*Article*

Language *and* Speech

# A Corpus-based Study of Fillers among Native Basque Speakers and the Role of *Zera*

Language *and* Speech
2014, Vol. 57(3) 338–366

DOI: 10.1177/0023830913506422
las.sagepub.com
SAGE

**Xabier Urizar**
Basque Center on Cognition, Brain and Language, Spain; San Fermin Ikastola, Spain

**Arthur G Samuel**
Basque Center on Cognition, Brain and Language, Spain; Ikerbasque, Basque Foundation for Science, Spain; Stony Brook University, USA

**Abstract**
Although speakers often transmit their messages clearly and concisely, their speech also includes disfluencies, including filler words. We have analyzed the kinds of filler-like words (hereafter fillers) that native Basque speakers produce and the role that these fillers have within the discourse. We recorded six Basque L1 speakers in a natural setting designed to trigger spontaneous speech. Because Basque is an agglutinative language it may offer speakers certain options for filler use that have not been observed in studies of languages that do not have such a rich agglutinative morphology (e.g. English). When speakers are close to the retrieval of a to-be-produced word, but not quite able to access it, they may use the agglutinative morphology to give the listener clues to the syntactic category of the intended word. In Basque such clues could be provided by modifying the surface form of a filler.

Our corpus includes approximately 300 filler tokens. We provide analyses of the kinds of fillers this population produces and the contexts in which these appear. Certain fillers tend to be produced before beginning large units (e.g. sentences), whereas others usually precede smaller units. One filler (/zera/) behaves differently than the others. In particular, it assumes context-based forms that offer listeners partial information about the almost-retrieved word.

**Keywords**
Discourse processing, disfluencies, hesitation markers, filled pauses, spontaneous speech, Basque

**Corresponding author:**
Xabier Urizar, Basque Center on Cognition, Brain and Language,Paseo Mikeletegi 69, 2nd Floor, 20009 Donostia, Spain.
Email: urizar.xabi@gmail.com

## Introduction

Although talking to one another is an extremely natural activity, producing a fluent utterance is an impressively complex task. Speakers have to activate concepts, select words, develop the proper syntax, retrieve the correct phonology, and initiate the complex motor patterns needed to move the articulators in exactly the necessary patterns. Some speakers are able to convey their messages clearly and succinctly, while others are more disfluent, prone to *um*s and *eh*s, repetitions and corrections (Fehringer & Fry, 2007). Fillers are typically monosyllabic segments that cannot be associated with any existing words. Phonetically, they most often consist of vocalic segments, mainly [ə,e,a] (Belikova & White, 2009), sounds that are relatively easy to make. In some cases, speakers simply elongate the final vowel of a word, rather than introduce a new segment. In English, the most common instance of lengthening (cf. Fox Tree & Clark, 1997, cited in Clark, 1994) occurs when "the" is pronounced as "thee", with the ending vowel sound drawn out past its usual duration. There is evidence that such vowel lengthening can actually facilitate the listener's processing of language (Arnold, Tanenhaus, Altmann, & Fagnano, 2004).

Fox Tree (1995) estimated that as many as 6% of words uttered are, or are affected by, some form of disfluency (see also Bortfeld, Leon, Bloom, Schober, & Brennan, 2001). There are various types of disfluencies, such as pauses, extended words and filler words such as "*um*", "*uh*", etc. As Arnold et al. (2006) have pointed out, such disfluencies have often been ignored in the psychological study of language production and comprehension. Traditionally, disfluencies have been viewed as "speech errors", distractions from the main content of language and, therefore, not a to-be-studied item. In fact, a common belief was that disfluencies are not even centrally processed, but instead filtered out by the listener to get to the real linguistic input (Bear, Dowding, & Shriberg, 1992; Lendvai, Van Der Bosch, Krahmer, & Canisius, 2004; Stolcke et al., 1996). More recent theories have allowed for the possibility that listeners use disfluencies (Bailey & Ferreira, 2003; Clark & Fox Tree, 2002), but it is not clear exactly how they are handled in language comprehension (Arnold et al., 2004).

The function of these hesitation markers is also still unsettled. Some researchers center theories on fillers' pragmatic or social function in discourse (Maclay & Osgood, 1959), but more work is needed to have a clear idea of what the functions within the discourse are. Maclay and Osgood argued that the function of hesitation markers is to keep the speaker's turn of speech: if the speaker produces a silent pause that is too long, the chance that the listener will interrupt increases. However, it has been noted that hesitation markers are very common in monologues, where there is no possibility of interruption (Schachter, Christenfield, Ravina, & Bilous, 1991), which is a problem for Maclay and Osgood's suggestion. Kunzel (1999) offered an interesting point concerning this issue. He noted that in "telephone speech" there were more fillers than in face-to-face conversation. Because of the absence of a visible listener, the speaker must orally signal to an invisible listener that he or she has not yet finished speaking. This intention can be communicated visually in face-to-face conversations, but not in telephone speech situations (Künzel, 1997).

A number of corpus analyses and behavioral studies suggest that listeners are sensitive to disfluencies (c.f. Brennan, 2001; Corley, MacGregor, & Donaldson, 2007). De Leeuw (2007) contrasts two ways to think about fillers: the symptom hypothesis and the signal hypothesis. Under the symptom hypothesis, hesitation markers are symptomatic of some kind of cognitive process on the part of the speaker. Maclay and Osgood's (1959) suggestion that fillers are used to win a bit of time speech planning can be viewed as a symptom. A related view is that hesitation markers may indicate that the speaker is formulating the next part of an utterance (Crystal, 1992). The approaches share the idea that the fillers are essentially reflecting operations within the speaker. Under the

signal hypothesis, fillers are more directed toward the listener. Consistent with this signaling idea, Fox Tree (2001) found that hesitation markers allowed addressees to recognize an upcoming word more quickly. Clark and Fox Tree (2002) argued that speakers use *uh* and *um* to announce either that they are initiating what they expect to be a minor (*uh*) delay, or a major (*um*) delay, in speaking (Brennan and Schober, 2001).

A similar signaling function can be served by the choice of one or another somewhat more complex discourse marker. For example, Fox Tree and Schrock (2002) found that the seemingly empty phrases "you know" and "I mean" actually tend to occur under different circumstances, potentially providing information for the listener. The phrase "you know" is used when the speaker wants the listener to make inferences, whereas "I mean" is used when the previous statement by the speaker is about to be clarified and adjusted. Similarly, fillers can help listeners understand sentences in which the meaning may be temporarily ambiguous. Fillers occurring at the point immediately before a clause begins aid in the comprehension of the clause as a cohesive statement (Bailey & Ferreira, 2003).

In fact, the most common locus for a filler is at the beginning of an utterance or phrase, presumably as a consequence of the greater demand of planning processes at these junctures (Maclay & Osgood, 1959; Barr, 2001). Disfluencies, particularly filled pauses and repetitions, are more frequent at the beginning of these constituents rather than in the other positions (Clark & Wasow, 1998; Shriberg, 1994). Fillers are more likely to occur before content words (Maclay & Osgood, 1959), a pattern that we will see in our corpus. Shriberg (1994) stated that the longer or more complex a constituent, the higher the disfluency rate immediately before or at the beginning of the constituent. Consistent with this claim, Watanabe, Den, Hirose and Minematsu (2004) found that clauses immediately after filled pauses contained more words than those without preceding filled pauses.

According to Bailey and Ferreira (2007), filled pauses occur in a particular distribution with respect to syntactic (Clark & Wasow, 1998), semantic (Schacter, Christenfield, Ravina, & Bilouset, 1991) or pragmatic (Smith & Clark, 1993) structure. In the case of syntactic structure, filled pauses (and other disfluencies, such as repetitions) are most likely to occur immediately prior to the onset of a complex syntactic constituent (Clark & Wasow, 1998; Shriberg, 1996). Filled pauses are also likely after the initial word in a complex constituent, especially after function words (Clark & Wasow, 1998), and listeners might be able to use the presence of a recent filled pause to predict that an ambiguous structure should be resolved in favor of a more complex analysis (Bailey & Ferreira, 2003; Shriberg, 1994).

There are corpora of fillers in several different languages, including Japanese (Watanabe, Hirose, Den, & Minematsu, 2008), English-German-Dutch (de Leeuw, 2007), American-British English (Wasow, 1998), French (Candeá, 2000), etc. For these languages the corpora provide information about the kinds of fillers, their behavior, distributional patterns, etc. We report here a comparable study using Basque, an isolated language with different syntactic and morphological properties. We have analyzed the kinds of filler words that Basque speakers produce and the role that these fillers have within the discourse. Because Basque is an agglutinative language it may offer speakers certain options for filler use that have not been observed in studies of languages that do not have such a rich agglutinative morphology. When speakers are close to the retrieval of a to-be-produced word, but not quite able to access it, they can potentially use the agglutinative morphology to give the listener clues to the syntactic category of the intended word. In Basque such clues could be provided by modifying the surface form of a filler. As we shall see, there is in fact one filler that shows exactly this property of assuming different forms that are tied to the syntactic properties of the to-be-uttered word (Candea et al., 2005; Clark and Wasow, 1998).

We should note that there is no reason to believe that speakers have an intent to give listeners clues when they produce informative fillers. Rather, the set of activations of various linguistic components (recent and upcoming phonological, lexical and syntactic units) can produce effects that are seen in the instantiations of fillers (cf. Bock and Cutting, 1992). Speakers could also potentially use the rich agglutinative morphology in order to aid access of the target word, similar to what some gestures do (e.g. Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Melinger & Kita, 2007).

In the current study, we examine the kinds of disfluencies a particular Basque L1 population produces; we also examine the distributional pattern of these fillers. Our set of disfluencies includes filler words and vowel lengthenings. Both of these types of disfluencies involve the insertion of extra spoken material, providing the speaker with additional time, and the listener with overt cues. As noted above, we will refer to all of these as "fillers", reflecting the extra spoken material. Our analyses of Basque fillers show some commonalities with analyses of fillers in other languages, as well as the very interesting behavior of a filler that is unique to Basque.

# 2 Methods

## *2.1 Participants*

Six participants took part in this experiment. We selected Basque L1 speakers who have not developed their L2 (Spanish in this case) very much. Thus, the participants were virtually monolingual Basque speakers. They come from a rural area called Basaburua (northwest of Navarra, Spain), and all speak the same variety of Basque. Five of the six speakers were at least middle-aged because this rural group has had less contact with Spanish. The ages of the speakers were: 35, 55, 67, 74, 76 and 89 (mean = 66).

## *2.2 Materials*

The first author (a native Basque speaker) recorded three participants with a Panasonic RR-US360 recorder. Recordings from three other speakers (from the same population) were extracted from a Basque spoken-language recordings database (http://mediateka.euskarabidea.es/en/home). The recording conditions for the speakers in the database were the same as those used in our recordings; the recordings contain spontaneous speech obtained at the speakers' houses. To create an informal atmosphere for the speaker we included a third participant in the conversation (this was also the procedure used for the database recordings). The additional person was the speaker's friend or another person close to him/her.

## *2.3 Procedure*

We recorded each speaker in a familiar environment so that the speaker would feel relaxed, without any pressure. We started the conversations with some simple questions so that speakers could forget they were being recorded. These initial questions were related to their working environment to provide an easy topic to talk about. After some minutes, after they had become comfortable with the recording situation, they started to speak fluently. All of the analyses we report here are based on the speech produced after that point. We transcribed what each speaker said, simply typing what was on the recordings. From the transcriptions we identified any fillers and, for each such filler, we noted the context for its occurrence.

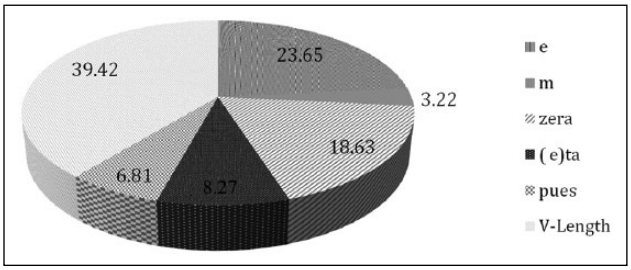


**Figure 1.** Occurrence rate for each of the six fillers

## 3 Results

The material contained 279 fillers out of 16,674 words within 167 minutes of recording. This means that the speakers uttered fillers with a frequency of 1.67 per min and that 1.7% of the total words were fillers. Appendix 1 lists each of the fillers, placed within the context in which it was produced. The 279 tokens were sorted into six category types. The primary fillers used by our speakers included /e/, /m/, "pues" (a common filler in Spanish), and "(e)ta" ("eta" means "and" in Basque). We also observed final vowel lengthening, and a number of variants of the Basque filler "zera". Vowel lengthening differs from the other cases because there is no additional element, just lengthening of an existing segment. Also, as we shall see, "zera" differs from the other fillers in several ways. Figure 1 shows the distribution of these six filler types in our corpus. As the figure illustrates, /e/ and vowel lengthening were the most common fillers, together comprising about two-thirds of the total.

Basque L1 speakers use a number of other fillers in their speech (es que, ba, and so on), but we focus here on the six fillers shown in Figure 1 because these items were used more frequently. Among the more common remaining fillers were: /*ba*/ (three times), "*es que*" (five times) and /*o*/ (eight times). In contrast, the numbers for the six analyzed fillers were: vowel lengthening, 110 times; /e/, 66 times; /*zera*/, 52 times; /(e)ta/, 23 times; "pues", 19 times; and /m/, 9 times. Although /m/ was not much more frequent than some of the unanalyzed fillers, we include it here because in the existing literature it has been reported relatively frequently. In fact, /e/ and /m/ have been widely reported in the literature of speech disfluencies (Clark & Fox Tree, 2002; Corley, 2008; O'Connell et al., 2005; Corley & Stewart, 2008 and O'Conell & Kowal, 2005).

We used a very conservative criterion for identifying a vowel as being lengthened – it had to sound very noticeably longer than normal to the coder to be included. As a check on this selection, for each vowel that was identified as lengthened, we measured the duration of the next instance of that vowel that matched in position within the word and syllabic structure of the word. As Figure 2 shows, the durations of the lengthened and control vowels formed non-overlapping distributions, confirming that the cases of vowel lengthening were indeed disfluencies.

We noted in the Introduction that the location of fillers in discourse is not random – they generally occur when the speaker is having some difficulty planning a linguistic unit. They have been observed immediately prior to the onset of a complex syntactic constituent (Shriberg, 1996). In some cases, the speaker may produce the initial word of the complex constituent, especially if it is a function word, and then produce a filler before being able to continue with the content of the constituent (Clark & Wasow, 1998). To see whether this pattern held for our Basque speakers, we sorted the fillers on the basis of the syntactic category of the following word. If the pattern holds,

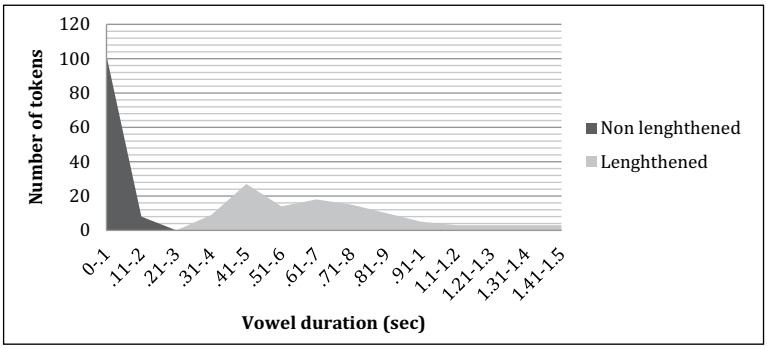


**Figure 2.** Distribution of lengthened and non-lengthened vowels.

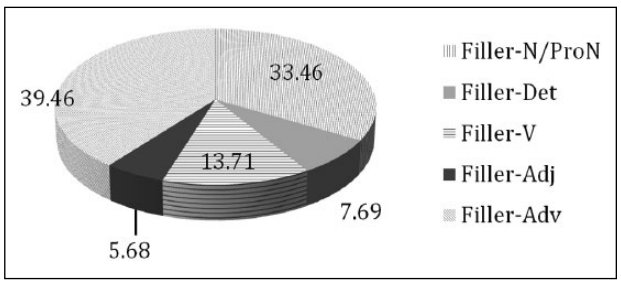


**Figure 3.** Observed frequency of filler ocurrence before each syntactic type.

we should see a different distribution of words following fillers than the overall distribution of words in the corpus. In particular, we should see more fillers before content words, and fewer before function words like determiners. There are also presumably effects of words beyond the one following the filler, but the following word is most clearly defined, and has been shown to matter in previous studies Shriberg, 1994.

Figure 3 shows that about three-quarters of the fillers came before adverbs or nouns, and less than 8% preceded determiners. However, if adverbs and nouns make up three-quarters of the words in our corpus and determiners occur about 8% of the time, then the distribution of fillers would only reflect the distribution of possible places for the fillers to appear. To estimate the distribution of the syntactic categories shown in Figure 4, we categorized the words in a sample of about 500 words from the corpus (see Appendix 2 for a transcription of the sample). The sample was taken from three of the speakers. For each speaker we took a chunk of about 40 words near the beginning of each recording, another 40 about 25% of the way through it, another 40 about halfway through it, and a final 40 about 75% of the way through it. Figure 4 shows the estimated percentage of words in each category that would by chance be expected to follow a filler (Appendix 3 shows the criteria used to categorize each word). As the figure shows, nouns and adverbs only comprise about half of the corpus, much less than the three-quarters of the time that they followed fillers. Conversely, determiners occurred about three times as often in the corpus as in the position following a filler. This pattern is consistent with previous reports that fillers tend to occur after function words, rather than before them. This result suggests that speakers sometimes can generate an initial function word of a complex constituent before stalling on the production of more contentful nouns and adverbs.

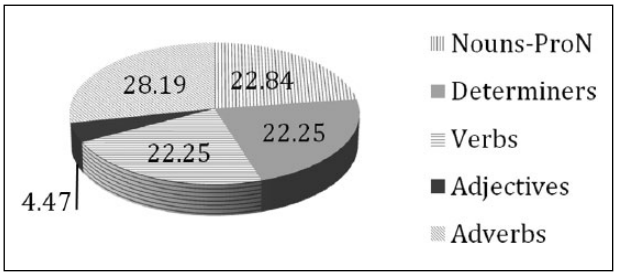


**Figure 4.** Relative frequencies of syntactic types.

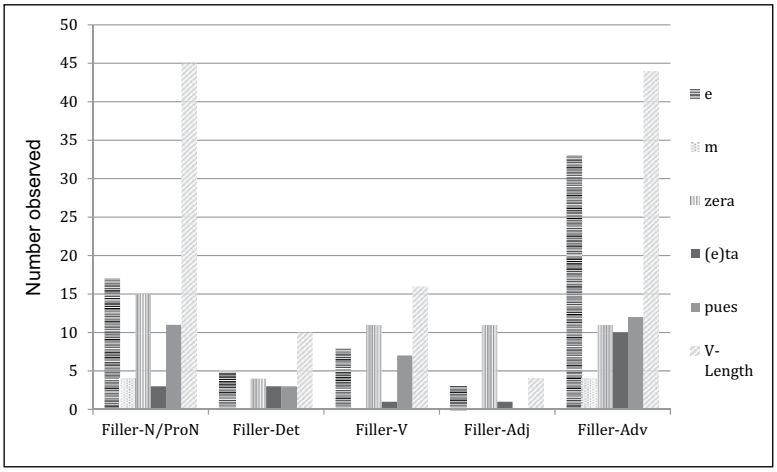


**Figure 5.** Distribution of fillers as a function of the syntactic role of the following word.

Recall that Fox Tree and her colleagues have reported that speakers use different fillers in different contexts. For example, they may use *uh* to signal a minor delay, and *um* to signal a major delay (Clark & Fox Tree, 2002). Similarly, Fox Tree and Schrock (2002) found that speakers use "you know" and "I mean" under different circumstances, potentially providing information for the listener. To see if there was a comparable distribution of roles to different fillers in our corpus, we broke down the occurrences of each filler by the type of word that followed. Figure 5 shows these filler distributions. As is clear in the figure, the overall dominance of vowel lengthening shown in Figure 1 is driven by its appearance before nouns, and especially before adverbs. Similarly, /e/ is heavily used before adverbs but is not particularly dominant in any other context.

a disfluency at the beginning of the constituent, and that Watanabe et al. (2004) found that clauses immediately after filled pauses contained more words than those without preceding filled pauses. These findings suggest that fillers often occur as a result of the extra planning needed when a speaker is about to produce a complex structure. With this in mind, we partitioned the observed filler occurrences based on whether the filler preceded a sentence (i.e. intersentential) or instead occurred during the production of a major constituent, presumably when less long-range planning was required (i.e. intrasententially). Table 1 presents this analysis. Four of the fillers (vowel lengthening, "pues", /zera/ and /eta/) primarily appeared within sentences, while the other two (/e/ and

**Table 1.** Intersentential versus intrasentential occurrences of fillers.

| | Vowel lengthening | *pues* | /zera/ | /eta/ | /e/ | /m/ |
|---|---|---|---|---|---|---|
| Intersentential | 30 | 4 | 7 | **10** | **42** | **6** |
| Intrasentential | **80** | **15** | **45** | **16** | 24 | 3 |
| Intersentential (%) | **27%** | **21%** | **13%** | **38%** | **64%** | **67%** |

**Table 2.** Durations of fillers and any surrounding gaps.

| Filler | Number of tokens | Mean duration of filler and the preceding/following gaps (ms) | | | Range of duration (ms) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Before | Filler | After | Before | | Filler | | After | |
| | | | | | Min | Max | Min | Max | Min | Max |
| /e/ | 66 | 500 | 694 | 260 | 0 | 2670 | 270 | 2620 | 0 | 1440 |
| Vowel lengthening | 110 | – | 748 | 410 | 0 | 0 | 300 | 1660 | 0 | 2660 |
| /zera/ | 52 | 120 | 367 | 121 | 0 | 450 | 150 | 870 | 0 | 970 |
| *pues* | 19 | 345 | 444 | 277 | 0 | 1120 | 240 | 980 | 120 | 890 |
| /eta/ | 23 | 303 | 638 | 381 | 0 | 800 | 340 | 1180 | 0 | 1070 |
| /m/ | 9 | 527 | 451 | 258 | 0 | 1150 | 230 | 780 | 0 | 450 |

/m/) were much more likely to be produced between sentences. Thus, we suggest that /e/ and /m/ are related to planning large units of an utterance, while "pues", /zera/, /eta/ and vowel lengthening are associated with difficulties in accessing smaller phrases or individual words. This is particularly clear for /zera/, with 87% of its occurrences located intrasententially. As we will discuss shortly, production of /zera/ can clearly be linked to production of a particular following word, consistent with the analysis shown in Table 1. Recall that Shriberg (1994) indicated that more complex constituents tend to have a disfluency at the beginning of the constituent, and that Watanabe et al. (2004) found that clauses immediately after filled pauses contained more words than those without preceding filled pauses.

We used a waveform editor to examine the speech including and surrounding each filler in the corpus. Table 2 presents the results of this examination, showing the duration of each type of filler and the duration of any preceding or following gap in speech production. Note that, for vowel lengthening, there is no preceding gap because the filler itself is just an extension of an existing segment. If we consider the sum of the filler duration and the duration of any surrounding pauses to be an estimate of the total filler event, then /zera/ again seems to be rather different to the other fillers: a /zera/ event takes only about 600 ms, about half the time of the other filler events. In that sense, a sentence with /zera/ as a filler is most similar to a fluent sentence. We thus see that /zera/ differs from the other fillers in being so consistently used intrasententially, and that it is in some sense more fluent than other disfluencies. We now turn to a property of /zera/ that is particularly distinctive – its ability to take different surface forms that depend on the syntactic role of the word that is being delayed.

Syntactically /zera/ is a demonstrative, but as a filler it has the ability to take on different syntactic roles, such as verb, noun, adjective, etc. That is, in addition to appearing in its "bare" form,

it can be conjugated in the same way as verbs are and it can take on different declination endings. We suggest that Basque speakers use /zera/ when they are encountering some difficulty in word retrieval, but not a total failure. That is, just as people in the tip-of-the-tongue state can often correctly report information about the difficult-to-retrieve word (e.g. its first sound and its number of syllables), speakers may know the syntactic role of a word that they are having trouble retrieving. This syntactic information can be attached to /zera/, taking advantage of the agglutinative nature of Basque. In this way, the speaker can use the filler /zera/ to provide the listener with potentially useful information. The accurate cueing of this information, together with the much shorter duration of /zera/ events, suggests that /zera/ is selected as a filler only when the speaker is close to achieving lexical access of the desired word.

We are not aware of any previous reports of a filler that changes form to this degree in a way that matches the syntactic properties of a to-be-retrieved word. As this behavior seems to be tied to the agglutinative nature of Basque, it may not be surprising that it has not been seen before because almost all work on fillers has been on non-agglutinative languages like English. The two most prominent other agglutinative languages are Turkish and Hungarian. Gosy has examined filler use in Hungarian, but although Gosy (2007) comments on affixes in the agglutinative language, there is no specific mention of a filler behaving like /zera/. In the case of the Turkish, although Bosker (H.R. Bosker, personal communication, July 5, 2013) has very recently found preliminary evidence for a filler that may have some of the properties of /zera/, he found too few occurrences of this filler to allow analyses of its distribution.

The following examples illustrate how /zera/ can appear in different forms (see Appendix 1 for the complete set of observed forms, and the context for each).

If the to-be-produced word is a verb, the speaker will produce something like this:

(a) *zertu* (literally, zer-tu = zer(a) + verb suffix);

If it is a noun, the speaker might say:

(b) *zera* (without declining);

If the speaker is going to make a comparative sentence, the speaker could produce:

(c) *zeragoa* (literally, zera-goa = zera + comparative suffix);

For a postposition sentence with Basque *inesibo* case, the speaker will produce:

(d) *zean* (literally, zera + inesibo case ending);

To produce a verb in future tense, the speaker may produce:

(e) *zeratuko* (literally, zera + future tense ending).

Thus, /zera/ provides a potentially useful predictor for the listener of what is going to follow. In our corpus, the to-be-produced word was successfully retrieved 63% of the time after /zera/ appered. The successfully retrieved words were usuallly produced immediately after /zera/; delayed access was observed for 12 of the 52 occurrences of zera (23%). In eight of these 12 cases one word intervened between /zera/ and the syntactically-matched word. Overall, /zera/ was almost always followed by a content word (49/52 cases), again consistent with the idea that this particular filler gets used when the speaker is close to successful lexical access, assuming function words have a different source in production than content words.

# 4 Discussion

As we noted in the Introduction, views on the nature of disfluencies have evolved from treating them as noise (Bear et al., 1992; Lendvai et al., 2004; Nakatani et al., 1993; Stolcke et al., 1996) to considering them to be potentially useful cues that speakers provide to listeners (e.g. Arnold et al., 2004; Barr, 2001; Brennan & Schober, 2001; Clark & Fox Tree, 2002; Corley et al., 2007; Fox Tree & Schrock, 2002). In the current study, we examined filler properties in the speech produced by essentially monolingual Basque speakers. This population is of interest not only because Basque is an isolate, and thus potentially different than other languages in which disfluencies have been studied, but particularly because of the agglutinative nature of Basque (Hirschberg & Nakatani, 1993; Stolcke & Shriberg, 1996).

As one might expect, in a number of ways the occurrence of fillers in Basque is comparable to the occurrence of fillers in previously examined languages. For example, as in previous work (e.g. Clark & Wasow, 1998), we found that fillers were produced disproportionately often before content words compared to function words. We also found that Basque speakers provide listeners with strong probabilistic cues by virtue of their selection of particular fillers. That is, just as Clark and Fox Tree (2002) reported that speakers may use *uh* to signal a minor delay and *um* to signal a major delay, when our speakers used /e/ or /m/ as a filler, listeners could have a strong expectation that a new sentence would follow because these fillers primarily occurred intersententially. In contrast, if one of the other fillers was produced, listeners should expect the sentence to continue because these other fillers were much more likely to be produced intrasententially. In this respect it seems that /e/ and /m/ are associated with planning complex syntactic units, whereas the other fillers get generated when encountering difficulty with lexical access.

Together with these commonalities with filler use seen in previous studies, we also observed a number of differences. At a global level, the filler rate here (approximately 1.7% of the words in the corpus) is lower than Fox Tree's (1995) estimate of filler occurrence (up to 6%). While there could be many sources for this difference, we speculate that it might reflect the relaxed and unpressured setting that we deliberately created for our speakers; fillers should occur at a higher rate when the speaker is trying to produce speech at a high rate. Our speakers were older than those typically recorded in previous work, a second factor that would generally yield slower speaking rates. We suspect that our older speakers may also tend to use vowel lengthening as a filler more often than younger speakers do, consistent with a slower speaking rate and the relatively high rate of vowel lengthening we observed. Vowel lengthening was by far the most common type of disfluency in our sample, more common than in most other studies (cf. Fox Tree & Clark, 1997). Conversely, the filler /m/ was quite rare in our sample, comprising only about 3% of the observed filler tokens. Prior studies of speech disfluencies (e.g. Clark & Fox Tree, 2002; Corley, 2008; O'Connell et al., 2005) had reported higher occurrence rates of /m/ (and *um*). This difference likely reflects the low frequency of occurrence of nasals in Basque (Corley & Stewart, 2008; O'Conell & Kowal, 2005).

We believe the most interesting difference between previous observations and those in our corpus involve the filler /zera/. We noted at the outset that, in principle, the agglutinative nature of Basque offered speakers a mechanism that could be used to provide listeners with partial information about an upcoming word when the speaker's access to that word was only partially complete. Our corpus provides a large set of examples demonstrating that Basque speakers take advantage of this mechanism: The filler /zera/ was the third most common filler used by our speakers, and its surface form often reflected the syntactic properties of an only partially accessed following word. /zera/ was almost always used intrasententially, consistent with its connection to a particular word

that was in the process of being accessed. The partial results of that lexical access can be signaled using the agglutinative properties of Basque, leading to the range of surface forms in our corpus. Interestingly, this process is apparently quite efficient in the sense that the marked forms of /zera/ were produced more fluently than the other filler forms – the total time involved in producing the filler, including any preceding and following pause, was half as long for /zera/ than for the other fillers.

As we noted above, there is some very preliminary evidence that Turkish might employ a filler that has some of the properties that we have found for /zera/, but there is not yet enough data to be sure. As such, we believe that in future research it would be quite informative to examine the nature of fillers in agglutinative languages to look for fillers that take on specific forms that reflect the syntactic roles of words that have been almost, but not quite, accessed from the lexicon. Disfluencies play a crucial role in signaling the difficulties and flow of spoken language. They are a consequence of language mechanisms that can provide a particularly informative window into the process of producing spoken language.

## Acknowledgements

We would like to express our appreciation to Susan Brennan and to Marie Huffman for their valuable and constructive suggestions on an earlier version of this paper. We also thank two anonymous reviewers for their helpful comments.

## Funding

This research received support from grant PSI2010-17781 from the Spanish government's Ministerio de Economia y Competividad.

## References

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: disfluency and reference resolution. *Psychological science*, *15*, 578–582. doi:10.1111/j.0956–7976.2004.00723.

Bailey, K. G., & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language*, *49*, 183–200.

Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In S. Santi, I. Guaïtella, C. Cave, & G. Konopczynski (Eds.), *Oralité et gestualité, communication multimodale, intéraction* (pp. 597–600). Paris, France: L'Harmattan.

Bear, J., Dowding, J., & Shriberg, E. (1992). Integrating multiple knowledge sources for detection and correction of repairs in human-computer dialog. In *Proceedings of the 30th annual meeting* of the Association for Computational Linguistics (pp. 56–63). Newark DE.

Belikova, A., & White, L. (2009). Evidence for the fundamental difference hypothesis or not? Island constraints revisited. *Studies in Second Language Acquisition*, *31*, 199–223.

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech*, *44*, 123–147.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, *44*, 274–296.

Candea, M., Vasilescu, I., & Adda-Decker, M. (2005). Inter- and intra-language acoustic analysis of autonomous fillers. In *Proceedings of DISS 05, Disfluency in spontaneous speech workshop* (pp. 47–52).

Clark, H. H. (1994). Managing problems in speaking. *Speech communication*, 15(3), 243–250.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111.

Clark, H. H, & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201–242. doi:10.1006/cogp.1998.0693

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: hesitations in speech affect language comprehension. *Cognition, 105*, 658–668. doi:10.1016/j.cognition.2006.10.010

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, *2*, 589–602.

Crystal, D. (1992). *Introducing Linguistics*. Harlow, UK: Penguin.

De Leeuw, E. (2007). Hesitation markers in English, German, and Dutch. *Journal of Germanic Linguistics*, *19*, 85–114.

Fehringer, C., & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica*, *41*, 37–72.

Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *34*, 709–738.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, *29*, 320–326.

Fox Tree, J. E, & Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition*, *62*, 151–167.

Fox Tree, J. E., & Schrock, J.C. (2002). Basic meanings of *you know* and I mean. *Journal of Pragmatics*, *34*, 727–747.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516–522.

Gosy, M. (2007). Alternative routes of lexical retrieval in spontaneous speech. *Studia Slavica Hung*, 52 (1–2), 129–136.

Hirschberg, J., & Nakatani, C. (1993). A speech-first model for repair identification in spoken language systems. In *Proceedings of the third European conference on speech communication and technology*.

Künzel, H. J. (1997). Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics*, *4*, 48–83.

Lendvai, P., Van Der Bosch, A., Krahmer, E., & Canisius, S. (2004). Memory-based robust interpretation of recognised speech. In *SPECOM 2004: 9th conference speech and computer*, St. Petersburg, Russia, September 20–24, 2004.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*, 19–44.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, *22*, 473–500.

O'Connell, D. C., & Kowal, S. (2005). Uh and um revisited: are they interjections for signaling delay? *Journal of Psycholinguistic Research, 34*, 555–576. doi:10.1007/s10936–005–9164–3.

Schachter, S., Christenfield, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, *60*, 362–367.

Shriberg, E. (1994). Preliminaries to a theory of speech disfluencies. Doctoral dissertation, University of California, Berkeley, CA.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*, 25–25.

Stolcke, A., & Shriberg, E. (1996). Statistical language modeling for speech disfluencies. In *Proceedings of international conference on acoustics, speech and signal processing* (pp. 405–409), Atlanta, GA.

Watanabe, M., Den, Y., Hirose, K., & Minematsu, N. (2004). Clause types and filed pauses in Japanese spontaneous monologues. In *INTERSPEECH 2004*, (pp. 2981–1984).

Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication*, *50*, 81–94.

**Table 3.** Filler durations measured using PRAAT software.

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| /e/ | | | |
| *beak esaten zizten eee, pues esenii*<br>they told me eee, so the milk… | 0.4 | 0.61 | 0.07 |
| *ordun gauzei eee kentzen*<br>then removing eee to the things | 1.08 | 0.41 | 0.1 |
| *Zoo eeee musen*<br>You eee at mus | 0.47 | 0.32 | 0.06 |
| *Ez, urrungoak eee Ezkurretik asten zeen*<br>No, the next ones eee started from Ezkurra | 0.21 | 0.75 | 0 |
| *bat eelle zen eee Martina*<br>one walked eee Martina | 0 | 0.88 | 0.62 |
| *Una gozada terrible, eeee ez, ez…*<br>It is a terrible pleasure, eeee no, no… | 0.89 | 0.56 | 0 |
| *Bai eeeeeeee baldin bazittun xaldik eo olaaa*<br>Yes eee if he/she had horses or so | 0.31 | 0.76 | 0.06 |
| *ta ze ejn bear zinun? Eee besteatzuk eelliko zeen*<br>What were you going to do? Eee others will do… | 0.63 | 0.27 | 0.26 |
| *Bai biño eeeee larraintzarra*<br>Yes but eee to Larraintzar | 0.13 | 0.6 | 0 |
| *pixkat joan nintzen eee leitzea*<br>I went a little bit eee to Leitza | 0. | 0.53 | 0.1 |
| *Igoatik eta, emengo neskak,eee igoakoak eta bewtekoak eta*<br>From Igoa and, girls from here, eee from Igoa and Beruete and | 0.8 | 1.25 | 0 |
| *azkenin iunes eee etten dozu*<br>In the end during the day eee you do | 0 | 0.67 | 0 |
| *beak akostum eeee gauza asko*<br>They us-(used to)eee a lot of things | 0.23 | 0.71 | 0 |
| *besteati etzaio eeee ain e nabari*<br>it isn't so evident eee to another one | 0 | 0.72 | 0 |
| *biñooo eeee ne ama*<br>but eeee my mother | 0.16 | 0.46 | 0.13 |
| (Question) *Eee a ver, eee, goajn ez zajt bururatu bear gañea 1*<br>(Question) eee, let's see, eee, I'm not going to remember it right now… | 0.6 | 1.17 | 0 |
| (Question) *Eee a ver, eee, goain ez zait bururatu bear gañea 2*<br>(Question) eee, let's see, eee, I'm not going to remember it right now… | 2.67 | 0.66 | 0.47 |
| *Ta tiendak ee bai, supermerkado aundik*<br>And shops eee yes, big markets | 0.47 | 0.5 | 0.11 |
| (Question) *eeeee eztakit nola esaten den*<br>(Question) eeeee I don't know how to say it | 1.99 | 0.55 | 0.33 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| (Question) *eeee ne bai*<br>(Question) eeee me yes | 0.93 | 0.75 | 0 |
| *Eee, ez, eee bajno ojk dee manic*<br>Eee, no, eee but still those are whims | 0 | 0.44 | 0.45 |
| *Eee, ez, eee bajno ojk dee manic*<br>Eee, no, eee but still those are whims | 0.81 | 0.93 | 0.81 |
| *kostatzen zakiten asko eeee eskreitzii*<br>It took me a while eee to write | 1.22 | 0.26 | 0.6 |
| (Question) *eeee nek hitzeitten dut asko ta rapido*<br>(Question) eeee I speak a lot and fast | 1.65 | 0.8 | 0.27 |
| (Question) *Eee nek baj, ne eta besti*<br>(Question) eee me yes, me and the other | 1.25 | 0.92 | 0.14 |
| *ta karo eeeeeeeeeee ez jakinke*<br>of course eeeee without knowing | 0 | 1.4 | 0 |
| *Emen ala etten zen, etten ginuzen þues eee Matematicas, Naturales, Sociales*<br>Here we did that way, we used to do well eeee maths, nature, social science | 0.06 | 2.62 | 0.3 |
| *geo atzo eon nintzen eee nekin ikasi zun mutil batekin*<br>then yesterday I met eee a boy who learned with me | 0.45 | 0.54 | 0.58 |
| *ez gañea nik eeee gañea*<br>no moreover I eeee moreover | 0.12 | 0.3 | 0.12 |
| *ez ez eee aziendatik bizianik ez ta*<br>no no eee there is nobody who lives from his/her own animals | 0.25 | 0.37 | 0 |
| *eztiote batua batua etten e eeeeeee nerii*<br>They don't speak batua batua eeeee to meee | 0.4 | 0.65 | 0.12 |
| *gauza trastek esaten ttu baino eeee baino*<br>He/she says bad things but eee but | 0 | 0.4 | 0 |
| *Guk len bordatxo zenen eeeee bordatxora*<br>When it was "Bordatxo" we eeee to "Bordatxo" | 0.21 | 0.38 | 0 |
| *hamar ume eeeee bai bai*<br>Ten children eeee yes yes | 0.56 | 0.63 | 0.07 |
| *þues hemen eeeee beste*<br>well here eeee other | 0.48 | 0.76 | 0.27 |
| *þues eeee herri batetik besteaaaa*<br>well eeee from one town to another | 0.08 | 0.45 | 0.09 |
| (Question) *eeeeeee hitzek eta asko sartzen ttut*<br>(Question) eeee I introduce a lot of words | 0.6 | 1.05 | 0 |
| *ne? Eeeee genealin*<br>Me? Eeee Normaly | 1.61 | 0.88 | 1.44 |
| *dute hor zenotzen eeeee ta ordun*<br>They have right there at Zenotz eeee and then | 0.5 | 0.46 | 0 |
| *ttikii eeeeee ttiki bai*<br>little eeee little yes | 0.07 | 0.4 | 0.05 |
| *zaldik or eeee ein dute hipika*<br>Horses right there eeee have done equestrian sports | 0.02 | 0.45 | 0.75 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *ez zeatik eeeee esan nahi dut*<br>No because eee I mean | 0 | 0.91 | 0 |
| *zertan izan eeeee eskolan hasi ginen*<br>At what eeee we started at school | 0.12 | 0.33 | 1.1 |
| *Bueno hoi eeeee esan duten bezala*<br>Anyway this eeee as I have said before | 0 | 0.76 | 0.2 |
| *Gainera pues eee amuarraia*<br>Furthermore well eee the trout | 0.73 | 0.43 | 0 |
| *ia claro eee argi: dago sartzen den bezela*<br>of course eee the light is as it enters | 0 | 0.65 | 0.76 |
| *aundiak eeeee egunean behin*<br>the big ones eeee once a day | 1.05 | 0.75 | 1.02 |
| (Question) *eeeeee pues egiten da*<br>(Question) eeeeee well it is done | 0.3 | 0.84 | 0 |
| *pues eeeeeee garbiketa*<br>well eeeeeee the washing | 0 | 0.26 | 0.54 |
| *Eskalera bat bezala eeeee goian*<br>Like a stair eeee above | 0.47 | 0.67 | 0.73 |
| *eta hoiek eeee astin esan duten bezala*<br>and they eeee as I have said before | 0.52 | 0.92 | 0 |
| *Ez ez ez, gu: gara pues eeee, guk etten dugu txiki guztia ez?*<br>No no no, we, are well eeee, we do all the little things don't we? | 0.45 | 0.78 | 0.35 |
| *pues eeeeee kriaderoan*<br>well eeeeee at the hatchery | 0.6 | 1.06 | 0 |
| *ta fario hori da pues eeeee errekan errekako amuarraia bezalakoa ez?*<br>And this is well eeeeee it is like the trout from the river – at the river | 0.54 | 0.67 | 0.89 |
| *ta hoik pues eeee ez dira berdin asetzen*<br>And they well eeee don't satisfy identically | 0 | 0.88 | 0.44 |
| *yesara eee batzuk*<br>To Yesa eee some | 0.34 | 0.46 | 0 |
| *Eeee, batzuek baj, andienak baj, gero ttikiak joaten dira pues, Riezu*<br>Eeee, some of them yes, the biggest ones yes, then the little ones go, well, to Riezu… | 0 | 0.34 | 0.44 |
| *Eee ez Ezkurretek eta ez ziren etortzen*<br>Eee no, they didn't come from Ezkurra | 0 | 0.61 | 0.52 |
| *Baj eee baj, ordun ezen bestejk ardoa ta:, zerbeza e izaten zen*<br>Yes eee Yes, some time ago there wasn't anything but wine, also there was beer | 0.16 | 0.66 | 0.43 |
| *Geo: zeara ere bai eeee, Maier…*<br>Then to that as well eeee, Maier… | 0.22 | 0.53 | 0.44 |
| *Ez ez, a-eeee aziendetatik bizi denik ezta*<br>No no,a-eee There is nobody who lives from the animals | 0 | 0.33 | 0.38 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Eeee, ne lawne ttikii den oj ez da allatzen tetxora*<br>Eeee, my little friend doesn't reach the ceilling | 0 | 0.66 | 0.39 |
| *Zalditan eeee batzuk entenditzewte*<br>Some are experts eeee on horses | 0.23 | 0.44 | 0.54 |
| *eeee ez, batzutan igual baj*<br>eeee no, sometimes maybe yes | 0 | 0.46 | 0.33 |
| *azkenekoiken eee iikiuten eskeostik.*<br>One of the last ones eeee since they opened it | 0.16 | 0.53 | 0.37 |
| *ez eeeeee erdeara pastu behar*<br>no eeeee it must change to another language | 0.12 | 0.61 | 0 |
| Median values for /e/<br>/Vowel-lengthening/ | 0.3 | 0.66 | 0.09 |
| *Iruñin daen tienda koxkorraik ereeee, danak ez bearkoute ertsi*<br>Those big shops that are in Pamplona as well, they mustn't close | - | 0.3 | 0.42 |
| *Bai barraaaaa barrra*<br>Yes exhausted exhausted | | 0.7 | 0.14 |
| *apartekoa* (…) *nombaitte, eztooo ez eztawke martxik*<br>apparently, there isn't, it isn't working | | 0.95 | 0.39 |
| *Bai, ta geooooo besteat*<br>Yes, and then another | | 0.48 | o.5 |
| *Oso rato majoak zeen aik, gaur bañooo diruu-bate etzen ordun*<br>They were great times, there wasn't as much money as we have today | | 0.74 | 1.62 |
| *oso ona, baniooo "cuidao" arrek zer edo zer esaten zunin*<br>so good, but be careful whenever she/he told something | | 0.83 | 0.14 |
| *Gu etxiin, amak gejna*<br>In our home, our mother the most | | 0.67 | 0.29 |
| *utzi zeen, baino, ordun meza eta bezpeetaaa ta, zorrotz*<br>they left, but, then the mass and the vespers and, strict | | 0.53 | 0.7 |
| *hogei hogeitabi urtekoak beste ilaran, bakoitxaaa be zean joaten zen*<br>People of 20–22 years old in the other row, each went in their own "that" | | 0.99 | 0.25 |
| *Lenao neska ta mutil bat beaaaa konpaaziora*<br>Some years ago a boy and a girl in comparison…. | | 0.49 | 0.38 |
| *Ta geo etten zen ikatzaaa lurreee lurreee lurreiken*<br>And then it was made the coal with ground | | 0.8 | 1.065 |
| *Subentzioak eta badaijj beñooo or, or printzipala lana*<br>There are foundings but the job is the first | | 0.4 | 0 |
| (1) *Oain baittuuu, oitaaa oitaemezortzi urte ooo*<br>(1) Now he/she has, 38 years or so | | 0.41 | 0.91 |
| (2) *Oain baittuuu, oitaaa oitaemezortzi urte ooo*<br>(2) Now he/she has, 38 years or so | | 0.3 | 0.85 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *oain delaaaaaa hogeitamar urte*<br>30 years ago | | 0.4 | 2.05 |
| (1) *Ez, ez, besteee eztajit nongoaen, besteee apartekoa nombajtte*<br>(1) no, no, I don't know were is he from, from another site apparently | | 0.65 | 0 |
| (2) *Ez, ez, besteee eztajit nongoaen, besteee apartekoa (…) nombajtte*<br>(2) no, no, I don't know were is he from, from another site apparently | | 0.6 | 0.36 |
| *Oi zen lenao felipeeeen anaiena*<br>It was formerly from Felipe's brother | | 0.99 | 0.12 |
| *Ta olaaaaa, ola joan zen*<br>And it went that way | | 0.43 | 0.2 |
| *Oixe bai, eeeesan dizutena*<br>Indeed, what I have told you | | 1.11 | 0 |
| *Majo paatzen zun, ondarrekooo ondarreko eoten zana*<br>He paid a lot, for the rests | | 0.78 | 1.5 |
| *Ta, eskeee ganaderoak ze esaten do?*<br>And, well what do farmers say? | | 0.64 | 0.41 |
| *Batek Iruñin ta bestiiii, or badaaaaai*<br>One in Iruña and the other, is somewhere around | | 1.2 | 2.66 |
| *fuera! Eskeee, ala baita*<br>Get off! Well, that's the case | | 0.6 | 0.16 |
| (1) *Oajn delaaa…oajn dela berrogeiiiii… berrogeita bost bat urte*<br>(1) Forty–forty-five years ago | | 0.65 | 1.86 |
| (2) *Oajn delaaa…oajn dela berrogeiiiii… berrogeita bost bat urte*<br>(2) Forty–forty-five years ago | | 0.89 | 1.4 |
| *Bajna bueno eee ongi dao etxetik oiiii joan eta etorri*<br>But well eee he is ok coming and going at home | | 0.87 | 1.49 |
| *Etzen oiiiii, arek etzun segitu*<br>It wasn't this, he didn't continue | | 0.66 | 0.56 |
| (1) *Hoi osoooo gauze onauk osooooooo ona*<br>(1) This is such a good thing, very good | | 0.4 | 0.34 |
| (2) *Hoi osoooo gauze onauk osooooooo ona*<br>(2) This is such a good thing, very good | | 0.49 | 0.64 |
| *apartekoa* (…) *nombaitte, Eztooo ez eztauke martxik*<br>apparently, there isn't, it isn't working | | 0.91 | 0.4 |
| *Pa SS o pa irun ooooooooo eonora*<br>To San Sebastian or to Irun or wherever | | 1.44 | 0.25 |
| *Oain jendek pezta botzekooo lanak*<br>Nowadays people hardly give money | | 0.78 | 0.68 |
| *Pixkittoat geixooooooo badiote*<br>If they say a little bit more | | 1.37 | 0.09 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Ezkurre, Saldis, baaaaaa Leitze*<br>Ezkurra, Salias, well Leitza | | 0.46 | 1.15 |
| *Ok salduu ooo desaparezittu*<br>Sell or disapear | | 0.41 | 0.43 |
| *Ta oain haik ateaaa taaaaaa*<br>And now they get off and | | 0.61 | 0.16 |
| *Ta olakooooooo jokadatzuk izaten zeen*<br>And there were similar actions | | 0.8 | 0.07 |
| *Ta Tolosaaaaaaan, Tolosan*<br>And in Tolosa, in Tolosa | | 1 | 0.44 |
| *Xaldik taaa giooooooo olaaaa* (question)<br>Horses and then that way | | 0.72 | 0.12 |
| *Txendak eta artzen ttute bainooo, txendak ee utzii ta*<br>They go through pathways but they also leave pathways | | 0.46 | 0.05 |
| *Zalditan ereeee batzuk entenditzewte*<br>Some are experts on horses | | 0.74 | 0.22 |
| *Baj, ze usteozooo? Komedii izaten zen ordun*<br>Yes, what do you think? It was not easy | | 1.39 | 0.07 |
| *Bate! Guuu, guk etten degun lanaj*<br>No one! We, the work we do | | 0.57 | 0.2 |
| (Question) *Eeeeeeeske ni bete erdeas*<br>(Question) well I always speak Spanish | | 0.46 | 0.06 |
| *eta ja etten dozooo besteatzuki bejs etten duzu ewskeas*<br>And I speak Basque with others | | 0.45 | 0.35 |
| *saltzen dizkizu þues arraultziiik, bolloak, makarronak*<br>He sells eggs, bread, macaroons, | | 1.57 | 0.07 |
| *Baj… biñooo… emeeen ferii þrintziþala etten zena*<br>Yes… but… but here the principal showing that was made | | 0.61 | 1.4 |
| *Bai, biñooo bai, onge*<br>Yes, but yes, well | | 0.8 | 0.1 |
| *mantsoagoan asten naz ordun ta batuagoan, bajnooo genealin neri re*<br>The slower I speak the more I use Basque, but in general me too. | | 0.6 | 0.47 |
| *Bajno bueno, bakoitxaaaak ee esan naj dut*<br>Yes well, each one as well I mean | | 1.27 | 0.2 |
| *Ezta difeentzi andiik izaan bainooo bai nabaitzen da difeentzi bai*<br>There isn't too much difference but it is noticeable | | 0.85 | 0.26 |
| *beste aldea ejn bearkozu zeorreee moldatu*<br>You have to turn to the other side, do it by your own | | 0.65 | 0 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *subentzjoak eta badajjj, beñooo, or printzipala lana*<br>There are foundings but the job is the first | | 0.47 | 0.07 |
| *difeentziik andina emendik igoara eonen dela, beñooo gu goartzen ga*<br>That the biggest difference will be from here to Igoa, but we are aware of it | | 0.54 | 0.32 |
| *Eztakit nola nola esatewte ez? Beñooo askotan ezteket*<br>I don't know how – how do they say, no? But, most of the times I don't know | | 0.7 | 0.3 |
| *Ne irujttun baj, gajnea oajn ala itzejttewte, beñooo ala izanen da igual*<br>From my point of view yes, moreover nowadays they speak that way, but, it will probably be that way | | 0.51 | 1.51 |
| *sartzen izaan ttut itzek eta ola beñooo baj itten dot*<br>I use to introduce words and so but yes I use to do it | | 1.66 | 0 |
| *besteat ematen zizuten ta, itxen amaiiiii bere gajn e diruu*<br>They gave you another and, our mother controlled money as well | | 0.68 | 0.18 |
| *gue etsairik haundina, eeeesan nahi dut, etsaie*<br>Our biggest enemy, I mean, the enemy | | 0.58 | 0 |
| *Bueno, eeeeeenteintzen baldin badezu*<br>Well, if you understand | | 0.53 | 0 |
| *eo iwn bat bakarrik eooo eztakit nola etten izaan duten*<br>Or just one day or I don't know how do they do it | | 0.81 | 0.08 |
| *Orren kostunbrii artu dozu eta ja etten dozoooo besteatzuki bejs*<br>You get used of it and you do the same with other people | | 0.99 | 0 |
| *Biñooo eztaaaaaa, ezta*<br>But neither, neither | | 1.17 | 0 |
| *Biño geooo ta gutxio*<br>But increasingly less | | 0.75 | 0.17 |
| *igual geooooo etorriko dee*<br>maybe they will come later | | 0.6 | 0.5 |
| *guartuko leikeeeeeee bai*<br>He/she could realize | | 0.67 | 0.44 |
| (Question) *neeeeee auxiliar administrative*<br>(Question) Meeee administrative assistant | | 0.77 | 0.04 |
| *Oi jon zeeen, azkenekoa joan zen Larrajntzarra bizitzea*<br>He/She went, the last one went to live to Larraintzar | | 0.75 | 0.42 |
| (1) *Orrek ee beti ematen dizuuu, beee, bado be gawze*<br>(1) That gives you, it's it has something | | 0.57 | 0.48 |
| (2) *Orrek ee beti ematen dizuuu, beee, bado be gawze*<br>(2) That gives you, it's it has something | | 0.52 | 0.44 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *ta geooooo astin bi aldis*<br>And then twice per week | | 0.79 | 1.07 |
| *etten duzu moztu ta, asten zaaa, asten za*<br>You cut it and, you start, you start | | 0.58 | 0.32 |
| *Baj, eta ordun areaaa, area Joan zen*<br>Yes, and then there, he/she went there | | 0.53 | 0.33 |
| *tokatu izandu najz joaten eh, bainooo, ez ko-nik ez dut kostunbreik*<br>I have had to go eh, but, I am not used to | | 0.43 | 0.56 |
| *oain udan asten da au eladukin eta giooo, panaderoa etortzen dena eunero*<br>Right now in summer this one starts working with ice cream and then the baker that comes every day | | 1.35 | 0 |
| *etten giñuzen Matematicaaaaaaas natura*<br>We studied maths, nature | | 0.8 | 0.2 |
| *Buenooo, saltzekooo, artu bear dute*<br>Well, to sell, they have to get | | 0.6 | 0 |
| *Bueno enpresa honeeeek beak*<br>Well this company also | | 1.27 | 0.26 |
| *baino ori da pues oriiii erabiltzen dute errepoblatzeko erreka*<br>But this is well they use it to restock the river | | 0.47 | 0.42 |
| *Baj baj gertatu da kasi urtero jendeak erretzen du erreka, errekaaa ombre badira beti*<br>Yes it has occurred that every year somebody burns the river, the river well there are always | | 0.77 | 0.48 |
| *Ez daaa ez da negozio familiarra bajna familia guztia or gawde lanean baj*<br>This is not a familiar company but we work all the family together | | 0.75 | 0 |
| *ez ez guuuuuk beno*<br>no no we well | | 0.47 | 0.32 |
| *Ta geroooooo harina de pescau*<br>And then fish flour | | 0.77 | 0.3 |
| *Ta holaxeeeeeee* (Question)<br>And that is (Question) | | 0.43 | 0.6 |
| *gero zeara ere baiiii Iraizotza*<br>And then to that as well, to Iraizotz | | 0.45 | 0.22 |
| *Xaldik taaa giooooooo olaaaa* (Question)<br>Horses and then that way | | 0.44 | 0.43 |
| *Ta, noa yoan dee gasolle, noa yoan dee beste gañeako gauzek, dena:… bua… gora*<br>And, where has gone the petrol, where have gone the rest of the stuff, all… bua… up | | 0.47 | 0.51 |
| *Aw, ganaderoa ola:… olaxe aj da e, olaxe aj da*<br>This one, the farmer that way… is doing that way e, he is doing that way | | 0.39 | 0.61 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Or denetatik e:otentzeen, a:rdoa, olioa, ta azukarik eta naran[x]ak*<br>They had everything, wine, oil, sugar and oranges | | 0.44 | 0.51 |
| *Or denetatik e:otentzeen, a:rdoa, olioa, ta azukarik eta naran[x]ak*<br>They had everything, wine, oil, sugar and oranges | | 0.56 | 0.45 |
| *Aaa, baj, bajno eztao, eztao, eztao:… eztao dirujk artzejk*<br>Aaa, yes, but there is not, there is not, there is not… there is not possibility to take money | | 0.46 | 0.34 |
| *Gawzek ola dizte, eztawke gawze:… sujeto eztao*<br>Things go that way, it is not subject to | | 0.39 | 0.56 |
| *mendik, azendak, eztute:… eztao, eztao, eztao*<br>The mountains, the animals, they haven't… there isn't, there isn't, there isn't | | 0.36 | 0.61 |
| *Besteatzuk: Bejko bentaajno*<br>Others till the lower house | | 0.54 | 0.77 |
| *Baj..biñooo..emeeen ferii þrintziþala etten zena*<br>Yes… but… but here the principal showing that was made | | 0.58 | 0.61 |
| *Ta arrek: beraa, karri onea denak*<br>And that one, down, bring here everything | | 0.46 | 0.2 |
| *Ne: ango etxiartakoa nun nek atautxi Joan joxee*<br>Meee my grandfather Juan Joxe was born at that house | | 0.36 | 0.44 |
| *Geo itxe bakojtza, baserrii::, baserri bakojtzan zortzi ume*<br>Then each house, the farm house, eight children per farm house | | 0.4 | 0.47 |
| *esejni falta bazajzkizu gawza ojtatik, ta geo:: astin bi aldis etortzea þeskadu*<br>If you have lack of milk, and then the fisherman cames twice a week | | 0.7 | 0.66 |
| *Biño bueno, be: been modun kanþotik jendii etorri¡*<br>But well, people come from other sites as he did | | 0.5 | 0.47 |
| *Eske:, balitzeke tabernaa joaten zala*<br>Well, if he went to the bar | | 0.35 | 0.23 |
| *Gero Yesan ere egiten dute txiki þixka bat baino: ez Arrarasen hobe*<br>Then they do in Yesa some litle stuff but better in Arrarats | | 0.62 | 0.44 |
| *Baj baj gertatu da kasi urtero jendeak erretzen du erreka, errekaaa ombre: badira beti*<br>Yes it has occurred that every year somebody burns the river, the river well there are always | | 0.67 | 0.58 |
| *Da:: mm:: baj, no:r zengatik ezawtzen duzun eo zergatik ajtzen zan eo ola*<br>And mm yes, who do you know why or why did he/she did it or so | | 0.48 | 0.28 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Da:: mm:: baj, no:r zengatik ezawtzen duzun eo zergatik ajtzen zane o ola*<br>And mm yes, who do you know why or why did he/she did it or so | | 0.46 | 0.21 |
| *Geo: zeara ere bai e::, Iraizotza*<br>Then to that as well, to Iraizotz | | 0.33 | 0.3 |
| *erreko luke:: erreka guztia*<br>He would burn the whole river | | 0.35 | 0.22 |
| *eta gero vitamina::k produktu asko ditu*<br>And then the vitamins have a lot of products | | 0.67 | 0.54 |
| *A:: egurran eske? San Juan bezperan joaten ziren.*<br>To get wood? They went at San Juan's eve | | 0.81 | 0.67 |
| Median values for /vowel-lengthening/<br>/Zera/ | | 0.56 | 0.33 |
| *Behi nabarrak esnea saltz-zeatzekoak eurtzunen*<br>Motley cows the ones for sal(e)-that the milk in Irurtzun | 0.45 | 0.56 | 0.22 |
| *Beko zea hartan*<br>In the lower that | 0 | 0.15 | 0 |
| *Elizan kontra PEGANTE zea dao*<br>It is close to the church that it is | 0 | 0.32 | 0 |
| *Juan Joxe zeaaaaa, BASERRII*<br>Juan Jose that, the farm | 0 | 0.53 | 0.67 |
| *Leioak beak zeaaa ZURESKOAK*<br>The windows are made of that of wood | 0.06 | 0.42 | 0.24 |
| *Oain eztao zeaik eh*<br>Now there isn't that eh | 0 | 0.25 | 0.06 |
| *Oi zea da, Aimar þilotarin TIO*<br>He is that, uncle of the pelotari Aimar | 0 | 0.3 | 0 |
| *Ultzamako zea hortan egoten da*<br>He uses to be at that of Ultzama | 0 | 0.18 | 0 |
| *Ta hor zea, ABEATSA da jendigo ta*<br>An there that, millionaire is people | 0.15 | 0.57 | 0.12 |
| *Ta zea ttu-IRUÑEKOAK ttuk*<br>And that ar-from Pamplona are | 0 | 0.24 | 0 |
| *An eoten ginen beiko zea artan*<br>We used to be in the lower that | 0 | 0.19 | 0 |
| *Ab-zea BENDATU egiten ginun*<br>Ab-that we used to have tea | 0.18 | 0.32 | 0 |
| *Geo etorri zen zeaaaaaa BALSII*<br>Then it came that, Vals | 0.45 | 0.57 | 0.51 |
| *musu bat eta beste e guk ematen ginun bejn o beste ikaatzen bajzen, zea, orduun ikaatzen bajzen*<br>Sometimes we gave them a kiss, that, then they were scared | 0.39 | 0.5 | 0 |
| *Lenao etten zituzten gauzek zeagoak, trastegoak*<br>Previously they did things that, worse | 0.2 | 0.5 | 0.97 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Bakoitxa be zean joaten zen*<br>Each one went on his that | 0 | 0.25 | 0 |
| *Zuuk etzeate zeatuko baño*<br>You are not going to that but | 0 | 0.4 | 0 |
| *gero zeara ere baiiii eee IRAITZOTZA*<br>Then as well to that to Iraizotz | 0.3 | 0.34 | 0 |
| *denak, denak, zeakoak RAZAKOAK*<br>All, all of that, of race | 0 | 0.43 | 0 |
| *goardik hemen zeuden zeatzeko goardik*<br>Guards were here to that, guards | 0 | 0.35 | 0.45 |
| *Jendea kokatzeko zean KOLEKZIOAN*<br>To situate people, in that, in the collection | 0.18 | 0.25 | 0.32 |
| *oain plata zea ahaztu ein zait*<br>Now I the "plata", that, have forgotten | 0 | 0.24 | 0 |
| *ta geo zea NESKA gaztik hasi zeen*<br>And then, that, young girls started | 0 | 0.18 | 0 |
| *ta nik emandako zeak orduun*<br>And the that I gave then | 0 | 0.3 | 0.13 |
| *zea a ver den txekorra kobratu ziona*<br>That let's see if it's the lamb what he/she has charged | 0 | 0.24 | 0.15 |
| (Question) *zean aitzen da, PRIMAVERAN aitzen da*<br>(question) in that she does, in spring she does | 0.2 | 0.18 | 0 |
| (Question) *zea egiten omen zen*<br>(question) that was probably done | 0 | 0.32 | 0 |
| (Question) *Zea KLARA*<br>(question) that, Klara | 0 | 0.48 | 0 |
| (Question) *Zea,TXIKII nintzelakos*<br>(question) Because that, because I was little | 0 | 0.5 | 0.32 |
| (Question) *Ta, zea nor izan zen?*<br>(question) and, that, who was? | 0.32 | 0.44 | 0.45 |
| (Question) *zeakoa YESAKO Juan jose*<br>(Question) Juan Jose from that, form Yesa | 0 | 0.25 | 0 |
| (Question) *Zeakoak, GARTZAONDARRAK*<br>(Question) From that, from Gartzaron | 0.15 | 0.46 | 0 |
| (Question) *Zean, BARTZELONAN*<br>(question) In that, in Barcelona | 0 | 0.33 | 0 |
| *amuarraia zeatzeko ta zeatzeko eeltzen zeen*<br>It was used to that and that the trout | 0.23 | 0.87 | 0 |
| *ta amuarraia zeatzeko egiten zen*<br>and it was used to that the trout | 0.14 | 0.45 | 0 |
| *Zurezko zera batzuk…KARRAKAK*<br>Some wooden thats… rattles | 0 | 0.21 | 0 |
| *Orduan koheterik ez ziren izanen, zean, HERRIAN*<br>Then there weren't going to have rockets, in that, in the town | 0.41 | 0.25 | 0.35 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *lenbixko urtetan behiekin ta, zeatzen, behiei KASUITTEN*<br>During the first years with cows and, doing that, taking care of cows | 0.31 | 0.32 | 0.41 |
| *ongi tratatzen tzieten gañea, zea bazkaltzea*<br>They treated them well, furthermore that, to eat | 0.23 | 0.16 | 0 |
| *jende asko joaten zen zea hoitaa, KANPEONATO hoitaa*<br>There was a lot of people going to that, to the competition | 0 | 0.19 | 0 |
| *Pues, maíz zeatzen zittuten, PRESTATZEN tzittuten*<br>Well, they that frequently, they prepared | 0 | 0.55 | 0 |
| *beti izaten da zea þixka bat,*<br>There is always a litle bit of that | 0 | 0.17 | 0 |
| *Geo zea haundikoa, FAMA*<br>Then of big that, fame | 0 | 0.18 | 0 |
| *ya zea, beti bazien bertsolari onak*<br>already that, there were always good singers | 0.12 | 0.21 | 0.33 |
| *Fama zeakua zen ERRAZKINDARRA re*<br>The fame was from that, from Errazkin as well | 0 | 0.43 | 0 |
| *Claro, beti izaten da zea mouko bat, baño ikaratu beñez, inpresionatu beñez.*<br>Of course, there is always some sort of that, but they never get scared | 0 | 0.19 | 0 |
| *ya gutxi aittu zienak eta zea, biño beti*<br>Few people was active and that, but always | 0 | 0.2 | 0.3 |
| *Bai, bai hala uste dut, claro, oaiñ zeatugo da eh? ILUSTRATUO eo gehio ikasik*<br>Yes, yes I think that way, of course, now is more that isn't it? It's more illustrated or more learnt | 0 | 0.45 | 0 |
| *nik aittu ttuten zeatan, KANPEONATOTAN*<br>In those thats I've listened to, in those championships | 0 | 0.41 | 0.32 |
| *Klaro, nik uste goain zeatuo daola, MODERNIZATUO eo, esan nahi dut*<br>Of course, I think upper it is more that, more modern, I mean | 0 | 0.33 | 0 |
| *ikasi in behar da aurrena eta falloik gabe, þues, zeatzen, hitzek behar den bezala*<br>You must learn and without failures, well, doing that, forming words properly | 0.3 | 0.42 | 0.4 |
| *konpletogokoa ote den euskera mintzatzeko moduu o beste, beste, zea bat ote den*<br>If one way of speaking is more complete or not, it's other, other, that | 0.32 | 0.19 | 0 |
| Median values for /zera/<br>/pues/ | 0.09 | 0.32 | 0,13 |
| *biñooo, þues emen bai*<br>But, well, here yes | 0.8 | 0.71 | 0.12 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Badaezpada pues, etortzen baldin bada gaitza bat*<br>Just in case, well, if a disease comes | 0.2 | 0.45 | 0.34 |
| (Question) *pues ooo pues bai*<br>(Question) Well or well yes | 0.65 | 0.55 | 0.14 |
| *Beak esaten zizten pues mmmmm etxea erose beauzu*<br>They told us to buy the house | 1.12 | 0.24 | 0.45 |
| *etten ginuzen pues eeeeeeee asko*<br>We did well eeee a lot | 0.65 | 0.66 | 0.3 |
| (Question) *pues hemen eeeee beste*<br>(Question) well here eeee another | 0.49 | 0.64 | 0.36 |
| (Question) *pues eeee herri batetik besteaaaa*<br>(Question) well www from one town to another | 0.3 | 0.38 | 0.23 |
| *Gainera pues eeeeee etxean*<br>Moreover well eee at home | 0.45 | 0.52 | 0.31 |
| *eeeeee pues egiten da*<br>eeee well it is done | 0.3 | 0.84 | 0.2 |
| (Question) *pues garbiketa*<br>(Question) well the washing | 0.16 | 0.26 | 0.54 |
| *Ez ez ez, guuu gara pues eeee, guk etten dugu txiki guztia ez?*<br>No, no, no, we are well eeee, we do all the litle stuff no? | 0.3 | 0.78 | 0.45 |
| *diputazioak pues kriaderoan*<br>The deputation well at the hatchery | 0.6 | 0.98 | 0.51 |
| *Eta farion oi da pues errekan*<br>And that "fario" is well at the river | 0.54 | 0.67 | 0.89 |
| *ta hoik pues ez dira berdin asetzen*<br>And those well aren't equally satisfied | 0 | 0.77 | 0.44 |
| *Eeee::, batzuek baj, andienak baj, gero ttikiak joaten dira pues, Riezu*<br>Eeee, some of them yes, the biggest ones yes, then the little ones go, well, to Riezu… | 0.22 | 0.51 | 0.35 |
| *da interna o externa ta ajk pues tratatu egiten ditugu*<br>It is internal and external and we treat all | 0.12 | 0.36 | 0.33 |
| *Amuaraj bakotxak, ta ojk pues ez dira denak berdin asetzen ez*<br>Each trout, and they well they don't satiate equally | 0.16 | 0.4 | 0.18 |
| *kalidade andikoa ta Yesan pues pantanoko ura da ta eztu kalidade oso ona ez?*<br>It is of high quality and in Yesa well it is swamp water and it doesn't have high quality | 0.23 | 0.43 | 0.34 |
| *Baino hori da pueeeees etxekoa*<br>But it is well from our home | 0.06 | 1.27 | 0.26 |
| Median values for /pues/<br>/eta/ | 0.37 | 0.65 | 0.35 |
| *Baldin baziñuken errege taaa xaldiii ta*<br>If you had the king and the horse and | 0.08 | 0.44 | 0 |

*(Continued)*

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Arraras taaa igoaaaa*<br>Arrarats and Igoa | 0.13 | 0.65 | 0.33 |
| *Tolosa, Iurtzun, taaa ok izaten zeen*<br>Tolosa, irurtzun, and those used to be | 0.33 | 0.36 | 0.25 |
| *Ardoa taaaaa cerveza*<br>Wine and beer | 0.11 | 0.5 | 0 |
| *naranjak etaaa danatatik ewkitzen zuten*<br>They had oranges and everything | 0.15 | 0.77 | 0.6 |
| *Gutxio kostatzen da taaa*<br>Because it is cheaper and | 0.6 | 0.35 | 0 |
| *Igoara taaaa Lekunberrea*<br>To igoa and to Lekunberri | 0.13 | 0.45 | 0.78 |
| *Majo afaldu taaaa ameiru pezta*<br>We had a great dinner and 13 pesetas | 0.66 | 0.39 | 0.55 |
| *afaltzea joan ta: launbete askotan 12tako*<br>We went for a dinner and lots of Saturdays by 12 o'clock | 0.12 | 0.8 | 0.12 |
| *Lasei taaaa apuroik gabeee*<br>Relaxed and without problems | 0.6 | 0.34 | 1.03 |
| *Rapido taaaa ligero*<br>Fast and lightweight | 0.11 | 0.66 | 0.7 |
| *Ezpalibauzu lau o bost bizitze ein, taaa, casa rural*<br>If you haven't done five–six lives, and, country house | 0.57 | 0.56 | 0.33 |
| *apezan itxee salduta, taaa ez da saltzen eta ola dabille gawze*<br>Once the priest's house is sold, and, it is not sold, this i how it is | 0.8 | 1.18 | 0 |
| *Bai, taaa geooooo besteat*<br>Yes, and then another | 0.37 | 0.44 | 0.43 |
| *Ta oajndik azteje, berrojtaaa… amabi urte olaxe izaan zittun*<br>Yet they grow, fifty… two years or so | 0 | 0.55 | 0.5 |
| *batek iruñin taaa bestik or badaaa eztajt ba*<br>One of them in Iruña and the rest somewhere around | 0.36 | 0.94 | 0.66 |
| *Ta oain haik atea taaaaaa…*<br>And right now they get out and… | 0.23 | 0.49 | 0 |
| *Xaldik taaa giooooooo olaaaa* (Question)<br>Horses and then that way | 0.29 | 0.45 | 0.22 |
| *ordun ta mantsoagoan asten naz ordun taaa batuagoan.*<br>The slower I speak the more I use Basque | 0.16 | 0.6 | 0.47 |
| *Biño aaa gutxio*<br>But ever less | 0.2 | 0.58 | 0.17 |
| *salistarrak ta difeente etten dute, taaa biwtearrak gugatik difeente etten dute*<br>People from Salidas speak different, and the people from Beruete differ from us | 0.19 | 0.53 | 0.33 |

**Table 3.** (Continued)

| Text | Time before filler (s) | Filler length (s) | Time after filler (s) |
|---|---|---|---|
| *Zapetillek eta*:: [*A bai e?*] *Bai, bai, danatatik*<br>Shoes and [Yes?] Yes, yes, everything | 0.34 | 0.43 | 0.23 |
| *Igoatik eta*::, *emengo neskak*<br>From igoa and, girls from here | 0.14 | 0.46 | 0.29 |
| *Baj eee baj, ordun ezen bestejk ardoa ta:, zerbeza e izaten zen*<br>Yes eee yes, then there wasn't choice there were wine and beer | 0.22 | 0.35 | 0.36 |
| *Taarnan afarii in ta kordionistee ta*:: *eztakit zenbat*<br>We ate in the bar and the accordionist and I don't know how many | 0.34 | 0.44 | 0.33 |
| *Taaa geroooooo harina de pescau*<br>And then fish flour | 0.s | 0.47 | 0.32 |
| Median values for /eta/<br>/m/ | 0.22 | 0.53 | 0.32 |
| *Baj, ba, beak esaten zizten e*::, *pues mm esenii*<br>Yes, well, they call this e:, well mm milk | 1.12 | 0.24 | 0.45 |
| *Oain dala berrogei mmmmm berrogeita bost*<br>Forty mmm forty-five years ago | 0.9 | 0.46 | 0.42 |
| *besteatzuki beis etten duzu euskeas, daaa mm bai*<br>You speak Basque with others, and mm yes | 0.84 | 0.28 | 0.3 |
| *Libre? Mmmm bueno*<br>Free? Mmmm well | 0.33 | 0.23 | 0 |
| (Question) *mmmm buuu eztakit*<br>(question) mmm buuu I don't know | 0.83 | 0.28 | 0 |
| *nei herriko mmm ez bazaie tokatu*<br>For me from the town mmm if they haven't happened to… | 0 | 0.54 | 0.3 |
| *Gio ya artzen diozte ez mmm*:: *beti gauzek beak esaten dee*<br>Then they get him didn't they? Mmm he always tells everything | 0 | 0.78 | 0.4 |
| *Da*:: *mm*:: *baj, no:r zengatik ezawtzen duzun eo zergatik ajtzen zan eo ola*<br>And mm yes, who do you know why or why did he/she did it or so | 0.33 | 0.47 | 0.28 |
| *taaaa mmmm jendiii plazaa ateazeko*<br>And to make people get out to the square | 0.2 | 0.8 | 0.23 |
| Median values for /m/ | 0.33 | 0.28 | 0.3 |

## Appendix 1

## Appendix 2

### *Speaker 1*

*Ojxe bai, esan dizutena oajn dela ogejtamar urte, litroat esenek gawr baño diru gejo balio zun ta. Bestea baño gawzek, ario artaa, noa yoan dee? Ua bajta komedii, obrero batek buzoa sartzen du ta zortze ordu fabriken paastu ejn bear dute, egii da, baño arrek dawke yornala fi[x]oa* (46 words)

*Ta, ne ajtte zena itxea etorre zen, ta esan zun, "oajñ* [*e*:], *Basaburun, den dirujk gejna arraatsen, Kontu guzik ejñ, ta, onze millones, sobra, nahikoa sobrante", onze millones gelditzen zeela, kontuk atea e, oajn dela berrogeita bost urte olaxe* (38 words)

An beste baserre bat badao, soro koskor batzuk bajttu iruñeako batek, ta, lengoan ojn atteekin amistadi awndii nun nek, ta beti esaten zeen, buah, zuk iilli bearko duzu Juan Jose sorotto guzik ta ardiikin zuk iilli bearko duzu (35 words)

Txondorra, su ematen zitzaionin txondorra baakizte alda aldea nola ejtten den eo ez? Baj baakizute. Ta:, geo gojtik botatzen zitzaion suu, ola rejillek ejtten zitzajzkion ola rejillek paatuz gora gora gora tximinik, ta geo suu andik botatzen zioten (38 words)

### Speaker 2

*itzek asko sartzen ttut erdeaskoak, igual batun aj najzenin iwal sajatzen naz pixkitto bat gejo, bueno nola esaten ez? Telebisioa nola esateute? Eztakit nola nola esaten dute ez?* [*Beño*::] *askotan ez deket* (30 words)

[*Ta*::] *biwtearrak gugatik difeente etten dute, biw-biwtearrak guañin gejo nabajtzen dee. Ojk difenteo, eta ultzamaa bajatzen bazaa difenteo. Emendik gartzaona ere difeentziik badao, neetzat gartzaondarrak eta eltzaburukoak eta, klase osora etten dute* (30 words)

Igoan baj, igoan eonen dee gejo. Igoan de ene bajno gazteagoak, ne edadekoak eta ezta. Bajno ni oajn emezrtzi eo ogej urtekoak igual eonen dee zortzi eo… Eta geo igual igoakoa ama eo ajtte dutenak eta geo itzootzen deenak eo, baj kuadrille politta badai baj… baj baj, dozena bat eo baj (50 words)

Oi jon [ze::n], azkeneko eondu Joan zen Larrajntzarra bizitzea, baj, gizonak an ejtten zun lana eta ordun [m::], eta iruñea joateko estudiatzea joateko, ooki gelditzen zaien, awtobuse bada an. Baj, eta ordun area:, area Joan zen (35 words)

### Speaker 3

*Baj baj, eskalera bat bezala da ez?* [*m*], *barnen egoten dira, o sea, badugu nabe bat aundia, eta beste bat txikia, arraultzak an patzen ditugu, eta gero ateratzen ditugu kanpora, eta gero esan dizuten bezela eskalera bat bezela da ez?* (39 words)

*Baj, pentsuan enpresak badauka bere laboratorioa Burgosen, ta pues ikusten baldin badugu gajtza dawkatela bidaltzen ditugu area muestra batzuk eta esaten digute gajtza aw dawkae eta tratatu onekin edo arrekin* (30 words)

Baj, amuarraje au da raza bat arco iris du izena eta au oso ongi adaptatzen da kriaderoan azitzeko. Gero beste bat badago farion izenekoa, baino ori da pues ori erabiltzen dute errepoblatzeko erreka, eta oj diputazioak jartzen ditu bere piszifaktoriak (40 words)

[Ta:] beti bildurra igual bakarron bat igual pasako da egunen batean eta pues erreko [luke::] erreka guztia, zergatik? Oj ibiltzen dira kasi udan ez? Neguan kantidade asko doa errekan ta, zaile da kalte egitea ez? (35 words)

## Appendix 3

The criteria used to calculate relative frequencies of cases are as follows:

1. Nouns:
   a. Included: nouns that have subject or object function (*Obrero batek buzoa…*), pronouns (*nek, zuk, oi, ua*, and so on), attributes (*Itzek sartzen ttut erdeaskoak*);
   b. Excluded: adverbs (etxera, larraintzarra, and so on), volumes (*LITROat esenek, duela ogejtamar URTE*);
2. Determiners:
   a. Included: articles *(-a, -ak, -ek, -( r)ik*, and so on), numbers (*ogej, bat, batzuk*, and so on), demonstratives (*au*);
   b. Excluded: relatives (*dizutena, aite zena*), verbal nouns (*egitea*);
3. Verbs:
   a. Included: synthetic and periphrastic verbs, verbs of relative clauses;
   b. Excluded: verbs that form the root of verbal nouns;
4. Adjectives: all adjectives that qualify nouns and pronouns;
5. Adverbs: all adverbs that modify verbs.