

*Metathesis of aspiration as the source
of anticipatory voicelessness in Basque*

Ander Egurtzegi – Institute of Phonetics and Speech Processing, Ludwig
Maximilian University of Munich

To be published in: *Journal of French Language Studies* 29 (2019).

Abstract

In Basque, there is evidence, especially in early loans from Latin, that a sequence #DV(R)T... where D is a voiced stop and T is a voiceless (aspirated) stop was optionally produced with devoicing of the first of these stops. An additional particularity of this sound pattern is that the devoiced word-initial stop typically surfaces with aspiration, while the previously aspirated stop loses it: #DV(R)T^(h)... > #T^(h)V(R)T... This typologically uncommon sound pattern has been described as assimilation of voicelessness in the literature, or spread of [-voiced]. I propose that this sound pattern is triggered by aspiration, not voicelessness, and that it is a case of metathesis, not assimilation. Under the proposed analysis, aspiration of the second stop in the word is reanalysed as originating in the first stop, an instance of *perceptual metathesis*. This approach accounts for the

distribution of aspirated stops before and after the optional change, and the failure of post-sibilant stops to trigger. This account also has implications for the chronology of aspiration-loss in Western dialects: at the time the earliest Latin loans were borrowed, all Basque dialects still maintained a historical series of aspirated stops. Only later, after this process of optional metathesis, did the Western dialects lose *h and stop aspiration.

1. Introduction¹

Basque shows a sporadic sound pattern that affects stops in non-contiguous positions within the word. As observed by Michelena (1977 [2011]: 197-199), Basque dialects show a high degree of variation in word-initial #DV(R).T^(h)V- sequences (D: a voiced stop, T: a voiceless stop). While some varieties have maintained this sequence as #DV(C).T^(h)V-, other varieties show an innovative sequence #T^(h)V(C).TV- in the same words, and within a single dialect, both variants can sometimes be found.

In previous literature, this sound pattern has been analysed as phonological long-distance assimilation of voicelessness or [-voiced]. Here, I propose that it is better analysed as a case of non-local metathesis of stop

¹ The grammatical abbreviations used in the examples are: gen. genitive, nom. nominative, pl. plural, sg. singular. Language abbreviations are: Ar. Arabic, Gas. Gascon, Germ. Germanic, Gr. Greek, Lat. Latin, L.Lat. Late Latin, PIE Proto-Indo-European, Rom. Romance and Sp. Spanish. The Basque dialects mentioned in the examples are: A Alavese, AE Aezkoan, B Bizkaian, G Gipuzkoan, HN High Navarrese, L Lapurdian, LN Low Navarrese, R Roncalese, S Salazarese, Z Zuberoan. I use >, < for regular sound change, >>, << for borrowing, * for reconstructed and ** for unattested forms.

aspiration. Evidence for this reanalysis can be found in the examples of this sound change as attested in the Basque dialects of France (Lapurdian, Low Navarrese and Zuberoan, see the map in Figure 1), also referred to as North-Eastern or Continental dialects. Continental Basque dialects are the only modern varieties that have aspirated voiceless stops and they are the only dialects that continue historical *h as /h/ in specific contexts.

Stop aspiration has been deemed historically secondary and it is still considered so by, e.g. Hualde (2018: 6), primarily due to its limited dialectal distribution and its limited distribution word-internally, in dialects where it does occur. However, Michelena (1977 [2011]: 200) assumed phonetic aspiration in some contexts for Proto-Basque voiceless stops (Martinet's 1950 fortis series), and additional evidence for a series of aspirated voiceless stops in Proto-Basque can be found in Lakarra (2018: 119-121), Blevins (2018) and Egurtzegi (in prep.). Zuberoan, the easternmost dialect of Basque, is the variety that makes the most extensive use of this feature. Some correspondences between voiced/voiced and voiceless/voiceless aspirated stops across Eastern and Western dialects are illustrated in (1). This paper will present additional evidence that the contrast between voiced and voiceless stops illustrated in (1) was historically a contrast between unaspirated and aspirated stops.

(1) Correspondences in stops across Western and Eastern dialects

	Western (Bizkaian)	Eastern (Zuberoan)
Voiced	<i>ibili</i> ‘walk’, <i>edan</i> ‘drink’, <i>egon</i> ‘be’	<i>ebili</i> , <i>edan</i> , <i>egon</i>
Voiceless / Voiceless aspirated	<i>etorri</i> ‘come’, <i>ikusi</i> ‘see’, <i>kate</i> ‘chain’	<i>ethorri</i> , <i>ikhusi</i> , <i>khate</i>

[Insert map here]

Figure 1: Historical map of the Basque dialects (19th century) from Egurtzegi (2014: 7).

The #DV(R)T^(h)... > #T^(h)V(R)T... sound change is most common in Romance loans that appear to be borrowed with initial voiced stops, as in *biper* (*piper*) ‘pepper’ << Lat. *piper*. In the Continental dialects, the words that maintain the arguably comparatively older state show a #DV(R).T^hV- sequence, with an aspirated voiceless stop in the second syllable, as in Low Navarrese *bipher*. Words with #DV(R).T^hV- are mainly found in the Lapurdian and Low Navarrese dialects. In contrast, the equivalent words in the Zuberoan dialect show an aspirated voiceless stop in the first syllable, while the stop in the second syllable is realized as a plain, unaspirated

voiceless stop (a non-contrastive allophone of the aspirated voiceless stop).²

Thus, the innovative forms found in Zuberoan Basque show the structure #T^hV(R).TV- as exemplified by the words in (2), where, the older form in the left-most column may reflect borrowing of initial voiceless stops as D-series stops in older Basque (e.g. Lat. *piper*, borrowed as Basque *bip^her*).

(2) DV(R).T^hV... > T^hV(R).TV... in Zuberoan

Older form	Innovative form	Gloss
<i>barkha(tü)</i> (L, LN, Z)	<i>pharka(tü)</i> (Z)	‘forgive’
<i>bipher</i> (L, LN)	<i>phiper</i> (L, LN, Z)	‘pepper’
<i>dithi</i> (L, LN)	<i>thiti</i> (Z)	‘tit, nipple’
<i>gakho</i> (L, LN)	<i>khako</i> (Z)	‘key, hook’
<i>gorphutz</i> (L, LN)	<i>khorpitz</i> (Z)	‘body’
<i>gathea</i> (L, LN)	<i>khatea</i> (Z)	‘chain’
<i>golkho</i> (L, LN)	<i>kholko</i> (Z)	‘breast, lap’

I suggest that the variants in (2) are a consequence of *perceptual metathesis* (Blevins and Garrett, 1998: 510-527; 2004: 121-125; see Egurtzegi, 2011; 2014: §8.2 for more examples in Basque), a sound change by which the aspiration in the onset of the second syllable was interpreted by listeners as originating in the initial stop. Blevins and Garrett (2004: 121-125) argue that perceptual metathesis is a natural phonetic sound change:

² The inherited two-way laryngeal contrast in Basque is maintained in all dialects. Loans and dialect borrowing have given rise to extremely rare cases in aspirating dialects where there may be contrast between voiceless unaspirated and voiceless aspirated stops, as in *merkhatü* ‘market’, from Latin *mercātu(m)*, vs. *merkatu* ‘to become cheap’, from *merke* ‘cheap’, possibly from Latin *merce(m)* (Larrasquet, 1939: 180; Hualde, 2003: 21).

aspiration is one of several phonetic features that show long or multi-segmental domains in speech. Here it is imagined that a speaker of Basque would have produced the word /gathea/ ‘chain’ as something close to [g̥at̪ea], creating a perceptual ambiguity regarding the segmental origin of phonetic aspiration. After hearing an ambiguous sequence such as [g̥at̪ea], a listener faces the decision of which of the two stops was intended to be aspirated by the speaker. If the etymologically unaspirated stop is deemed to be the source of phonetic aspiration, the listener can reinterpret the word as /khatea/ instead of /gathea/, and the perceptual metathesis will have occurred.

This sound pattern of word-initial devoicing of stops followed by a stop in the next syllable seems to be found in all Basque dialects, but initial stop aspiration is only observed in the Continental varieties, which provide the primary evidence for this proposal.

2. Common Basque phonotactics and anticipatory voicelessness

An important factor in understanding the distribution of voiced and voiceless obstruents in the history of Basque relates to a feature of inherited Basque vocabulary: inherited historic forms suggest that in Common Basque, the stage of the language that preceded modern dialectal division

(roughly placed in the 5th -6th centuries, see Michelena, 1981 [2011]), the T-series of voiceless stops were prohibited from occurring word-initially, though they were licit word-medially, after vowels and sonorants in pre-vocalic position (Michelena, 1977 [2011]: 200). The absence of a D vs. T contrast in Common Basque has been attributed to a regular process of word-initial fortis (or aspirate) stop debuccalization whereby T^(h) > h (see Martinet, 1950: 231; Michelena, 1957 [2011]: 66-70; and, more recently, Lakarra, 2018: 119-121; Blevins, 2018). Thus, inherited Basque words with an arguably old word-initial voiceless stop are extremely rare and have been argued to belong to the expressive vocabulary (Gavel, 1920: 317 ss.), e.g. *thu* ‘saliva’, or to instantiate analogical restoration of a root-initial voiceless stop that is preserved in non-initial position, e.g. *thegi* ‘stable, shelter’ (cf. *hegi* ‘roof’, with initial debuccalization, but *aletegi* ‘granary’, *ale* ‘grain’, where /t^(h)/ was non-initial, and therefore, did not debuccalize; Michelena, 1977 [2011]: 206-207; see Lakarra, 2018: 119-121 and Blevins, 2018 for more examples).

As a consequence of this sound change only the D-series was found word-initially. This phonotactic resulted in a clear pattern in loans: all old loanwords with initial oral stops are borrowed into Basque as D-series stops, regardless of the voicing of the stop in the donor language (Martinet, 1950: 224-225; Gavel, 1920: 317 ss.). This pattern is evident in the Basque loans from Romance shown in (3).

(3) Word-initial D << T, D in old loanwords

Latin/Romance		Basque	Gloss
<i>paradīsu</i>	>>	<i>baradizu</i>	‘paradise’
<i>torre</i>	>>	<i>dorre</i>	‘tower’
<i>camellu</i>	>>	<i>gamelu</i>	‘camel’
<i>ceresia</i>	>>	<i>gerezia</i>	‘cherry’

While the antiquity of some of these loans might be questionable, for others, like *gerezia*, the non-palatalized initial velar suggests borrowing prior to the 3rd-5th century (3rd for Straka, 1979: 201; 5th for Allen, 1978: 14). In sum, I suggest that the oldest borrowings reflect this phonotactic, and only later, in words of the form DV(R).T^hV..., did optional perceptual metathesis take place.

More examples of the devoicing sound pattern in loanwords can be found in (4), with the proposed loan source in the leftmost column. Each variant is followed by the list of dialects where it can be found, according to the *General Basque Dictionary* (Michelena and Sarasola, 1987-2018). Note that examples in (4) are meant to showcase the differences in voicing of the word-initial stops, and, therefore, are not representative of the full range of variants for each lexeme. For example, I show *gorputz*, *korputz* ‘body’, but I ignore *korpitz*, *khorpitz*, etc.

(4) DV(R).T^hV... > T^hV(R).TV... in loanwords

Latin/Romance	Borrowed as:	Innovative form	Gloss
Lat. <i>piper</i>	<i>biper</i> (L, LN)	<i>piper</i> (general)	‘pepper’
Rom. <i>pipi</i>	<i>bipi</i> (L, LN)	<i>pipi</i> (B, G, AN, L, AE, S)	‘moth’
Lat. <i>pācem</i>	<i>bake</i> (B, G, HN, L, Z, S, R, A)	<i>pake</i> (B, G, R)	‘peace’
Lat. <i>peccātum</i>	<i>bekatu</i> (B, G, HN, L, LN, Z, S, R, A)	<i>pekatu</i> (B, G)	‘sin’
Lat. <i>decuma/decima</i>	<i>dekuma</i> (AE, R)	<i>tekuma</i> (S)	‘tithe’
Lat. <i>digitale</i>	<i>ditare</i> (HN, L, LN, Z, A)	<i>titare</i> (B, G, HN, L, AE, S, R)	‘thimble’
Germ. * <i>titta</i> or Gr. <i>τίθη</i> , via Rom.	<i>diti</i> (L, LN, A)	<i>titi</i> (B, G, HN, LN, AE, S, Z, R)	‘tit, nipple’
Lat. <i>corpus</i>	<i>gorputz</i> (B, G, HN, L, LN, AE, A)	<i>korputz</i> (B, G, LN, Z, S, R)	‘body’
Lat. <i>cortem</i>	<i>gorte</i> (L, LN)	<i>korte</i> (Z)	‘court’

Lat. <i>parcēre</i>	<i>barkatu</i> (B, G, HN, L, LN, Z)	<i>parkatu</i> (B, S, R, Z)	‘forgive’
Lat. <i>cortīna</i>	<i>gortina</i> (L)	<i>kortina</i> (B, G, HN)	‘curtain’
Lat. <i>catena</i>	<i>gate</i> (G, L, LN, A)	<i>kate</i> (B, G, HN, S, Z)	‘chain’
L.Lat. <i>cattum</i>	<i>gatu</i> (HN, L, LN, AE, S, Z, R)	<i>katu</i> (general)	‘cat’
Ar. <i>kutub</i> via Rom.	<i>gutun</i> (L, LN, Z)	<i>kutun</i> (B, G)	‘letter’
Lat. <i>cepulla</i>	<i>gipula</i> (A)	<i>kipula</i> (B, G, HN)	‘onion’
Lat. <i>colpu</i>	<i>golko</i> (L, LN)	<i>kolko</i> (B, G, HN, AE, S, Z)	‘breast, lap’
Lat. <i>catillu</i>	<i>gatilu</i> (L, LN, Z, A)	<i>katilu</i> (B, G, HN, R)	‘bowl’
Lat. <i>cupa</i>	<i>gupel</i> (L, LN)	<i>kupel</i> (G, HN, S)	‘barrel’

Words with a word-initial voiced stop are considered comparatively older based on the lack of word-initial voiceless stops in older stages of the language. Recall that inherited voiceless stops underwent debuccalization, and were continued as /h/. It should also be mentioned that the oldest attestations of particular words tend to show a voiced initial stop: although the dialectal distribution of the variant *bip(h)er* is nowadays much more

restricted than *p(h)iper*, the oldest attestation of this word may be *Johan Biperr* (1360).

The same process of aspirate metathesis leading to innovative forms with voiceless stops is also found in potentially inherited lexical items, including those in (5). Most of these are compound words or derived forms where the voicelessness of the initial consonant of the second member is a consequence of regular historical compound-boundary devoicing (e.g. *ogi + bil > og+bil > oTbil > op(h)il*). For example, for 5a, the second root is from *gor ‘hard’, for 5b, the second member is *gain* ‘top, uppermost’, and for 5c, the second member of the compound is *gorotz* ‘manure’.

(5) DV(R).T^hV... > T^hV(R).TV... in potentially inherited words

Older form	Innovative form	Gloss
<i>bikor</i> (G, HN, L, LN, R; cf. <i>bihi</i> ‘seed’)	<i>pikor</i> (G, HN, S, R)	‘grain, grape’
<i>bikain</i> (B, G, HN)	<i>pikain</i> (L, LN)	‘best, first’
<i>bekorotz</i> (general; cf. <i>behi</i> ‘cow’)	<i>pekorotz</i> (B, G)	‘cow dung’
<i>galte</i> (R, A; cf. <i>galdu</i> ‘lose’)	<i>kalte</i> (B, G, HN, L, LN)	‘harm, damage’
<i>gako</i> (L, LN)	<i>k(h)ako</i> (B, G, R, Z)	‘key, hook’
<i>goipe</i> (L, LN, Z)	<i>koipe</i> (B, G, A)	‘fat, oil’

<i>galpar</i> (G)	<i>kalpar</i> (G, HN)	‘hair’
<i>gurpil</i> (G; cf. <i>gurdi</i> ‘cart’)	<i>kurpil</i> (G, HN)	‘wheel’
<i>girten</i> (B, G)	<i>kirten</i> (B, G, HN)	‘handle’
<i>gurtu</i> (LN, S, Z, R)	<i>kurtu</i> (HN, LN)	‘bend the head, worship’
<i>gokots</i> (LN)	<i>k(h)okots</i> (HN, L, LN, AE, Z)	‘chin’
<i>gotor</i> (S, R, L, LN, Z)	<i>kotor</i> (S)	‘solid, firm’
<i>gatazka</i> (L)	<i>katazka</i> (G, LN)	‘fight, fighting’

Notice that initial aspiration is only shown for two forms in (5): *k(h)ako* and *k(h)okots*. This is because the other lexical items do not show the initial voiceless variant in Zuberoan or other aspirating varieties. At present, I have no explanation for the observed lexical split: Latin loans are more likely to show aspirate metathesis in Zuberoan than inherited compounds. A further mystery is the failure of expected metathesis to occur in several common words which, superficially, satisfy the structural description of the change, like native *bete* ‘fill, full’, and *gertu* ‘sure, ready, nearby’ (<< Latin *certu(m)*). Neither of these words shows an initial voiceless stop in any known variety of Basque. Finally, as noted earlier,

although common and widespread, this apparent pattern of devoicing/aspiration is not systematic in any Basque variety.

While many mysteries remain, adopting an aspiration-based approach instead of a voicing approach allows us to explain one regularity in the pattern that has not been previously noted. While the metathesis of aspiration can occur over a preceding consonant (e.g. *gorphutz* > *khorputz* ‘body’, *barkhatu* > *pharkatu* ‘forgive’, *golkho* > *kholko* ‘breast’),³ if the potential trigger segment is preceded by a sibilant, there is never metathesis of aspiration/devoicing. The examples in (6) illustrate words that might be expected to undergo initial devoicing/aspiration, but where there is no evidence of a variant in any dialect.⁴ This seems related to the sibilant preceding the medial voiceless stop.

(6) Words with sibilants that do not show devoicing

Word	Gloss
<i>gaztaina</i>	‘chestnut’
<i>gazta</i>	‘cheese’

³ Note that, in most dialects, stops following a nasal tend to undergo voicing (Michelena, 1977 [2011]: 189).

⁴ An anonymous reviewer correctly points out that the General Basque Dictionary includes some of these words with an initial voiceless stop, such as *kastelu*, *kastaña* (HN, R) and *testatu* (R). Nevertheless, there is a single attestation of *kastelu*, and *kastaña* is also rare in the historical sources. In addition to the initial voiceless stop, *kastelu*, *kastaña*, and *testatu* show other signs of recent borrowing into the language, such as the apical sibilant preceding the stop, when Latin sibilants are regularly adapted to a laminal, and the vowel /e/ (instead of /a/) in *dastatu*.

<i>gazte</i>	‘young’
<i>guzti</i>	‘all’
<i>gostu</i>	‘pleasure’
<i>gaizto</i>	‘bad’
<i>gastu</i>	‘expense’
<i>gaztelu</i>	‘castle’
<i>bazter</i>	‘corner’
<i>buztan</i>	‘tail’
<i>distira</i>	‘brightness, sparkle’
<i>dastatu</i>	‘taste’
<i>beste</i>	‘other’ (but cf. <i>bertze</i>)
<i>bost</i>	‘five’ (but cf. <i>bortz</i>)
<i>busti</i>	‘wet’ (but cf. <i>musti</i>)

Our analysis treats these exceptional cases as a consequence of regular deaspiration: as argued by Michelena (1977 [2011]: 209), from the oldest stages of Basque, there is evidence that stops were deaspirated after sibilants ($T^h > T/S_$). Inherited words showing this phonotactic in Zuberoan include *esku* ‘hand’, *asko* ‘many, a lot’, *asto* (< *arsto*) ‘donkey’, *azpi* ‘lower part, thigh’. Since there is no aspiration in any of the medial voiceless stops

in (6), the structural description of metathesis is not satisfied, and it is not in evidence.⁵

Another environment where the apparent devoicing is regularly absent is where stop sequences occur in synthetic verbs. In contrast to analytic verbs, synthetic verbs are not stressed in Basque. One constraint on aspiration in aspirating dialects is that it can only occur in the onset of historically stressed syllables (or before). This results in the lack of phonetic aspiration in the stops of these verbs. Compare the synthetic forms *dator* ‘he comes’, *dakar* ‘he brings’ and *dakusa* ‘he sees’ with the main verbs of their analytic counterparts *ethorri* ‘to come’, *ekharri* ‘to bring’ and *ikhusi* ‘to see’. The lack of aspiration in forms like *dator*, *dakar*, *dakusa*, in turn, prevents metathesis that would have given rise to unattested forms such as ***thator*, ***thakar* or ***thakusa* in the North-eastern Basque dialects.

3. Comparison with previous analyses

The analysis above makes two important assumptions: first, the *T-series in Proto-Basque was aspirated; second, the *D-series in Proto-Basque was, most likely, voiceless unaspirated word-initially, so that metathesis of

⁵ For a similar instance of deaspiration leading to apparent exceptions to sound change, see Iverson and Salmons (1995) on Grimm's Law.

aspiration to D resulted in a voiceless aspirated stop. The analysis differs from earlier proposals in these two details, and in positing metathesis of aspiration (or [spread glottis]) as opposed to an assimilatory process (whether one adopts long-distance assimilation, spreading or correspondence of devoicing or [-voiced]) from the trigger consonant to the word-initial target consonant.

Under the current analysis, metathesis of aspiration is a one-step process, though it is optional. In contrast, older proposals (such as Michelena 1977 [2011]: 198-199; Trask, 1997: 130; Egurtzegi, 2013: 159, etc.) implicit or explicitly assume a number of changes. First, there is assimilatory devoicing of the word-initial stop: *bipher* > *pipher*. Second, there is aspiration of the newly devoiced stop: *pipher* > *phipher*. Third, there is dissimilation of the aspiration of the stop in the second syllable due to the newly created aspirated stop: *phipher* > *phiper*.

While one might argue, on grounds of simplicity, that metathesis of aspiration is simpler, the last stage of the 3-step process above might be justified as an instance of general laryngeal feature dissimilation (v. MacEachern, 1999). Indeed, Michelena (1977 [2011]: 175) proposed a constraint for derived TV(C)TV- sequences created after the ‘voicelessness assimilation’: if a word has two voiceless stops, only the first stop can be aspirated. Further, the main evidence for this constraint are examples

analysed here as aspirate metathesis. In addition, note that a better known dissimilation in Basque, which involves *h...h sequences, is anticipatory. For example, in compounds like /hil+herri/ (*hil* ‘dead’, *herri* ‘town’) a common form is *ilherri* ‘cemetery’, with loss of initial aspiration (cf. Michelena 1977 [2011]: 174; Egurtzegi, 2014: §4.45). Clearly, one general dissimilatory process cannot be used to account for *hilherri* > *ilherri* and **phipher* > *phiper*.

However, there is a more serious problem with the assimilatory analysis. Assimilation of voicelessness or [-voiced] in obstruents is typologically common as a local process, but almost unheard of as a sporadic long-distance process. In her typology of voicing assimilation, Cho (1991; 1999) only lists examples of local voice assimilation. Local assimilation of voicing between adjacent obstruents is found in a wide range of languages including English, Swedish, Sanskrit, Russian, Polish, Catalan, French and Spanish. Evidence for assimilation of both values of voicing ([+voiced] and [-voiced]) can be found: compare English *leaf/leaves* (with lexically aberrant voicing in the plural) with English *cuff/cuffs*, where the regular devoicing of plural /-z/ occurs after a voiceless obstruent. Though obstruent clusters are limited in Basque, they show the same pattern: after coda /s/ or /z/, voiceless apical and laminal sibilants respectively, one finds only /p/, /t/, or /k/ within the phonological word: *esker* ‘thanks’ (**esger), *ezker* ‘left’ (**ezger); *beste* ‘other’ (**besde), *uzta* ‘harvest’ (**uzda); *ospe*

‘fame’ (**osbe), *ezpain* ‘lip’ (**ezbain), etc. While many of these clusters are lexical, the pattern of cluster devoicing is evident in transparently compositional phonological words like *ez[t]akit* ‘I don’t know’, from *ez* ‘no’ + *dakit* ‘I know’.

In contrast, few if any languages are reported with sporadic or regular long-distance assimilation of voicelessness, as suggested by Michelena (1977 [2011]: 198-199) and others (Trask, 1997: 130; Egurtzegi, 2013: 159, etc.) for Basque. While long-distance consonant assimilation or harmony is wide-spread, the features involved are either non-laryngeal (e.g. secondary place features of palatalization, labialization, or pharyngealization; [+/-anterior] for coronals) or laryngeal features assimilating are [+voiced] along with potentially relevant tone, with [-voiced] a default value, not showing evidence of assimilation. Between Hansson (2001, 2004a, 2004b) and Rose and Walker (2004) only a handful of potential long-distance assimilation of voicing in obstruents are found. These include sound patterns in Kera, Ndebele, Zulu, and Yabem. Hansson (2004a) argues that these cases either show extensive tone-voicing interdependence, with low tone triggering agreement in voicing, or that agreement in voicing is explicitly patterned after agreement in ‘at least one other laryngeal feature’.

In contrast, the sound pattern examined in Basque above does not involve voicing. An older ‘voiced’ stop from the D-series becomes aspirated, and there is no interaction with tone (or pitch accent). Instead of proposing an unattested type of long-distance assimilation of voicelessness together with two additional changes of aspiration and deaspiration, I suggest a one-step, well-attested change of featural perceptual metathesis, where the feature in question is aspiration.

Further, perceptual metathesis of aspiration is well described in some of the world’s languages. A clear case, supported by recent phonetic studies, occurs in Andalusian Spanish (Ruch and Harrington, 2014; Ruch and Peters, 2016). In this variety of Spanish, as shown in (7), a sequence of /hT/ (where h < s), is reinterpreted as Th, pronounced as a post-aspirated stop.

(7) Andalusian Spanish (Ruch and Harrington, 2014: 13)

<i>casco</i>	[ˈkahko]	>	[ˈkak ^h o]	‘helmet’
<i>caspa</i>	[ˈkahpa]	>	[ˈkap ^h a]	‘dandruff’
<i>pasta</i>	[ˈpahta]	>	[ˈpat ^h a]	‘paste, pasta’
<i>estaba</i>	[ehˈtaβa]	>	[eˈt ^h aβa]	‘she/he was’
<i>estado</i>	[ehˈtaðo]	>	[eˈt ^h aðo]	‘state’

estanco [eh'taŋko] > [e't^haŋko] 'kiosk'

pestaña [peh'taɲa] > [pe't^haɲa] 'eyelash'

While the process in (7) is local, and could be viewed as a segmental merger of articulations, with retiming of oral and laryngeal events, the laryngeal metathesis described for Cayuga, whereby /Vh/ → /hV/ in unstressed syllables, is more similar to the Basque metathesis described above, with aspiration in one syllable reinterpreted as a feature of the preceding syllable. Representative data from Cayuga is shown in (8).

(8) Cayuga (Foster 1982; Blevins & Garrett 1998: 510)

/kahwistaʔeks/ → [kħa'wisdʔaes] 'it strikes, chimes (a clock)'

/akekahaʔ/ → [a'gekhaaʔ] 'my eye'

/koʔnikooohaʔ/ → [gʔo'nikhwaʔ] 'her mind'

In addition to the shift of aspiration in Basque illustrated in (2), as in *bipher* > *phiper* 'pepper', another metathesis close to the Cayuga type has been described for Basque. Under this sound change, an /h/ in an unstressed position can be reinterpreted as the onset of a more prominent syllable. Examples include: Lat. *arēna* >> *areña > *harea* 'sand'; *hon* 'good' + *etsi*

‘deem’ *honetsi* > *onhetsi* ‘love, esteem’ (for more examples, see Lakarra, 2009: 217-219, 2015: 365-368; Egurtzegi, 2011: 52-53).

Finally, a somewhat parallel sound change to the Basque process of stop aspiration metathesis can be found in Monguor and Santa, two Mongolian languages (Svantesson et al., 2005: 207-208; Jatteau & Hejna, submitted). The comparison between the Old Mongolian reconstructions or the more conservative Shira Yugur forms and the Monguor and Santa forms in (9) demonstrates a sound pattern parallel to that found in Basque.

(9) Monguor (Svantesson et al., 2005: 207-208)

Old Mongolian	Monguor	Santa	Shira Yugur	Gloss
*tot ^h ara	t ^h utor		ht ^h ɔr	‘inside’
*təč ^h in	t ^h icin		təč ^h ɣn	‘forty’
*kak ^h ai	Xqai	q ^h uqei	kak ^h ai	‘pig’
*čok ^h i	č ^h uku			‘to fit’
*pat ^h u	p ^h ati	p ^h utu	pat ^h	‘firm’
*pič ^h i	p ^h uči	p ^h iči		‘to write’

A possible account for this change offered by Svantesson and colleagues (2005: 207) involves the reverse change of that attested in Andalusian Spanish in (7), namely post-aspiration to pre-aspiration of the second stop in the sequence, followed by a change analogous to that found in Cayuga in (8), with spread of the aspiration to the first stop through the

vowel. Another parallel of the Basque sound pattern in (2) can be found in European Romani varieties (Turner, 1959: 491, 494) such as Welsh Romani *p^huĉ* ‘to ask’ (< Prakrit *pucchāi* < Sanskrit *pr̥cchāti*) or Romani *k^habnī* ‘pregnant’ (< Prakrit *gabbhiṇī-* < Sanskrit *garbhiṇī-*).

While the existence of aspirate metathesis as a sporadic change seems well supported in early loans, there remain a small number of seemingly later cases of sporadic word-initial devoicing of stops in Basque dialects that do not follow this pattern. Many such cases have been attributed to a second wave of loanwords or to a late influence of the contact language on words that were already part of the Basque language. Consider words like: *kartzela* ~ *gartzela* ‘prison’ << Spanish *cárcel* and *kurutze* ~ *gurutze* << Latin *cruce(m)* ‘cross’. In these cases, we must assume that the older form of the loan (*gartzela* and *gurutze*) are original, but that high frequency of Spanish *cárcel* ‘prison’ and *cruz* ‘cross’ (or earlier related Romance terms), have resulted in ‘infection’ of these words, showing a shift of voiced to voiceless without a Basque-internal source.

In addition, there are a number of words that show initial devoicing and aspiration but cannot be explained by any of the processes discussed

thus far. These include inherited words such as *gar* ~ *k(h)ar* ‘flame’, *ge* ~ *k(h)e* ‘smoke’, *gorotz* ~ *k(h)orotz* ‘manure’.⁶

Finally, recall that the aspirate metathesis account is able to explain two large classes of exceptions to the process: it does not occur when the voiceless stop is preceded by a sibilant, because, just in this environment, stops are deaspirated; and it does not occur in synthetic verbs (or other unstressed words) because voiceless stops in these words are deaspirated as well. Under an account that assumes the spread of voicelessness or [-voiced], neither of these exceptional patterns is easily accounted for.

4. Implications

Though a sporadic process limited primarily to loanwords in Basque is unlikely to have serious consequences for our general understanding of

⁶ Blevins (2018) explains these forms in terms of etymological roots with medial *h or *n: *gahar, *kheni and *gohor, respectively. Nevertheless, there are other loans with unexpected initial voiceless stops. Words of this kind include: *bidaia* ~ *pidaia* << Bearnese Gascon *biadge* ‘travel’, *gaiola* ~ *kaiola* << Spanish *gayola* (<< Old French *gaiole*) ‘cage, jail’, *berde* ~ *perde* << Spanish *verde* ‘green’, *gidatu* ~ *kidatu* << Old Gascon *guidar* ‘guide’, *keiñu* << *guiño* ‘gesture, sign’, *gisu* ~ *kisu* << Lat. *gypsu* ‘gypsum’ and *kabarra* << Spanish *gabarra* ‘barge’. None of these words show the appropriate phonological context nor do they show any trace of a voiceless stop in the donor language. This implies that some of the cases in (4) and (5) might as well be due to this sporadic context-free change, although devoicing is far less frequent in words without a voiceless stop in the second syllable. Finally, a few words such as *toki* ‘place’ and *tegi* ‘stable’ show an unexpected word-initial voiceless stop, with the former word never showing aspiration in the continental dialects. These words could be late developments from the suffixes *-toki* and *-tegi*, which are, in most cases (i.e. when occurring after the second syllable), not expected to show aspiration, and might have become autonomous lexical items only recently (Michelena 1977 [2011]: 205-206).

Basque historical phonology, one implication of the aspirate metathesis suggested in this paper concerns the phonetic realization of word initial voiced or lenis stops in older stages of Basque. It could be the case that the previously voiced/lenis word-initial stop underwent devoicing as a consequence of metathesis of the aspiration (or long VOT). But, more likely, perhaps, is that the word-initial stop was already phonetically voiceless prior to the sequential reinterpretation of stop aspiration as initial. Thus, the process of metathesis of aspiration as described in (2) presents potential evidence for a time when there were no word-initial voiced (or pre-voiced) stops in Basque. Though Michelena described the word-initial series as ‘lenis’, and many authors (Hualde, 1999; Blevins, 2018, etc.) assume they were voiced, aspirate metathesis, as understood here, suggests a neutralized series of word-initial stops that were voiceless unaspirated word-initially.

The reconstruction of word-initial voiceless lenis stops was already part of Martinet’s (1950) proposal for the Proto-Basque stop system. Martinet reconstructed the opposition of Proto-Basque stops as being grounded on ‘strength’ (fortis vs. lenis) instead of voice (voiced vs. voiceless), as in the modern language. Martinet based the phonetic reconstruction of word-initial stops on the evidence provided by Latin loanwords such as those in (3), where both voiced and voiceless stops are adapted as lenis in Basque. This paper opens the way for revisiting the long-

standing discussion of the reconstruction of Proto-Basque stops in a future work (Egurtzegi, in prep.).

Another implication of the analysis is chronological. Given widespread forms like *piper* in non-aspirating dialects, it could be the case that large-scale loss of aspiration in the Western dialects followed the metathesis sound change.

A final suggestion of this analysis relates to the chronology of metathesis and other better studied sound changes. Following Michelena (1977 [2011]: 246-248, see also Igartua, 2015 and Egurtzegi, in press), a sound change taking $*n > \tilde{h} / V_V$ occurred before Common Basque (ca. 500 CE), and was still in progress during early contact with Romance (e.g. *añate* ‘duck’ << Latin *anate(m)*). Of interest is the form *k(h)oroa* ‘crown’ << Latin *corōna*, among others. The expectation is that this word should be borrowed with initial /g/, like the other examples in (3). However, no variant of this word with an initial voiced stop is attested in any Basque dialect. Given that this form is recorded with an aspirated stop in all continental Basque dialects, it could be that a metathesis of the segmental aspiration resulting from $*n > \tilde{h}$, produced a word-initial voiceless aspirated stop (Latin *corōna* >> **gorona* > **goroñ̥a* > *khoroa*). Potential evidence for this proposal comes from the compound *goronbil* ‘wooden ball adorning a

balaustrade',⁷ which might be composed by *goron-* and *bil* 'round', the former word being an archaic combination form of *koroa* 'crown' (< *gorona << Latin *corōna*). If a pre-Common Basque *gorona were the first member of a compound, it is expected to lose the final vowel, and this loss in turn would bleed *n > ĥ, which only occurs intervocalically.

5. Conclusions

Dialect variation within Basque shows a sound pattern that optionally devoices a word-initial voiced stop when followed by a voiceless stop in the next syllable. This process has been described as a distant assimilation of [-voiced] in the previous literature, yet sporadic distant assimilations of voicelessness are typologically rare. I have proposed an alternative account for this sound pattern: a process that was previously understood as an assimilation of [-voiced] has been accounted for by means of a metathesis of the aspiration of the stop (or [spread glottis]), a process that has been documented in a number of languages, and that also characterizes other patterns in Basque.

This new interpretation of the Basque sound pattern brings at least three important improvements with respect to previous accounts: First, it accounts for the non-occurrence of the sound change when the voiceless

⁷ I'm grateful to an anonymous reviewer for mentioning this word.

stop in the second syllable is preceded by a sibilant (**DVST^hV), due to deaspiration of this stop; second, it avoids an asymmetry regarding the direction of the dissimilation of aspirates in Basque (i.e., dissimilation of segmental aspiration has been observed to occur anticipatorily in Basque yet previous accounts assumed perseverative dissimilation of the aspiration in voiceless stops); and third, it offers potential evidence of the voiceless or devoiced phonetic realization of word-initial lenis stops, already proposed by Martinet (1950) based on independent observations. This analysis also has chronological implications, given that the metathesis sound change is expected to have begun prior to large-scale loss of aspiration in the Western Basque dialects.

This research also presents a broader contribution to theories of phonological typology, given that it provides an alternative analysis for a potential case of assimilation of voicelessness across a vowel. While local voice assimilation is cross-linguistically common in obstruent clusters, long-distance processes of the same kind are unknown.

Finally, critical data for this analysis come from Zuberoan, one of the few remaining dialects which preserve, what we consider to be, the archaic feature of stop-aspiration. With two dialects of Eastern Basque disappearing in the 20th and 21st century, the importance of studying the threatened varieties of the language cannot be underestimated.

Acknowledgements

Thanks are due to three anonymous reviewers as well as to the editors of this volume for their very helpful comments and suggestions. All errors are mine. This research has been partially funded by the Alexander von Humboldt Foundation and the Spanish Ministry of Economy and Competitiveness [FFI2016-76032-P; FFI2015-63981-C3-2].

Address:

Institute of Phonetics and Speech Processing (IPS), Ludwig-Maximilians-Universität, Schellingstr. 3, 80799 Munich, Germany.

E-mail: egurtzegi@phonetik.uni-muenchen.de

References

Allen, W. S. (1978). *Vox Latina: The pronunciation of Classical Latin*, 2nd edn. Cambridge: Cambridge University Press.

Blevins, J. (2018). *Advances in Proto-Basque reconstruction with evidence for the Proto-Indo-European-Euskarian hypothesis*. London and New York: Routledge.

- Blevins, J. and Garrett, A. (1998). The origins of consonant-vowel metathesis. *Language*, 74: 508-556.
- Blevins, J. and Garrett, A. (2004). The evolution of metathesis. In: B. Hayes, R. Kirchner and D. Steriade (eds), *Phonetically based phonology*. Cambridge: Cambridge University Press, pp. 117-156.
- Cho, Y. Y. (1991). A typology of voicing assimilation. In: A. Halpern (ed), *Proceedings of the 9th West Coast Conference on Formal Linguistics*. Stanford: CSLI, pp. 141-155.
- Cho, Y. Y. (1999). Parameters of consonantal assimilation. Munich: Lincom Europa.
- Egurtzegi, A. (2011). Euskal metatesiak: abiaburua haien ikerketarako. *International Journal of Basque Linguistics and Philology [ASJU]*, 45: 1-79.
- Egurtzegi, A. (2013). Phonetics and Phonology. In: M. Martínez-Areta (ed), *Basque and Proto-Basque. Language-internal and typological approaches to linguistic reconstruction = Mikroglottika 5*, Frankfurt am Main: Peter Lang, pp. 119-172.
- Egurtzegi, A. (2014). *Towards a phonetically grounded diachronic phonology of Basque*. PhD dissertation. University of the Basque Country.
- Egurtzegi, A. (In press). On the phonemic status of nasalized /h̃/ in Modern Zuberoan Basque. *Linguistics*, 56.

Egurtzegi, A. (In prep.). Reconstructing the Proto-Basque stop system.

Foster, M. (1982). Alternating weak and strong syllables in Cayuga words. *International Journal of American Linguistics*, 48: 59-72.

Gavel, H. (1920). *Éléments de phonétique basque*. Paris: Édouard Champion.

Hansson, G. Ó. (2001). *Theoretical and typological issues in consonant harmony*. PhD dissertation. University of California, Berkeley.

Hansson, G. Ó. (2004a). Long-distance voicing agreement: an evolutionary perspective. *Proceedings of the Berkeley Linguistic Society*, 30: 130-141.

Hansson, G. Ó. (2004b). Tone and voicing agreement in Yabem. In: B. Schmeiser, A. Kelleher, V. Chand, and A. Rodriguez (eds), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, pp. 318-331.

Hualde, J. I. (1999). Pre-Basque plosives. In: J. Franco, A. Landa and J. Martín (eds), *Grammatical analyses in Basque and Romance linguistics*. Amsterdam: Benjamins, pp. 77-104.

Hualde, J. I. (2003). Segmental phonology. In: J. I. Hualde and J. Ortiz de Urbina (eds), *A grammar of Basque*. Berlin and New York: Mouton de Gruyter, pp. 15-65.

Hualde, J. I. (2018). Aspiration in Basque. *Papers in Historical Phonology*, 3: 1-27.

Igartua, I. (2015). Diachronic effects of rhinoglottophilia, symmetries in sound change and the curious case of Basque. *Studies in Language*, 39: 635-663.

Iverson, G. K. and Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, 12: 369-396.

Jatteau, A. and Hejné, M. (submitted). Gradient dissimilation in Mongolian: Implications for diachrony. *Papers in Historical Phonology*.

Lakarra, J. A. (2009). *h3 > h1, *h2 > h1 eta horiei datxezkien zenbait fenomenoz. *Lapurdum*, 13: 211-231.

Lakarra, J. A. (2015). Hiru hasperen haboro. In: M.-J. Ezeizabarrena and R. Gómez (eds), *Eridenen du zerzaz kontenta: sailkideen omenaldia Henrike Knörr irakasleari (1947-2008)*. Bilbao: UPV/EHU, pp. 349-378.

Lakarra, J. A. (2018). La prehistoria de la lengua vasca. In: J. Gorrochategui, I. Igartua and J. A. Lakarra (eds), *Historia de la lengua vasca*. Gobierno Vasco: Vitoria-Gasteiz, pp. 23-244.

Larrasquet, J. (1939). *Le Basque de la Basse-Soule Orientale*. Paris: Klincksieck.

MacEachern, M. R. (1999). *Laryngeal co-occurrence restrictions*. New York, NY: Garland Publishing.

Martinet, A. (1950). De la sonorisation des occlusives initiales en basque. *Word*, 6: 224-236.

Michelena, L. (1957 [2011]). Las antiguas consonantes vascas. In: D. Catalán (ed), *Estructuralismo e historia. Miscelánea homenaje a André Martinet*, vol. 1. La Laguna: Universidad de La Laguna, pp. 113-157 [J. A. Lakarra and Í. Ruiz Arzalluz (eds). (2011). *Obras completas 7*, pp. 57-91].

Michelena, L. (1977 [2011]). *Fonética histórica vasca*, 2nd edn = *Publicaciones del Seminario 'Julio de Urquijo'*. Donostia-San Sebastian: Diputación Foral de Guipuzcoa [J. A. Lakarra and Í. Ruiz Arzalluz (eds). (2011). *Obras completas 6*].

Michelena, L. (1981 [2011]). Lengua común y dialectos vascos. *International Journal of Basque Linguistics and Philology [ASJU]*, 15: 291-313 [J. A. Lakarra & Í. Ruiz Arzalluz (eds). (2011). *Obras completas 7*, pp. 517-542].

Michelena, L. and Sarasola, I. (1987-2005). *Orotariko Euskal Hiztegia* [*General Basque Dictionary*] (16 vols.). Bilbao: Euskaltzaindia [<http://www.euskaltzaindia.net/oeh/>, accessed on 04/04/2018].

Rose, S. and Walker, R. (2004). A typology of consonant agreement as correspondence. *Language*, 80: 475-531. DOI:

<https://doi.org/10.1353/lan.2004.0144>

Ruch, H. and Harrington, J. (2014). Synchronic and diachronic factors in the change from preaspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics*, 45: 12-25. DOI: <https://doi.org/10.1016/j.wocn.2014.02.009>

Ruch, H. and Peters, S. (2016). On the origin of post-aspirated stops: production and perception of /s/ + voiceless stop sequences in Andalusian Spanish. *Laboratory Phonology*, 7: 1-36. DOI:

<https://doi.org/10.5334/labphon.2>

Straka, G. (1979). *Les sons et les mots : Choix d'études de phonétique et de linguistique*. Paris: C. Klincksieck.

Svantesson, J.-O., Tsendina, A., Karlsson, A. and Franzén, V. (2005). *The Phonology of Mongolian*. Oxford: Oxford University Press.

Trask, R. L. (1997). *The History of Basque*. London and New York: Routledge.

Turner, R. L. (1959). Transference of aspiration in European Gypsy. *Bulletin of the School of Oriental and African Studies*, 22: 491-498.