus, as Mitxelena points out, *bahitũ* ‘to seize, pawn’ has a variant *banitu*, and is attested in Navarrese Romance as *banido*, but it has no indication of nasalization in Larrasquet’s dictionary.
- b) Negative pro-forms: **e-no-*: *iñur(-ere)* ‘nobody’ (< **e-no-r*), *iñula* (< **e-nola* ‘no way’), *iñun* ‘nowhere’ (< **e-non*), etc. In this set of pro-forms, /n/ can also be securely established, given the corresponding question words *nor*, *nola*, *non*, etc., and cognates in other dialects.
- c) For most (synchronically) monomorphemic native words, the reconstruction of **/n/* for Zub. /ñ/ is the best hypothesis available; in some cases we find confirmatory evidence in Roncalese or in Old Bizkaian, which is generally mentioned in the corresponding entries in the *OEH* when available; in other cases, we can have less certainty (n=14): *añardi* ‘sow, female pig’, *añari* ‘ram’, *iñi* ‘reed, rush’, *iñitz* ‘dew’, *eñi* ‘easy’, *añalke* ‘shame, timidity’, *añizpa* ‘sister of a woman’, *añolku* ‘advice’, *xeñe* ‘small, fine’ (*zeñatũ* ‘to punish physically’, *zeñe* ‘a unit of measure, span’), *iñes* ‘fleeing’, *iñardets(i)* ‘to reply’, *añatz(i)* ‘to forget’, *iñarraus(i)* ‘to shake’, *añarrausi* ‘yawn’.

- d) Expressive words: *eñe* ‘no, nope!’, *eñelega* ‘to annoy’. In these examples it is not clear that /ñ/ continues an older /n/.
- e) Finally, we also find some cases of obviously non-etymological nasalization: *iñau* ‘yourself’ < *hi+hau*, *giñau* ‘ourselves’ < *gu+hau*, *eñausi* ‘barking’ (cf. *erausi* ‘to talk’ < **e-r(h)a-aus(i)*).

For comparison, we also find 60 lemmata with intervocalic *-h-* and no indication of nasalization, including the items in (6):

(6) Non-nasalized intervocalic /h/ (n=60)

- a) From aspiration after /r/ (VrV > VrhV > VhV): *ehi* ‘finger’ (< *eri*), *eho* ‘to kill, grind’ (< *ero*), *oha* ‘to knead; to suit’ (< *ora*); *soho* ‘meadow’ (< *soro*, Lat. SOLU), *behez-i* ‘to distinguish’ (< *berez-i*), *kehella* ‘wattle gate’ (< *kereta* < Rom. **cleta*), *uharte* ‘island’ (< *ur+arte*), *uhaitz* ‘river’ (< *ur+gaitz*), *zohi* ‘mature’ (cf. *zori*, *zoli*), *zuhoi* ‘tool socket’ (< *zur+ koi*).
- b) From vowel doubling: *lahardeki* ‘to skin’ (< *larrü* ‘skin’+ *edeki* ‘remove’).
- c) Loanwords. With /h/ (< /f/) in Gascon: *buha* ‘to blow’, *gaha* ‘sic (for a dog)’, *gahe* ‘ladle’, *kohat* ‘slap, smack’, *dehota* ‘barren animal’. From /-b-/ > /h/: *kahaka* ‘owl’ (Gasc. *cabaque*), *tahalla* ‘tablecloth’ (Gasc. *tabaille*), *taharna* ‘tavern’, *ahul* ‘insipid, bland’ (cf. Old Sp. *ávol*, *OEH*), *ziho* ‘tallow, suet’ (Gasc. *seu*, *sevós*, Sp. *sebo*).
- d) Old compounds with initial aspiration in the second member: *aihári* ‘supper’ (< *gau+hari*), *behatz* ‘thumb’ (< *be+hatz*).⁹
- e) Old compounds with aspiration (from glottalization) at morpheme boundaries: *gaiháldi* ‘nighttime’ (< *gau+aldi*), *gaihérdi* ‘midnight’ (< *gau+erdi*).
- f) Likely from /-n-/: *ahal* ‘power, possibility’ (cf. *añalke*), *bahe* ‘sieve’ (< **bane*).
- g) Other: *aiher* ‘prone to happen’, *aho* ‘mouth, opening’, *ahür* ‘palm of the hand’, *auher* ‘lazy’ (common *alfer*), *beha* ‘to listen’, *beharri* ‘ear’ (< **beha+(g)arri*), *behar* ‘necessity, need’, *behe-ra* ‘downwards’, *behi* ‘cow’, *bihi* ‘grain’, *bihotz* ‘heart’, *bohor* ‘mare’ (<

⁹ We consider that a compound may be old if it shows the application of phonological changes or its meaning is not transparently derivable from that of its components. Applying these criteria, a word like *artzain* ‘shepherd’, for instance, may be considered an older compound than *ardi-belhar* ‘a type of grass’, literally ‘sheep-grass’.

behor), *buhurri* ‘twisted’ (common *bihurri*), *eharrausi* ‘yawn’, *ehatz* ‘broom’, *ehe* ‘laundry water’, *eihar* ‘arid, shriveled’, *leher* ‘bursting’, *lehia* ‘haste, zeal’, *leiho* ‘window’, *lohi* ‘mud’, *ohe* ‘bed’, *ohil* ‘to flee’, *ohol* ‘plank’, *oihal* ‘cloth’, *oihü* ‘cry, shout’, *saihets* ‘side’, *sohaio* ‘indolent’, *xahal* ‘calf, heifer’, *zahar* ‘old’, *zahi* ‘dandruff’, *zohardi* ‘clear sky’, *zühür* ‘prudent’.

Of greatest interest are, of course, the items in (6g), whose etymology is unknown. In some of these items aspiration may have its source in /-n-/, even if nasalization was later lost. This is the case with *bahitü* ‘to seize’, already mentioned above. We have listed in (6f) a couple of items for which this development seems very likely. This is the cases with *bahe* ‘sieve’, which Mitxelena reconstructs as **bane*, a borrowing from a Romance source **van*. Larrasquet transcribes *ahal* ‘be able’ and its inflected (*ahalaz*) and compounds (*ahal-bezala*) without nasalization, but *ãhâlke* ‘shame; shyness’ with nasalized vowels, showing that these two words are synchronically unrelated: /ahal/ vs. /ãhalke/. In spite of this, there is a very likely etymological relation between the two words: *ahal* ‘power’ + *-ke* ~ *-ge* ‘without’ > *ahatge*, as Mitxelena proposed. Nasalization on vowels can be relatively unstable. As a comparison, notice that, in Portuguese, in both *lua* ‘moon’ and *verão* ‘summer’ an intervocalic /n/ has been lost, but in the former of these two words the vowels subsequently lost their nasal character. Thus, it could just be the case that in a number of items in (6g) intervocalic aspiration resulted from earlier /-n-/, with loss of nasalization at a later stage; we just don’t have the evidence.

In some other items /h/ may be etymological in the sense that it continues an old intervocalic fricative in Common Basque and older stages. For *behi* ‘cow’, the compositional form *bet-* (e.g. *betizu* ‘wild cattle’) suggests **/beɣi/*. In *zahar* ‘old’, the evidence of Vasconic onomastics (i.e., the name *umme sahar* in the Lerga inscription) guarantees the ancient character of the fricative. On the other hand, in *leher* ‘bursting’, there could be vowel doubling from earlier **ler*, just as in *lahardeki* ‘to skin’.

4. Post-sonorant aspiration: /nh/, /lh/, /rh/, /rh/

In Zuberoan and the other northern dialects, word-internal aspiration may occur not only after a vowel or glide, but also after a non-labial nasal or a liquid, as in *elhorri* ‘thorn’. It is clear that in some of these words, the aspiration goes back to Common Basque, since in medieval texts for the

Southwest of the Basque County we find toponyms like *Elhorriaga*. It is equally clear, however, that in other instances, the aspiration is not etymological, since we find it in loanwords that lack aspiration in this position in the source language, e.g., *solhas* ‘conversation’ < Gasc. *solatz*; *sorho* ‘field’ < Lat. *SOLU*, *onhest* ‘honest’, etc. In these words, aspiration appears to be a stress-induced phenomenon. That is, aspiration was inserted immediately before the stressed vowel at a time when the stress-accent fell regularly on the second syllable counting from the beginning of the word, as hypothesized by Mitxelena (1977). In many other cases, when the word is not obviously a borrowing, it is difficult to know whether the post-sonorant aspiration is a conservative feature or an innovation, as in, for instance, *alhaba* ‘daughter’, *ilhun* ‘dark’, etc.

In Larrasquet’s dictionary, we find 197 words with *nh*, *lh*, *rh*, which are reduceable to 96 lemmata, themselves reduceable to 75 or fewer roots. These sequences are always found between the first and second syllable of the word. In a few items where /rh/ was simplified to /h/, the rhotic tap is not noted by Larrasquet (e.g. *so(r)ho*). Classifying again by etymological source, we find the groups in (7):

(7) Post-sonorant aspiration (total n=96)

- a) Non-etymological aspiration, in loanwords (n=9): *anhua* ‘provisions’ (< Lat. *ANNŌNA*), *bilho* ‘hair’ (< Lat. *PILU*), *gelhari* ‘chamber maid’ (< Lat. *CELLĀRIU/A*), *uñhú* ‘onion’ (< Gasc. *onhon*), *so(r)ho* ‘field’ (< Lat. *SOLU*).
- b) Heteromorphemic sequences in compounds (n=14): *ilhérri* ‘cemetery’ (< *hil+herri*), *bürdünhúr* ‘ferroginous water’ (< *burdina+hur*) *bürhézür* ‘skull’ (< *buru+hezur*), *háurhùn* ‘good child’ (< *haur+hun*), *egarháu* ‘debilitating temperature’ (< *egarri+haro*), *hunhártü* ‘accept’ (< *hun+hartu*), *hílhòtz* ‘corpse’ (< *hil+hotz*).
- c) Non-etymological /h/ inserted across morpheme boundaries (n=6): *anhártino* ‘until then’ (< *han+artino*), *bürhás* ‘bare-headed’ (< *buru+as*), *bürhüngürü* [bürhüngű] ‘vertigo’ (< *buru+inguru*), *sínhts(-i)* ‘believe’ (< *zin+ets-i*).
- d) /k/ ~ /h/ alternations in suffixes (< *χ) (n=3): *bürhói* ‘egotistic’ (< *buru+ *χoi*), *aurhíde* ‘sibling’ (< *haur+(k)ide*), *janhári* (< *jan+*χari*, cf. *bazkári*).
- e) Other (n=43): *alha* ‘to graze’, *alhaba* ‘daughter’, *alhargün* ‘widow’, *a(r)han* ‘prune’, *a(r)he* ‘harrow’, *a(r)hin* ‘light, fast’, *añharba* ‘spider’, *añhara* ‘swallow’, *elhür* ‘snow’,

e(r)hi ‘finger’, *elhe* ‘word, speech’, *elhorri* ‘thorn’, *olho* ‘oats’, *o(r)he* ‘dough’, *ülhün* ‘dark’, *ürhe* ‘gold’.

Regarding the words in (7e), as mentioned, we know that the aspiration is etymological in at least *elhorri* ‘thorn’. The etymological challenge is to determine which additional words in this group also have original, as opposed to stress-triggered, aspiration. In words with *-lh-* or, in Southern dialects, intervocalic *-l-*, it is generally assumed that the lateral was fortis or geminate, which prevented the application of intervocalic rhotacism. It is however unclear if we should reconstruct, for instance, **eLur* ‘snow’, **iLun* ‘dark’ or **eLχur*, **iLχun* for Common Basque.

5. Aspirated plosives

In the standard reconstruction (following Mitxelena 1977), Common Basque as spoken in the first centuries of our era is reconstructed as having only *lenis* (voiced) plosives in word-initial position, whereas word-medially a two-way *lenis/fortis* contrast obtained. At some later point, word-initial *fortis* plosives started to be allowed. At this stage, these consonants may have been aspirated when on the onset of the stressed syllable and unaspirated elsewhere. Later, a shift of the accent with preservation of the aspiration *in situ*, led to phonemic split and the three-way contrast /b d g/ vs. /p t k/ vs. /p^h t^h k^h/ in both word-initial and intervocalic position that we find in Northern dialects (Mitxelena 1977). Mitxelena notices that in native words and old borrowings voiceless stops are generally aspirated when on the onset of the second syllable, concluding that this generally was the stressed syllable before the shift of the stress-accent.

5.1. Word-initial plosives

Had there been no lexical change since Mitxelena’s earliest reconstructed stage, we would expect to find no voiceless plosives word-initially, whether aspirated or unaspirated, and for voiceless plosives on the onset of the second syllable to be generally aspirated. As new words entered the lexicon, word-initial voiceless plosives should have acquired aspiration if the initial syllable was stressed and should lack aspiration otherwise. In fact, the picture that we find is far more complex. Mitxelena (1977: 208) already remarked that it is hard to know why we have, for instance, *phena* ‘pain’ with aspirated /p^h/, but *pera* ‘pear’ without it.

In Larrasquet’s dictionary we count 138 lemmata starting with aspirated /p^h t^h k^h/ (e.g., *pharte* ‘part’, *thuna* ‘stain’, *khogan* ‘beehive’) and 239 lemmata with unaspirated /p t k/ (e.g., *par-tü* ‘to

come across/meet’, *tilur* ‘linden’, *kobesor* ‘confessor’). When we consider words of well-known etymology, we find all possible connections between aspiration of a word-initial plosive and the location of the stress in the source language, as shown in (8):

(8) Word-initial /p^h t^h k^h/ vs. /p t k/ by the position of the stress in the source

- a) Words with /p^h t^h k^h/, where the initial syllable is stressed in the source language; e.g.: *phena* ‘pain’ (Gasc. *pena*), *pheska* ‘peach; fishing’ (Gasc. *pescà*), *thumba* ‘thomb’ (Sp. *tumba*), *khana* ‘cane’ (Gasc. *cana*), *khantü* ‘proximity’ (Sp. *canto*), *khanbara* ‘chamber’ (Gasc. *cramba*, Lat. *CAMERA*), *küña* ‘cradle’ (Gasc. *cua*, Sp. *cuna*).
- b) Words with /p t k/, where the initial syllable is unstressed in the source language: *kokí* ‘cheeky, trickster’ (Gasc., Fr. *coquin*), *kantü* ‘corner’ (Gasc. *canton*), *kanabera* ‘reed’ (Gasc. *canavèra*), *paper* ‘paper’ (Gasc. *papèr*).
- c) Words with /p^h t^h k^h/, where the initial syllable is unstressed in the source language: *khatiña* ‘chain’ (Lat. *CATĒNA*), *khandera* ‘candle’ (Gasc. *candèra*), *khantore* ‘song’ (Lat. *cant(a)-* + *-ore*), *kharrika* ‘street’ (cf. Gasc. *carrèra*, Lat. *CARRICĀRE*).
- d) Words with /p t k/, where the initial syllable is stressed in the source language: *pera* ‘pear’ (Gasc. *pera*), *pipa* ‘pipe’ (Gasc. *pipa*), *posta* ‘post’ (Gasc. *pòsta*), *tinta* ‘ink’ (Gasc. *tinta*), *tiro* ‘shot’ (Sp. *tiro*), *kobla* ‘verse’ (Gasc. *cobla*).

Regarding the aspirated or unaspirated character of the word-initial plosive, the expected patterns are those in (8a-b). Some near-minimal pairs like *khántü* (cf. Sp. *canto* ‘border’) vs. *kantü* ‘corner’ (cf. Sp. *cantón*), *kokí* ‘cheeky, trickster’ (cf. Gasc., Fr. *coquin*) are suggestive of an effect of stress. Nevertheless, we also find words like those in (8c), where the initial syllable is not stressed in the Romance source, so that the aspiration on the plosive must obey to a different reason; possibly analogical patterns created after the link between aspiration and stress was broken or weakened. This matter is further discussed in the next subsection. Regarding that, in examples like *paper*, *pipa*, *tinta*, *kantü*, etc., one of the two unaspirated voiceless plosives is necessarily on the onset of the stressed syllable. These appear to be all recent borrowings, since they are essentially identical to the corresponding words in the neighboring Romance languages.

5.2. Plosives on the onset of the second syllable

Under the aspiration-to-stress hypothesis, in native words and old borrowings, voiceless plosives on the onset of the second syllable should be aspirated (except when immediately preceded by a sibilant), since stress-accent on the second syllable was at some point the general rule, whether at the Common Basque stage (Mitzelena 1977; Martínez Areta 2004), or at a later stage, and only in the Eastern area (Hualde 1993b, 2007, 2012, 2022a, 2022b; Egurtzegi & Elordieta 2023). This is what we systematically find in participles: *ethorri* ‘to come’, *ekharri* ‘to bring’, *ikhusi* ‘to see’, etc., with the apparent exception of *jakin*, which is explained as **e.ákin*. It is also what we generally find in the native lexicon, *aphal* ‘low’, *okher* ‘twisted’, *ekhi* ‘sun’, and in old borrowings, *bekhatsu* ‘sin’, *zaphore* ‘flavor’, *aphez* ‘priest’. An interesting alternation in aspiration in noun/verb pairs that we find both in Larrasquet’s dictionary and Leizarraga’s work is between unaspirated plosive in the noun and aspiration in the verb, as in *maite* ‘beloved’ vs. *maitha-tü* ‘to love’, *jokü* ‘game’ vs. *jokha-tü* ‘to play, gamble’. This seems to reflect a former difference in stress pattern, which we still find in the High Navarrese variety of Goizueta (Hualde 2026), where stress was not shifted to the penult. In Larrasquet’s dictionary we also find the minimal pair *jokha* ‘to play’ vs. *joka* ‘hitting’ < *jo+-ka*, also explainable as **jokhá* vs. **jóka*. Mitxelena notices the existence of some clearly old borrowings such as *bake* ‘peace’ and *mutu* ‘mute’, whose intervocalic plosive is unaspirated both in Leizarraga and in Larrasquet. Again, these words have exceptional initial stress in Goizueta (Hualde 2024).

The automatic connection between plosive aspiration and stress-accent was broken after the shift of the stress to the penultimate syllable in the East, which took place at some point before the 16th century, since the contrast between aspirated and unaspirated plosives was already phonological for B. Etxepare and J. Leizarraga. Obviously, words entering the lexicon after the stress shift would not necessarily be adapted in this manner. The existence of stress alternations within word paradigms after the shift, as in *ákher*, *akhérra*, *akherrék* ‘billy goat’ (uninfl., abs. sg., erg. pl.) could also easily lead to restructuring of older words.

What is then the overall frequency of aspirated and unaspirated voiceless plosives on the onset of the second syllable in Larrasquet’s dictionary? Excluding words immediately preceded by a sibilant or followed by a liquid, where only unaspirated plosives are possible, we count 152 instances of /p^h t^h k^h/ and 352 of unaspirated /p t k/. Aspiration of plosives in this position, which,

as Mitxelena pointed out, is the general pattern in old words, has thus become a minority option in 20th c. Zuberoan.

Recall that originally all word-initial plosives were voiced. When initial voiceless plosives started to be allowed, most often it must have been the case that the initial syllable was stressed and therefore aspirated. Once stress-accent lost its role as a conditioning factor, new analogical patterns may have arisen favoring aspiration of word-initial (but not word-medial) plosives. This would account for the fact, noticed by Martinet (1950), that in words with voiceless plosives on the onset of both the first and the second syllable, if one of the two is aspirated, it can only be the first one, as in *khaka* ‘excrement’.

6. Word-initial /h/

Finally, word-initial /h/ is very frequent in Zuberoan, more so than in other dialects with this phoneme, to the extent that Trask (1997: 162-163) suspects non-etymological aspiration in many instances. Trask notices in particular monosyllabic words like *hur* ‘water’, *hun* ‘good’ and *huñ* ‘foot’, which lack aspiration in other dialects, respectively, *ur*, *on*, *oin* (see also Mitxelena 1977: 209). We find 583 distinct words with initial *h*- in Larrasquet’s dictionary. This is a sizable number, likely much greater than for any other consonant, but the dictionary contains more than three times as many words starting with a vowel, 1911. We may conclude that Trask is likely to be right, at least for some items. Outside of Zuberoan as well, we find instances where word-initial aspiration is an innovation, as in *hira* ‘ire’, *husu* ‘use’, *harma* ‘arm, weapon’, *harroka* ‘rock’, etc. This forces us to carefully consider all evidence before reconstructing word initial aspiration. Nevertheless, each example would need to be considered individually. Thus, in the *EHHE*, words like *hur* ‘water’ (Standard Bq. *ur*) and *hoin* ‘foot’ (Standard Bq. *oin*) are reconstructed with initial aspiration, and the seemingly unjustified initial aspiration of *harea* ‘sand’ is reconstructed as arising from metathesis.

7. Conclusion

The main findings of this paper are the following. First of all, the constraint against more than one aspiration per word domain postulated by Lafon is systematically obeyed in monomorphemic words and with the participial suffix *-tü*. On the other hand, there are very few examples of its

active application as a dissimilatory process in polymorphemic words, even when one of the instances of aspiration is not etymological and was inserted at the compound boundary.

Mitxelena's constraint against aspiration after the second syllable has only four exceptions in monomorphemic environments in Larrasquet's dictionary (only one more than the exceptions that Mixelena himself noticed). On the other hand, this constraint is freely violated in compounds. It is reasonable to assume that both constraints operated regularly at the diachronic stage where the stress-accent normally fell on the second syllable of the word and that they were weakened after the shift of the accent to the penultimate syllable, which weakened the link between aspiration and stress-accent (e.g. *ákher*). Compounds with more than one aspiration or with a single aspiration beyond the second syllable may either be recent formations or, if they are older, may have had two accented syllables.

The overall impact of the weakening of /-n-/ in the frequency of intervocalic aspiration is hard to gauge with precision, given the fact that there is evidence of nasalization having been lost in some instances, and, on the other hand, a few clear cases of spontaneous nasalization. We noticed the special problem presented by words containing both aspiration and vowel nasalization but also containing a nasal stop. Leaving aside these words, in Larrasquet's dictionary we find slightly over a dozen native stems where aspiration has most likely resulted from */-n-/ (21 if we count each negative word with **e-no-* separately and include also words of likely onomatopoeic origin; see (5) above).

Regarding aspirated stops, the pattern established by Mixelena for the old lexicon where the *akher* type is more frequent than the *jakin* type has been completely altered in the stage of language reported by Larrasquet.

With this detailed analysis of aspiration in Larrasquet's dictionary, we hope to have contributed to a better understanding of the historical causes for the restrictions that we find in the distribution of aspiration and the reasons that explain exceptions to the rules.

References

Axular, Pedro. 1643. *Gero*. Bordeaux: G. Milanges (Edition and Spanish translation by L. Villasante, 1964, Barcelona: J. Flors; Critical edition by B. Urgell, 2015, Bilbao: Euskaltzaindia & Nafarroako Gobernua).

- EHHE = Lakarra, Joseba A., Julen Manterola & Iñaki Seguro. 2019. *Euskal hiztegi historiko-etimologikoa (EHHE-200)* [Basque historical-etymological dictionary]. Bilbao: Euskaltzaindia. <https://www.euskaltzaindia.eus/ehhe>
- Egurtzegi, Ander. 2018. On the phonemic status of nasalized /h̃/ in modern Zuberoan Basque. *Linguistics* 56(6). 1353-1367. <https://doi.org/10.1515/ling-2018-0024>
- Egurtzegi, Ander. 2023. /h̃/ hasperen sudurkarituaren inguruan. *ASJU* 57(1-2). 275-294. <https://doi.org/10.1387/asju.25952>
- Egurtzegi, Ander & Gorka Elordieta. 2023. A history of the Basque prosodic systems. *Diachronica* 40(1). 30-72. <https://doi.org/10.1075/dia.20066.egu>
- Ettxepare, Bernart. 1545. *Linguae Vasconum Primitiae*. Bordeaux: François Morpain.
- Gonzalez-Eppherre, Jean-Philippe & Beñat Oihartzabal. 2025. L'aspiration en Basque: phonématique et analogy. *ASJU* 59(1). 177-211. <https://doi.org/10.1387/asju.27517>
- Hualde, José Ignacio. 1993a. Topics in Souletin Phonology. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *Generative Studies in Basque Linguistics*, 289-327. Amsterdam & Philadelphia: Benjamins. <https://doi.org/10.1075/cilt.105.12hua>
- Hualde, José Ignacio. 1993b. On the historical origin of Basque accentuation. *Diachronica* 10(1). 13-50. <https://doi.org/10.1075/dia.10.1.03hua>
- Hualde, José Ignacio. 2003. Segmental phonology. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 15-64. Berlin: Mouton De Gruyter.
- Hualde, José Ignacio. 2007. Historical convergence and divergence in Basque accentuation. In Carlos Gussenhoven & Tomas Riad (eds.), *Tones and tunes*, vol. 1, 291-322. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110207569.291>
- Hualde, José Ignacio. 2012. Two Basque accentual systems and the notion of pitch-accent language. *Lingua* 122(13). 1335-1351. <https://doi.org/10.1016/j.lingua.2012.05.003>
- Hualde, José Ignacio. 2018. Aspiration in Basque. *Papers in Historical Phonology* 3. 1-27. <https://doi.org/10.2218/pihph.3.2018.2602>
- Hualde, José Ignacio. 2022a. Accent shift and the reconstruction of Old Common Basque accentuation. In Junko Ito, Haruo Kubozono & Armin Mester (eds.), *Prosody and prosodic interfaces*, 29-62. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198869740.003.0002>
- Hualde, José Ignacio. 2022b. The reconstruction of Old Common Basque accentuation: Closing open issues. *ASJU* 56(2). 77-106. <https://doi.org/10.1387/asju.23983>
- Hualde, José Ignacio. 2024. Azentua eta hasperena: QED [Accent and aspiration: QED]. *Fontes Linguae Vasconum* 137. 235-241. <https://doi.org/10.35462/flv137.10>

- Hualde, José Ignacio. 2026. On the historical link between accent and aspiration in Basque. *Journal of Historical Linguistics*. Published online 09 Feb 2026. <https://doi.org/10.1075/jhl.24021.hua>
- Igartua, Iván. 2008. La aspiración de origen nasal en la evolución fonética del euskera: un caso de *rhynoglottophilia*. *ASJU* 42(1). 171-189. <https://doi.org/10.1387/asju.2305>
- Igartua, Iván. 2015. Diachronic effects of rhynoglottophilia, symmetries in sound change and the curious case of Basque. *Studies in Language* 39. 635-663. <https://doi.org/10.1075/sl.39.3.04iga>
- Lafon, René. 1948. Remarques sur l'aspiration en basque. *Mélanges offerts à M. le Professeur Henri Gavel*, 56-62. Toulouse: Privat.
- Lafon, René. 1958. Contribution à l'étude du parler basque de Larrau (Haute-Soule). In Diego Catalán (ed.), *Miscelánea homenaje a André Martinet*, vol. 2, 77-106. La Laguna (Repr. in Jean Haritschelhar & Piarres Charritton (eds.), *Vasconiana*, 113-133. Bilbao: Euskaltzaindia).
- Lakarra, Joseba A. 2009. **h3 > h1, *h2 > h1* eta horiei datxezkien zenbait fenomeno [On **h3 > h1, *h2 > h1* and some related phenomena]. *Lapurdum* 13. 211-231. <https://doi.org/10.4000/lapurdum.2091>
- Lakarra, Joseba A. 2011. Gogoetak euskal dialektologia diakronikoaz: Euskara Batu Zaharra berreraiki beharraz eta haren banaketaren ikerketaz [Thoughts in Basque diachronic linguistics: On the need to reconstruct Old Common Basque and on the investigation of its diversification]. In Irantzu Epelde (ed.), *Euskal dialektologia: lehena eta oraina* [Basque dialectology: Past and present] (*Supplements of ASJU* 69), 155-241. Bilbao: EHU.
- Larrasquet, Jean. 1939. *Le basque de la Basse-Soule Orientale* (*Collection Linguistique* 46). Paris: Société de Linguistique de Paris.
- Leizarraga, Joannes. 1571. *Jesus Christ gure Jaunaren Testamentu Berria* [The New Testament of our Lord Jesus Christ]. La Rochelle: Pierre Hautin.
- Manterola, Julen & José Ignacio Hualde. 2021. Old Basque had **/χ/*, not */h/*: medieval data, implications for reconstruction and Basque-Romance contact effects. *Journal of Historical Linguistics* 11(3). 421-456. <https://doi.org/10.1075/jhl.19041.man>
- Martinet, André. 1950. De la sonorisation des occlusives initiales en basque. *Word* 6. 224-236 (Repr. in André Martinet, 1964, *Économie des changements phonétiques*, 2nd ed., 370-388. Bern: A. Francke).
- Martínez Areta, José Miguel. 2004. El acento proto-vasco. *ASJU* 38(1). 135-206. <https://doi.org/10.1387/asju.8869>
- Mixelena, Koldo [Michelena, Luis]. 1951. La distribución de las oclusivas aspiradas y no aspiradas. *BAP* 7. 539-549.
- Mixelena, Koldo [Michelena, Luis]. 1958. À propos de l'accent basque. *Bulletin de la Société de Linguistique de Paris* 53. 204-233.

- Mitxelena, Koldo [Michelena, Luis]. 1977. *Fonética histórica vasca*, 2nd expanded ed. Donostia-San Sebastián: Diputación Foral de Gipuzkoa.
- OEH = Mitxelena, Koldo & Ibon Sarasola. 1987-2005. *Orotariko Euskal Hiztegia / Diccionario General Vasco*, 16 vol. Bilbao: Euskaltzaindia (11th online edition, 2023: <http://www.euskaltzaindia.eus/oeh>).
- Orduña Aznar, Eduardo. 2026. La aspiración en vascónico-aquitano: distribución y posibles restricciones. *Veleia* 43. 195-210. <https://doi.org/10.1387/veleia.27292>
- Trask, R. L. 1997. *The history of Basque*. London: Routledge.