

Phonological and phonetic aspects of whistled languages*

Annie Rialland

Laboratoire de phonétique et phonologie
UMR 7018 (CNRS-Université de la Sorbonne Nouvelle)
rialland@idf.ext.jussieu.fr

0. Introduction

This article examines phonological and phonetic aspects of whistled languages. A whistled language is a system of whistled communication which allows fluent subjects to transmit and exchange a potentially unlimited set of messages over long distances. In this respect, they are quite different from communication systems limited to a repertoire of stereotyped messages. For example, the whistled formulas used by certain herders or animal trainers do not constitute whistled languages as such.

A further respect in which true whistled languages differ from other types of whistled communication is that they encode auditory features of spoken languages by transposing key components of speech sounds. This article examines this second aspect of whistled languages. It will show that there are two types of whistled languages: 1) those based on non-tone languages, which transpose formants, and 2) those based on tone languages, which transpose tones. However, both types are similar in most other respects. Both transpose the basic amplitude envelope of the spoken utterance they are transposing. This envelope has a dual function, assigning whistled sounds to a small number of major classes while providing a frame for the alignment of whistled melodies with phone boundaries. In addition, both types of whistled languages have a phonological structure which is related to, but partly independent of that of the spoken language they transpose. This structure involves acoustically-defined features, similar to the acoustic correlates of spoken language features but subject to the specific constraints of the whistled medium. We will show that these features can be viewed as primitives organized into a simple hierarchical organization.

The discussion is organized as follows. Section 1 of this article examines formant-based whistled languages. General background information is given in section 1.1. The main phonological and phonetic features of Silbo Gomero, a whistled language of the Canary islands, are presented in section 1.2, and Silbo Gomero is compared with the Turkish whistled language of Kusköy in section 1.3. Section 2 of the article is devoted to tone-based whistled languages. Background information is given in section 2.1. Two case studies then follow, one involving a whistled language based on Hmong (a Hmong-Mien language of China and Indochina) in section 2.2, and the other a whistled language based on Moba (a Gur language of Togo) in section 2.3. Section 3 offers general conclusions.

1. Formant-based whistled languages

Formant-based whistled languages are based on non-tone languages, that is, languages that do not make lexical or grammatical use of tone. Before examining their phonological and phonetic aspects more closely, we give some background information on the best-known whistled languages of this type.

* I would like to thank my colleague Bernard Gautheron, to whom I am greatly indebted for the study of Silbo. This study is greatly based on data that we collected in La Gomera. He contributed his expertise in sound engineering, which was particularly valuable, in the difficult recording conditions. His experience in designing perception tests for deaf persons were an important asset in the organization of our perceptive test. His friendliness and insights were also very valuable, while conducting the interviews with the whistlers.

1.1. Formant-based whistled languages: previous studies and extant data

Among whistled languages practiced today, perhaps the best known is Silbo Gomero, the Spanish-based whistled language of La Gomera, an island of the Canary archipelago. Silbo Gomero is usually thought to have been inherited from the previous inhabitants of the island, the Guanches, a Berber population which was decimated, defeated, and enslaved by Spanish armies in the 15th and 16th centuries. Silbo Gomero was subsequently used by the herders of the island for communication among themselves, and it also served an important role in organizing the defense of the island against foreign invaders. It was still in general use as late as the 1950s. After this time it declined progressively, and by 1998 the estimated number of fluent whistlers (i.e., those able to converse in Silbo) had dwindled to fifty. It currently benefits from positive efforts to save it as part of the patrimony of the island. Since 1999, it has constituted a required subject in the primary and secondary school curriculum.

Silbo Gomero also occupies a special position in the history of studies of whistled languages. It was the first whistled language to be studied scientifically. The fact that Silbo Gomero was not an independent code but a transposition of Spanish from speech to whistling was first understood by Lajard (1891), who concluded his study: "Nous avons vu que [le langage sifflé] n'est pas particulier au point de vue linguistique, car c'est de l'espagnol" ["We have seen that the whistled language is nothing special from the linguistic point of view, because it is Spanish"].

Important later studies of Silbo Gomero are due to Classe (1957) and Trujillo (1978). Trujillo has played an important role in the survival of Silbo by developing an instruction program for schoolteachers. Several documentary films have also been made on Silbo. The main one is "Les derniers siffleurs de la Gomera", a 50-minute documentary directed by Marc Jampolsky, in which the present writer participated as scientific advisor. This film, first shown on French television (TV5) in 1999, includes interviews with whistlers, several whistled exchanges, excerpts from perception tests, and examples of classroom instruction in Silbo. Some of the initial results collected for this film have been presented in conferences (Rialland 2000, 2002, 2003). In addition to this work, Carreras et al. (2005) have recently published fMRI studies of Silbo.

Other whistled languages are less well known. A Turkish-based whistled language was discovered in Kusköy, a village located 30 kilometers from the Black Sea in the valley of Görele in Turkey, and a team of scientists led by the french acoustician R.G. Busnel went to Kusköy to study it. The result was a broad study based on recordings, films, and perception tests, most of the results of which were published in a special issue of the journal *Phonétique appliquée* (n° 14-15, 1970). The discussion below will be based on these publications as well as on films and recordings made by Busnel and his colleagues (Busnel 1967, 1968, 1970, Busnel and Classe 1979).

Another whistled language has been studied by Busnel, Moles and Vallancien: that of Aas, based on a Bearnese dialect in the French Pyrénées. However, this language was already vestigial when it was first reported in the late fifties. Few people were still able to whistle it, and their repertoire was limited to a small number of stereotyped formulas. A systematic study of 22 of these utterances was published as Busnel et al. (1962a, b).

A Greek-based whistled language was studied by Chrambilikis (1994) and by Xiromeritis and Haralampos C.Spyridis (1988-1989). It was in use in Antias, an isolated village on the island of Evia. Outside Europe, another whistled language of this type is the one based on Tepehua in the state of Hidalgo (Mexico), described by Cowan (1972).

1.2. Silbo Gomero

Our discussion of Silbo Gomero, or *Silbo* as it is called by its practitioners, is based on previous studies by Classe (1957) and Trujillo (1978) and on our own first-hand data. Trujillo's study includes 89 spectrograms. Our data have been drawn from recordings and films of whistled exchanges made in various situations, including teacher-student interactions in the classroom, interviews with fluent whistlers, observation of teaching methods, and a preliminary perception test involving two whistlers. The materials were collected by the present author and her colleague Bernard Gautheron in La Gomera in November, 1998.

1.2.1. Subjects

Our study is based primarily on three subjects, whom we identify as I, L, and A. The first two are in charge of teaching Silbo in primary and secondary schools, where they are playing a key role in the revival of Silbo. Both used Silbo on a daily basis in their childhood and young adulthood. The third is a female whistler who used Silbo in her childhood and later while raising a family of six. L was our main source of information on the history of Silbo, teaching methods, and judgements of the ease with which different sounds and words are encoded in Silbo. L and A were employed as subjects in the perception study described in the Appendix.

1.2.2. Transposition of speech in whistled language: an introduction

As an illustration of a whistled utterance in Silbo, let us compare spectrograms of the spoken and whistled versions of the sequence *epe* as produced by subject I. These are shown in Figure 1.

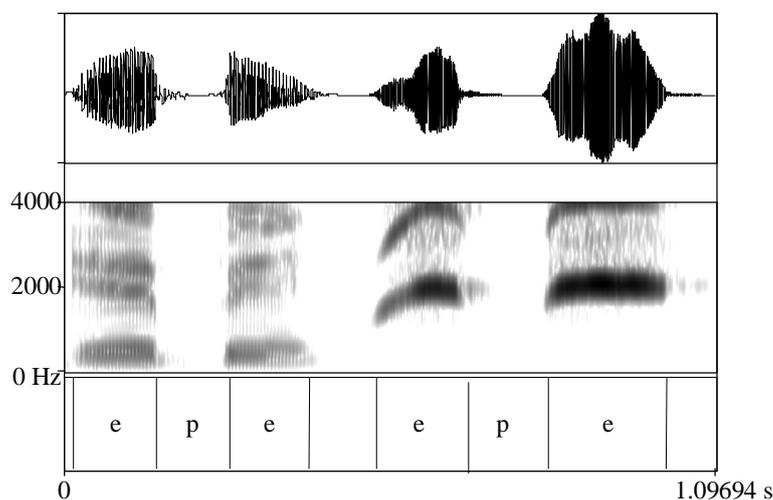


Figure 1. Waveforms (top) and spectrograms (bottom) of *epe* as spoken and whistled by whistler I. The spoken utterance is shown on the left and the whistled utterance on the right.

Spectrograms of spoken speech such as the one shown on the left display the formant structure of vowels as well as other information such as duration and relative intensity. In this example, for instance, we see the first three formants (F1, F2, F3) and part of a fourth. Spectrograms of whistled speech such as the one on the right are less familiar to linguists and require brief comment. In this example we observe two harmonics of which the first (i.e.,

lowest) one, H1, is located at about 2000 Hz . This harmonic constitutes the fundamental frequency (F0) of the whistled tone, and determines the perceived pitch of the whistled utterance. Higher harmonics are whole-number multiples of this fundamental; however, they do not appear to play a role in differentiating vowels, since under normal conditions of transmission only H1 is loud enough to emerge from background noise.¹

If we now compare the F2 contour of the spoken utterance on the left with the corresponding H1 contour of the whistled utterance on the right, we observe that both are similar: both fall towards a low-frequency locus at the edges of the consonant. As a first generalization, we can say that the F2 contour of a spoken utterance is transposed into a roughly similar H1 contour in its Silbo counterpart.

Intensity modulations are also parallel in the spoken and whistled utterances, as is shown by the waveforms at the top. Thus in both spoken and whistled utterances, vowels are high-amplitude sounds while stops are represented by silence.

The following sections will shown in more detail how Silbo encodes the vowels and consonants of spoken Spanish by transposing F2 contours and signal envelope modulations into whistled H1 patterns.

1.2.3. Transposition of vowels

Spanish has five vowels, *i*, *e*, *a*, *o* and *u*. Let us compare these vowels as spoken and whistled by the same person in the sequence /i e a o u/ (Figure 2).

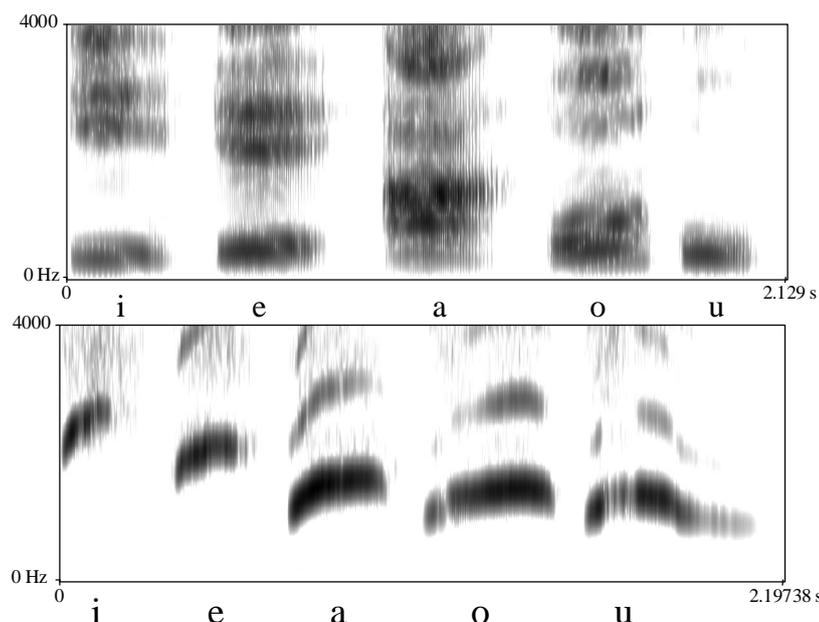


Figure 2. Spectrograms of the vowel sequence *i*, *e*, *a*, *o* and *u* as spoken (top) and whistled (bottom) by whistler I

Spoken vowels are shown on the top. For this subject, these vowels have F2 values ranging between about 800 Hz and 2300 Hz overall. Whistled vowels, shown on the bottom, have a first harmonic (H1) ranging between about 1200 Hz and 2700 Hz. A comparison of the two

¹ Our colleague, Bernard Gautheron, measured the whistled signal at around 6 dB over background noise across a distance of 1500m; at this distance, only the first harmonic can emerge. Recorded messages over a normal communicating distance also show that only the first harmonic emerges (Busnel and Classe 1979).

shows that the descending F2 pattern of the spoken vowels is paralleled by a descending H1 pattern of the whistled vowels; that is, both F2 and H1 fall as we proceed across the series *i*, *e*, *a*, *o*, *u*. No other vowel formant shows this pattern. It is apparent that F2 -- and only F2 -- is transposed into Silbo.

Closer study shows a small difference between the two patterns, however. While F2 descends in a roughly linear progression from *i* to *u* in the top spectrogram, H1 gradually flattens out as we approach *u* in the bottom spectrogram; indeed the H1 value of *o* and *u* are very similar. Thus the whistled vowels do not copy F2 exactly. The reason for this is that the pitch range of H1 has a "floor" value of around 1200 Hz for this whistler, while the pitch range of F2 drops to much lower values (around 800 Hz in *u*). As a result of this difference, lower values of F2 must be shifted upward when transposed into Silbo.

This limitation explains the difference we find in comparisons between front vowels on the one hand and between back vowels on the other. Spoken and whistled contours of front vowels resemble each other closely. Thus, comparing the front vowels *i* and *e* in Figure 2, we see that the H1 of whistled *i* (2620 Hz) is only slightly higher than the F2 of its spoken counterpart (2390 Hz), and corresponding values of *e* are nearly identical (H1 = 1930 Hz, F2 = 1990 Hz). In contrast, the values of whistled *o* (1380 Hz) and *u* (1270 Hz) are much higher than the F2 values of their spoken counterparts (890 Hz and 790 Hz -too weak to be seen on this figure-, respectively). The other subjects showed similar discrepancies, which are explained in the same way.

The best match between whistled and spoken vowels is found with the vowel *a*. Depending upon the whistler, the values of H1 and F2 can be about equal, as in our example (H1 = F2 = 1480 Hz), or H1 can be slightly higher (see also Trujillo 1978). The vowel *a* plays a special role in Silbo in that it provides a reference point for whistlers, who usually begin their messages by whistling *a* followed by the name of the addressee: *a Bernardo*, *a Maria*, *a Sebastian*, *a Domingo*. In this respect, the vowel *a* functions somewhat like the concert A to which the members of an orchestra tune their instruments.

Not all these vowels are distinguished in perception. Trujillo (1978) argues that there are only two distinct vowels in Silbo, an "acute" vowel corresponding to *i* and *e* and a "grave" vowel corresponding to the others. In the teaching methods developed for Silbo inspired by Trujillo's work, these are presented as the fundamental vowels. It seems possible, however, that other vowels may be perceived as well. In our perception test (see the Appendix), *o* and *a* were not merged in production, and they were sometimes correctly distinguished, with *a* receiving a higher identification score than *o*. Classe (1957) similarly observed that *Antonio* and *Antonia* can be whistled differently, but that values of *a* and *o* often overlap. It appears, then, that the distinction between *a* and *o* is often present in production and even in perception, but is relatively weak and inconsistent.²

We have so far established the following points: 1) Silbo transposes F2 to H1 in vowels, 2) no other formants are transposed, 3) Silbo has a fundamental contrast between "acute" and "grave" vowels corresponding to front vs. central and back vowels in spoken Spanish, and 4) further distinctions within the whistled vowel space appear possible but are less robust. The vowel space can be regarded as organized in terms of a hierarchy involving a primary division of the vowels into two sets distinguished by a robust acute/grave contrast, and a less robust secondary division between within the grave category.

² Not enough data were obtained to permit discussion of *e* and *u*, and the status of these vowels must be left for further work.

1.2.4. Transposition of consonants

The transposition of consonants involves two main components, consonant-vowel transitions, and signal envelope modulations. We consider these in turn.

1.2.4.1. Transposition of consonant-vowel transitions

Previous work has shown that Silbo and other formant-based whistled languages transpose consonantal-vowel transitions of the spoken language into the whistled waveform. We will show here that just as in the case of vowels, these transitions are not simply copied from the corresponding speech transitions but involve a partly conventionalized adaptation motivated by the more restricted range of the whistled melody.

Let us first examine the transitions at the edges of the dental consonant *t* in *ota* (Figure 3).

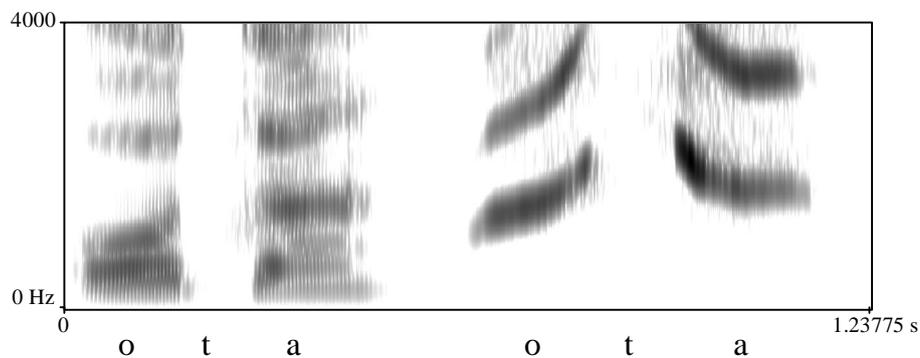


Figure 3. Spectrograms of *ota* as spoken (left) and whistled (right) by whistler I

We see that the F2 transition (left) and the H1 transition (right) both rise at the end of *o*, as is typical of this vowel before dental sounds. However, not all vowels show such a close parallel, as we see from a comparison of the spoken and whistled transitions in *ete* (Figure 4):

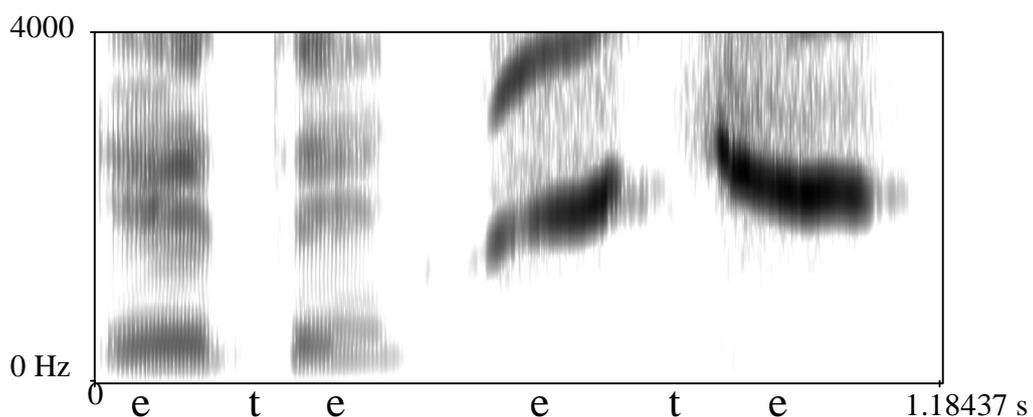


Figure 4. Spectrograms of *ete* as spoken (left) and whistled (right) by whistler I

Figure 5. Whistled H1 trajectories in *ete*, *iti* and *oto* (whistler I). Measurements were taken every 10ms. Durations of intervocalic silent intervals were averaged

Here the F2 contour falls in the first vowel of the spoken utterance (left), while the H1 contour rises in its Silbo counterpart (right). The same discrepancy can be observed when *t* is flanked by *i* in the sequence *iti*.

These discrepancies make sense if we posit that Silbo utterances do not simply copy F2 transitions but involve partly conventionalized patterns. To see this point, it is useful to model F2 and H1 contours in terms of abstract *loci*, that is, idealized points toward and away from which the contours appear to move. Although the notion "locus" was defined in terms of unique points corresponding to each consonantal place of articulation when this notion was first introduced, subsequent speech research has shown that loci for each place of articulation vary somewhat according to the identity of surrounding vowels. This is true of loci in Silbo as well, as can be observed in the traces in Figure 5, illustrating typical H1 trajectories in the vicinity of *t* in three vocalic contexts (whistler I).

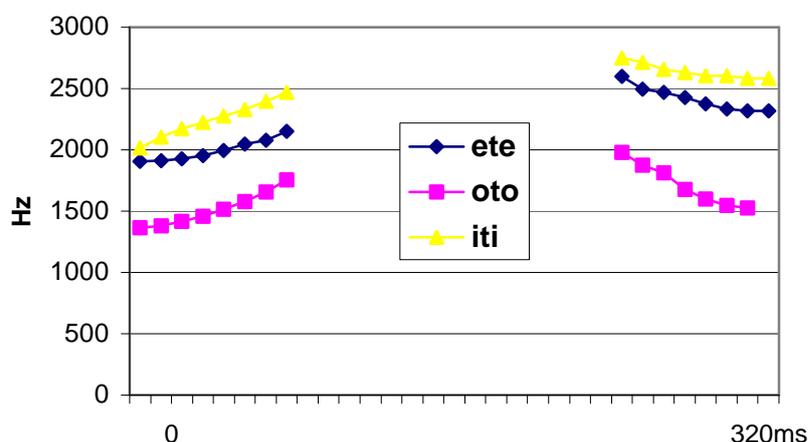


Figure 5. Whistled H1 trajectories in *ete*, *iti* and *oto* (whistler I). Measurements were taken every 10ms. Durations of intervocalic silent intervals were averaged

Each locus induces the same effect as a consonant locus in speech: transitions appear to move towards and away from this point, and have somewhat different values depending on the identity of the surrounding vowels. In our data for this whistler, loci for dental sounds are usually pitched within the 3000-3200 Hz range, which is well above the maximum H1 value of *i* (around 2700 Hz) and much higher than the average locus for dentals in spoken Spanish, around 1800 Hz. Other whistlers had similarly high loci for dentals, located at around 3000 Hz for whistler L and 3200 Hz for whistler A.³

We also observed a low-frequency locus at the edges of labial consonants, as shown in the spectrograms of *apa* (Figure 6).

³ Figure 5 also reveals an asymmetry in the shapes of pre- and post-consonantal trajectories suggesting that the temporal center of the locus occurs about two-thirds of the way through the consonant.

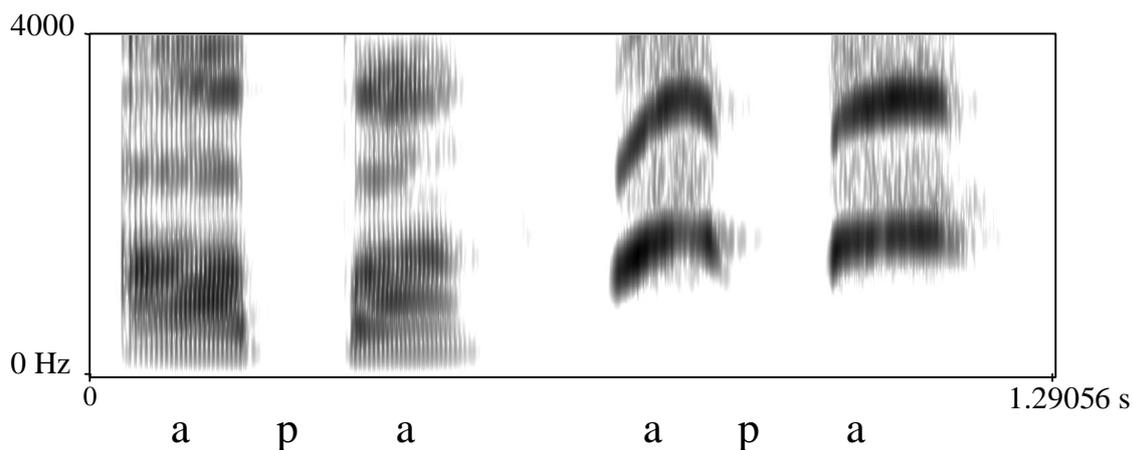


Figure 6. Spectrograms of *apa* as spoken (left) and whistled (right) by whistler I.

In this example, as well as in *epe* (Figure 1), transitional movements are similar in the spoken and whistled utterances. The fall towards a low locus is typical of labial consonants. Let us consider another context, however, that does not show this parallelism. Spoken and whistled productions of *upu* are shown for whistler I in Figure 7.



Figure 7. Spectrograms of *upu* as spoken (left) and whistled (right) by whistler I.

Comparing the spoken and whistled forms, we see a discrepancy between the relatively flat F2 pattern of the spoken utterance (which is not clear on the spectrogram as F1 and F2 of *u* are very close together) and rapidly moving H1 pattern of the whistled utterance. The whistled transitions point towards a locus at around 800 Hz, which is considerably lower than the H1 value of *u* (around 1200 Hz), the lowest vowel in the system. Thus the whistled labial locus is pitched at a lower level than that of any vowel of the system. Further data show that the labial locus is higher in the context of the front vowels *i* and *e* (around 1100 Hz), but still below the H1 of *u*. The two other whistlers (L and A) showed the same pattern. This extra-low locus is also observable in many spectrograms published by Trujillo (1978). Thus the whistled labial locus occupies a position symmetrical to that of the dental locus: the labial locus is located well below the range of vowel realizations while the dental locus is located

well above it. In other words, dental and labial consonant loci frame the upper and lower bounds of vowel realizations.

To summarize, the Silbo system does not simply copy F2 movements of speech, but adapts them to respect constraints specific to whistling. Consonant loci in Silbo are located at more extreme values than spoken loci, framing the "vowel space".

1.2.4.2. How many loci are there in Silbo?

A distinction between "acute" consonants with rising transitions and "grave" consonants with falling transitions was established in the work of Trujillo (1978). This two-way distinction is taught by teachers of Silbo. Rising transitions characterize the [+coronal] sounds of spoken Spanish, ranging from dentals to palatals (*t, d, r, rr, l, n, s, tʃ, j, ɲ, ʎ*) while falling transitions characterize [-coronal] sounds (*p, b, f, m, k, g, x*):

In our perception test (described further in the appendix), these two broad categories of consonants were reliably identified, as is shown by the following numbers :

Number of whistled [+coronal] consonants	Number of whistled [-coronal] consonants	Number of correct identifications of [± coronal]	% correct identification of [±coronal]
37	27	62	96%

The distinction conveyed by consonant transitions is a very robust one, and has been recognized as such by earlier scholars and teachers of Silbo. The results of our perception test confirm this robustness.

Silbo therefore distinguishes at least two loci. However, further loci have been found in other whistled languages (such as Turkish, which we examine in section 1.3), and we may ask whether other distinctions are made in Silbo as well.

In fact, it appears that a further distinction within the class of coronal sounds may exist, at least in the production of single words and minimal VCV sequences. Our data show that the post-alveolar and palatal sounds *tʃ, ɲ, j*, and *ʎ* have notably higher loci than the dental sounds *t, d, s, n*, and *l*. Thus, for example, whistled words such as *mañana* [maɲana] 'tomorrow', for which we have ten realizations by five different whistlers, palatal *ɲ* is realized with a consistently higher locus than its dental counterpart *n*. This is illustrated by the spectrograms in Figure 8.

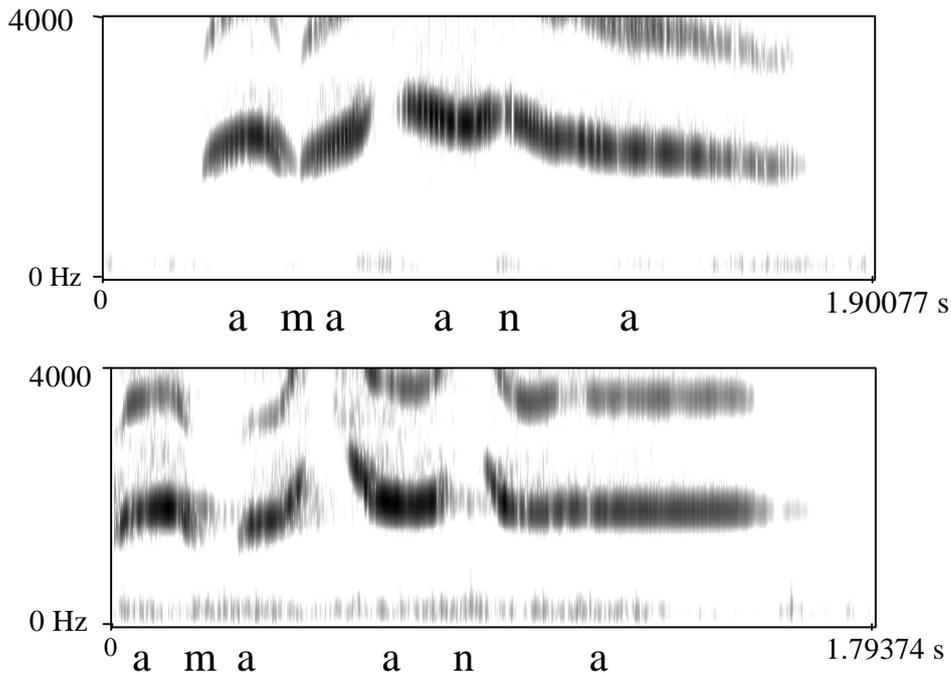


Figure 8. Spectrograms of [a ma ana] *a mañana* as whistled by two different whistlers (L and A).

We observe a similar difference in the transitions associated with the post-alveolar affricate *tʃ* and the dental *t*, as shown in Figure 9.

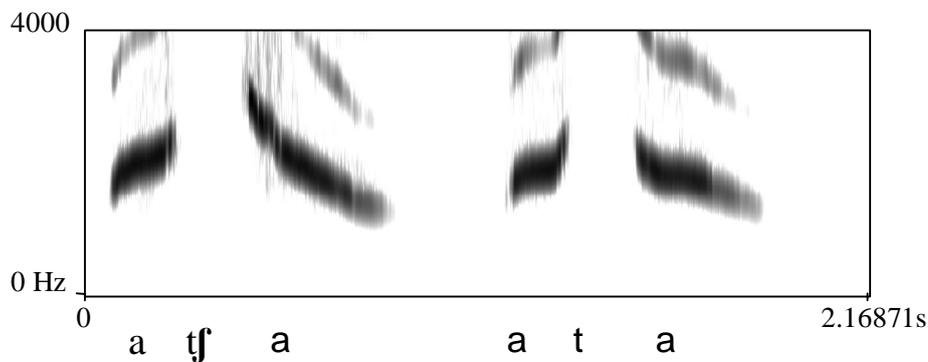


Figure 9. Spectrograms of *atʃa* and *ata*, as whistled by whistler L.

In the realization of *atʃa*, the transitional movements corresponding to the release of the affricate start higher than the movements at the release of the consonant *t* in *ata*. This difference indicates a higher locus for the post-alveolar than for the dental *t*, similar to what we find in speech. The minimal pairs *olo* and *oʎo* recorded in succession by whistler I illustrate a further distinction involving a high and an extra-high locus (Figure 10).

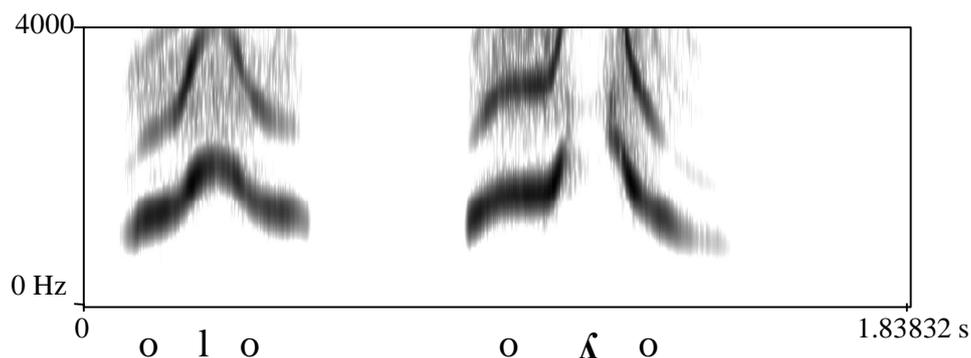


Figure 10. Spectrograms of *olo* / *oʎo* as whistled by whistler I.

ʎ has a higher locus than *l*.

We will call this extra-high locus the "sharp" locus, after the Jakobsonian term (Jakobson, Fant & Halle 1952). We will see below (§1.3) that a similar locus can be observed in the whistled language based on Turkish.

Thus, these data suggest the existence of a "sharp" locus in addition to an "acute" one, at least for some whistlers, under some conditions. However, this extra-high locus has not been recognized by previous authors, is not visible in previously published data and is not included in current teaching methods, suggesting that it is not fully robust.

We finally address the question whether velar consonants have distinct loci. Trujillo (1978) assumes that there is no difference in the realization of labials and velars, and his spectrograms support this view. In the teaching of Silbo, no distinction is made between velar and labial consonants, both of which are produced with falling transitions. In our perception test, velar sounds also tended to be confused with labials (see Appendix). They were realized with falling movements like labials, with one exception: the syllable *ka*, which was realized with no transitional movements (like spoken *ka*), and was recognized in 3 cases out of 4. With the possible exception of this syllable, our results support the view that whistled labials and velars form a single category in Silbo.

We conclude that Silbo has two main consonant loci, an "acute" one characterizing coronal sounds (both anterior and posterior) and a "grave" one characterizing noncoronal sounds (labials and velars). An additional "sharp" locus subdivides coronals into anterior (dental) vs. posterior (post-alveolar, palatal) sounds, at least for some whistlers. In contrast, the category of "grave" consonants is not subdivided further.

1.2.4.3. Transposition of the signal envelope

Let us now consider how characteristics of the signal envelope are transposed into Silbo. We first consider realizations of *apa* and *aba* shown in Figure 11. The reader should keep in mind that in spoken Spanish, the voiced obstruents written *b*, *d*, *g* are regularly realized as spirants, i.e. either as fricatives or approximants, in intervocalic position; thus *aba* is realized phonetically as [aβa], and so forth.

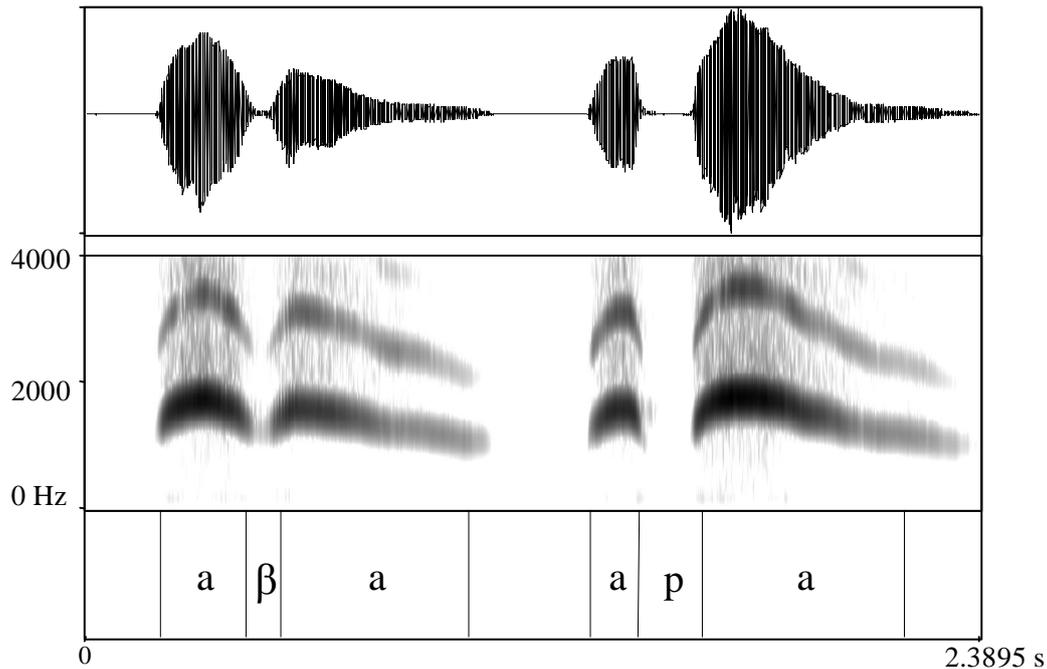


Figure 11. Waveforms and spectrograms of *aba* [aβa] and *apa* as whistled by whistler I.

This whistler makes no distinction between labials and velars; thus, the realization on the left could correspond to either *aba* or *aga* and that on the right to either *apa* or *aka*. We observe that the H1 contour is not fully interrupted in the realization of *aba* as it is in *apa*: it continues weakly even through the dip in the signal envelope. This dip corresponds to a similar dip in intensity in the spoken language. The distinction between voiceless stops and voiced spirants is encoded in the envelope as an *interruption* in the voiceless stop and as a *dip* in the voiced spirant (often with a brief period of silence, as discussed below). This distinction is realized in the same way in dentals; thus [ð] exhibits the same type of intensity dip as [β].

Now that we have seen how signal envelope modulations can be transposed in whistling, we may consider what contrasts are based on them.

1.2.4.4. What contrasts are based on amplitude modulations in Silbo?

Two types of whistled consonants were illustrated in Figure 11: "continuous" consonants showing an intensity dip, and "interrupted" consonants showing just an interruption without any dip. These two types of consonants were established by Trujillo (1978) and are also recognized and taught by teachers of Silbo. Continuous consonants encode spoken [+voiced] consonants while interrupted consonants encode [-voiced] consonants:

"interrupted" consonants	"continuous" consonants
<i>p, t, k, f, s, χ, tʃ</i>	<i>b, d, g, r, rr, l, n, ʝ, ʎ, j</i>

In our perception test (Appendix), these two categories of consonant had very good identification scores, as is shown by the following numbers:

Number of whistled [+voiced] consonants	Number of whistled [-voiced] consonants	Number of correct identifications of [±voiced]	% correct identifications of [±voiced]
32	32	56	87%

It will be noted that the continuous/interrupted distinction transposes the feature [±voiced] and not the feature [±continuant]. In voiceless fricatives, the signal envelope is different in speech and in Silbo: spoken fricatives and affricates are characterized by high-amplitude noise triggering a "bump" in the amplitude envelope, while in Silbo they are realized by a silence, just like voiceless stops. This difference shows again that whistlers do not reproduce the speech signal exactly "as is" but make use of a codified version of it.

However, our data again suggest that further distinctions may be possible, at least for some whistlers. Thus, an intermediate signal shape with a gradual decay followed by an interruption was regularly found in L's realizations of nasals. These nasals had a high rate of correct identifications (90%). The spectrograms in Figure 12 illustrate the difference between the nasal *n* and the voiceless stop *t*.

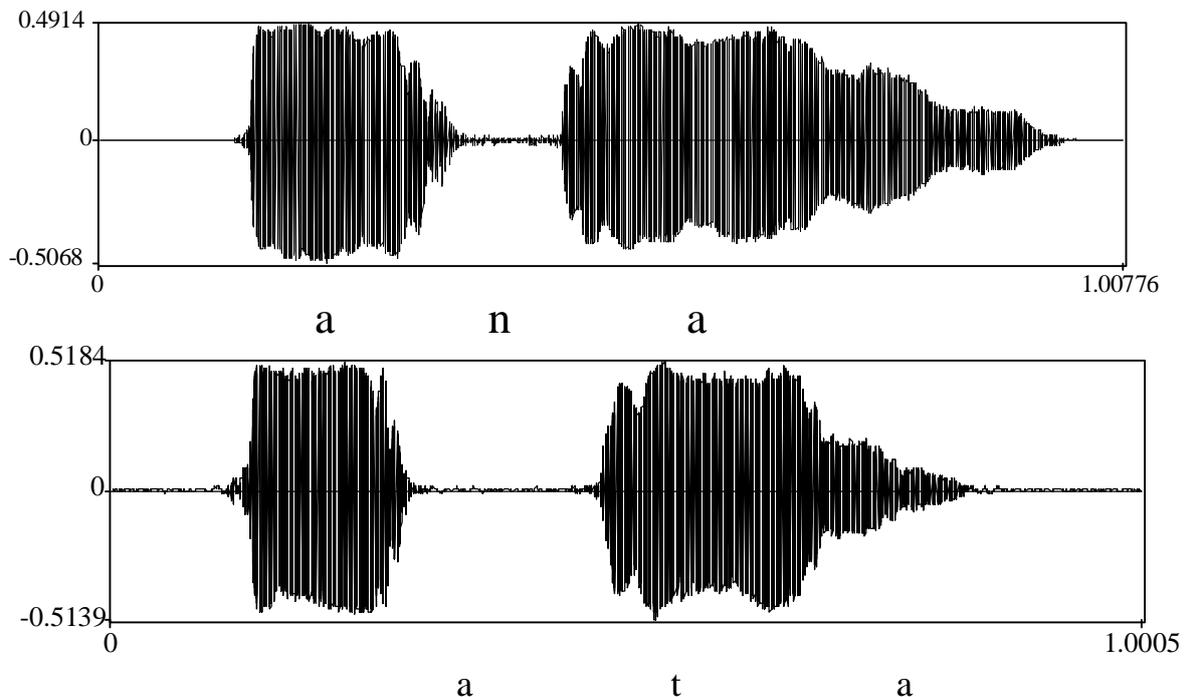


Figure 12. Waveforms of *ana* and *ata* as whistled by whistler L.

Two stages can be observed in the realization of *n*: a gradual decay in the intensity dip leading to a period of silence. The total duration of the two stages is similar to the duration of the silent interval in *ata*. In the realizations of our three main informants (L, A, and I), nasals were regularly realized with such a 'tapered' silent interval. Figure 13 shows the mean values of the silence durations in nasals, voiced spirants and voiceless stops, with standard deviations.

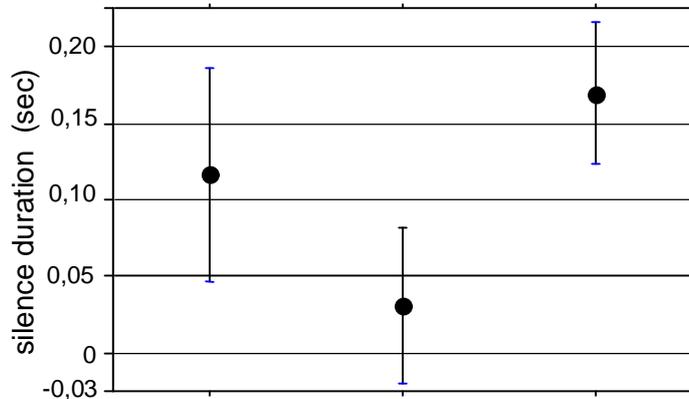


Figure 13 : Mean values of silence durations in nasals, voiced spirants and voiceless stops, with standard deviations. Measurements were made of all nasals (N = 14), spirants (N = 13), and voiceless stops (N = 16) produced and correctly identified by L. and A. in the perception test.

A single-factor ANOVA test showed that the three groups are distinct ($F(2.39) = 16$; $p < .001$). Post-hoc comparison (Fisher's PLSD) showed significant differences at the $p < .05$ level for all two-way comparisons, the lowest p-value ($p = .03$) being for the nasal vs. unvoiced stop comparison. To summarize, our whistlers realized nasals with a feature we term [gradual decay], which splits the class of [continuous] sounds into two (nasals vs. voiced spirants), while Trujillo's whistlers had only one category of [continuous] consonants, which contained both nasals and voiced spirants. These findings, including the variations among whistlers, are not too surprising with respect to other whistled languages. In Turkish, nasals are realized either as [gradual decay] or [continuous] as in Silbo (see §1.3), and nasals have been reported as "voiceless" (in our terms, [interrupted]) in Tepehua (see §1.4.). We hypothesize that the double phase of whistled nasals -- gradual decay + silence -- may reflect the extra complexity of these sounds, in which additional airflow passes through the nose.

Intensity modulations are also involved in the encoding of flaps (*r*) and trills (*rr*). In our perception test, however, despite these cues, flaps and trills were not correctly identified. Thus, the striking pattern shown in the realization of *arra* in figure 14 was not interpreted as a trill, but as *n* (twice) or *b* (once).

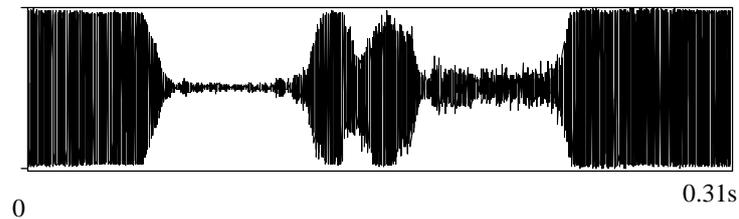


Figure 14. Waveform of *arra* as whistled by whistler L.

Leaving these sounds aside, then, three contrastive signal envelope shapes can be recognized, which we characterize as [interrupted], [continuous], and [gradual decay]. The first two are produced by all whistlers, while the last seems to be produced by only some. Thus, Silbo has a robust contrast between [interrupted] and [continuous] consonants, and a weaker one involving [gradual decay].

1.2.5. The overall architecture of the system

With this background, let us consider the overall architecture of the system.

1.2.5.1. The signal envelope as "frame" of the architecture

The important role of signal envelope modulations in transposing spoken language into whistled language may seem surprising. However, it is known that the signal envelope plays a central role in speech processing. Experiments involving stimuli in which spectral information has been removed but the envelope shape preserved have shown that while English-speaking listeners cannot recognize individual phonemes, they do identify contrasts between major classes of sounds, such as voiceless vs. voiced stops, liquids vs. glides, etc. (see Van Tassel et al. 1987, Shannon et al. 1995). Formant-based whistled languages, such as Silbo, give further evidence for the importance of these cues.⁴

Moreover, the signal envelope has another, even more basic function: it provides the "frame" for the temporal alignment of vowel and consonant realizations. H1 trajectories of consonants are timed in relation to the modulations of the signal envelope. Our spectrograms and waveforms have illustrated this type of relationship. For example, Figure 1 showed how H1 contours drop towards a "grave" locus in the vicinity of the consonant *p*. A similar frame is also present in spoken languages, where it provides landmarks for the realizations of feature contrasts. As Stevens points out, "In a sense, the acoustic discontinuities produced by forming a constriction with the primary articulators constitute the glue that binds together the various features of a consonant to constitute a "segment" (Stevens 1994, 250). In whistled

⁴ The neuroprocessing of these cues seems straightforward, as they are correlated with the amplitude of the movement of the basilar cells reacting to the signal (which basilar cells react depends mainly on the frequency of the signal). Cues such as VOT are more complex, as they involve a comparison between two events: the consonant burst and the onset of voicing. Thus, the robustness of the cues embedded in amplitude modulations may find at least a partial explanation in neurophysiological processing.

languages, amplitude modulations build up the only frame available for alignment, paralleling its role in speech.⁵

1.2.5.2. The separation between vowels and consonants

Vowels and consonants are radically distinguished from each other in both dimensions of Silbo: amplitude modulations and H1 frequency movements. In the amplitude domain, vowels correspond to peaks, and consonants to dips or silences. In the frequency domain, there is a double separation of vowels and consonants. First, vowels are realized as quasi-steady states while consonants are realized as rapid movements. Second, the frequency bands allocated to vowels and consonants are different, as consonant loci are located above and below the vowel range but not within it.

Interestingly, this division of the tonal space resembles the separation between tone realizations and intonational "loci" which can be observed in certain African tone languages. For example, in Gulmancema, in Moore or in Ncam (tone languages of the Gur family), intonation is characterized by targets located on lengthened vowels at the end of the utterance. These targets are situated higher or lower than the pitch range of tones and trigger rising or falling movements (Rialland 2004). In Silbo as well as in such tone languages, then, level tones and movements towards "extreme" targets have different functions in the system. These similarities point towards a common structuring of the melodic space, independently of its specific functions.

1.2.5.3 Feature analysis

We have seen that Silbo has a robust vowel contrast between an acute and a grave vowel (with possible subdivisions), and that it transmits place of articulation distinctions of spoken consonants by a three-way contrast among [grave], [acute] and (for some whistlers) [sharp] consonants. Consonant loci are situated outside the boundaries of the overall H1 "space", as they are higher (acute, sharp) or lower (grave) than those of any vowel. Silbo, at least in the hyper-articulated form produced in a test situation, also displays a three-way contrast signalled by the features [interrupted], [continuous], and [gradual decay]. We have retained Trujillo's features [acute], [grave], [interrupted], and [continuous], while adding two new ones, [sharp] and [gradual decay].

While the vowel system includes only two vowels, which we may represent /I/ and /A/, the consonant system is more complex. For Silbo consonants, we propose the feature analysis shown in Table 1. Silbo phonemes are shown in capital letters, and their spoken Spanish counterparts are given below in parentheses. "Optional" refers to features that are not shared by all whistlers.

	P (p, k, f, χ)	T (t, s)	CH (t)	B (β, γ)	D (ð, z, l)	J (j, ʎ)	M (m)	N (n)	Ñ (ɲ)
continuous				+	+	+	+	+	+
gradual decay (optional)							+	+	+

⁵ The signal envelope has a further function in encoding rhythmic aspects of the language, a point which we can only briefly remark on here. In Silbo, stressed vowels are longer and more intense than unstressed ones. Final lengthening at the end of utterances or at the end of major phrases is also transposed into Silbo. Intonation is very restricted in Silbo, but not totally absent. For example, final intonation sometimes occurs at the end of utterances, as is illustrated in Figure 2, in which a fall in pitch occurs at the end of the list.

acute		+	+		+	+		+	+
sharp (optional)			+			+			+

Table 1. Feature analysis of Silbo consonants (shown in upper case). Spoken Spanish counterparts of each consonant are given in parentheses.

Note that [gradual] implies [continuous] and [sharp] implies [acute] in this system (a hierarchical limitation). Given this intrinsic implication, and leaving aside the question of *r* and *rr*, *all theoretically possible feature combinations are used*. Silbo thus provides a nice illustration of feature economy, the maximal use of distinctive features.

1.2.5.4. Acoustic features

The features proposed for Silbo are acoustically-based, similar to the features of Jakobson, Fant & Halle (1954). They were chosen because they fit the Silbo data well. However, the reader might ask what articulatorily-based features for Silbo would look like. We have no articulatory data (X-rays, for example) that would help to answer this question.. It should be kept in mind, however, that Silbo is not usually whistled in the same way as a tune, with the lips rounded. Lip-rounding whistling is not powerful enough for long-distance communication. Our subjects usually whistle Silbo with the tip of of tongue curled up and grooved, placing the bent index finger into the mouth to maintain the base of the tongue in place. This tongue configuration is not ordinarily used in speech, and its various shapes in producing Silbo sounds are different from those of corresponding speech utterances. The tongue movements shaping H1 in Silbo are probably analogous to the corresponding tongue movements in speech, with parallel forward and backward displacements of the tongue. However, even with these similarities, there is no simple transfer of articulatory movements from speech to Silbo, but an adaptation. Other whistling techniques have also been reported for Silbo, especially whistling with two fingers (Classe 1963). X-rays of other whistled languages such as whistled Turkish (Busnel 1968) also show different global positions of the tongue in whistling and in speech, depending upon which whistling method is used.

Given this variation in the production of whistling, acoustically-defined features seem more appropriate than articulatorily-defined features.

1.2.6. Conclusion on Silbo Gomero

Silbo has a reduced phonological system and architecture, with phonological features, a single-formant vowel structure, a signal envelope frame for aligning formant movements, consonant loci cuing consonantal place of articulation, and amplitude modulations cuing major class features. Independent confirmation of the fact that Silbo has a linguistic structure comes from the results of brain imaging, which show that the processing of Silbo involves brain regions activated in speech processing, including the cortical language areas (Carreras et al. 2005).

1.3. Comparison with the Turkish whistled language of Kusköy

Turkish, with its 8 contrastive vowels and vowel harmony, has a more complex vowel system than Spanish. A whistled language based on Turkish was fully in use when it was studied by R.-G. Busnel and his colleagues in the late 1960s. The main results of their study

insofar as they bear on this article are presented in Leroy (1970), dealing with the phonetic study, and Moles (1970), presenting results of intelligibility tests. X-ray films have also been made (Busnel 1968).

While taking different approaches, Leroy and Moles both reached the conclusion that the vowel system of the whistled language is organized in terms of a three-way contrast. Leroy based this conclusion on the observation that vowels are realized within three frequency bands, as shown in Table 2.⁶

/a/	/o/	/ö/	/e/	/ü/	/i/	/i/
1750	1750	2100	2100	2650	2600	2600
A		E		I		

Table 2. Vowels and their H1 frequencies in the Turkish whistled language (Leroy 1970).

Thus, it seems that this whistled vowel system makes use of a ternary division of the vowel space.

As far as consonants are concerned, Leroy (1970) studied realizations of labial, dental and velar stops. She showed convincingly that there are (at least) four loci:

- a "labial" locus, lower than that of any vowel (as in Silbo)
- a "dental" locus, higher than that of any vowel (as in Silbo)
- an extra-high locus for *s*, *z*, *ʃ* and for the second part of the affricates *tʃ* and *dʒ* (as in Silbo)
- a velar locus, intermediate between the labial and dental loci, which does not exist in Silbo

Thus, a four-term opposition was found which differs on only one point from Silbo: the presence of a separate velar locus, located within the vowel space rather than outside it. (This locus can be observed in Figures 15 and 17 below.)

Leroy's study is based on spectrograms. It gives information on H1 trajectories but does not directly consider the signal envelope, though it includes useful observations such as the following: "/r/, /l/ et /j/ sont des phonèmes continus, c'est à dire qu'il n'y a pas d'interruption du signal pendant leur émission" (/r/, /l/ and /j/ are continuous phonemes, that is to say that there is no interruption in the signal during their emission).

In order to compare the use of the signal envelope in Silbo and the Turkish whistled language, we conducted an analysis of waveforms in the latter, examining data from a recording made by Moles containing 149 pairs of whistled and spoken words.⁷ We found three signal envelope shapes, similar to those of Silbo.

The first is the "interrupted" shape characterized by abrupt onset and offset, which is used to transmit voiceless sounds. This shape is illustrated by the tracings of spoken and whistled *okul* "school" shown in Figure 15.

⁶ We omit /u/ as it was grouped with high vowels, which is unexpected due to the fact that its F2 is low and this position not confirmed by the examples that we have studied (see figure 15)

⁷ The tape and its transcription were donated to our laboratory by its former director, the late Professor René Gsell.

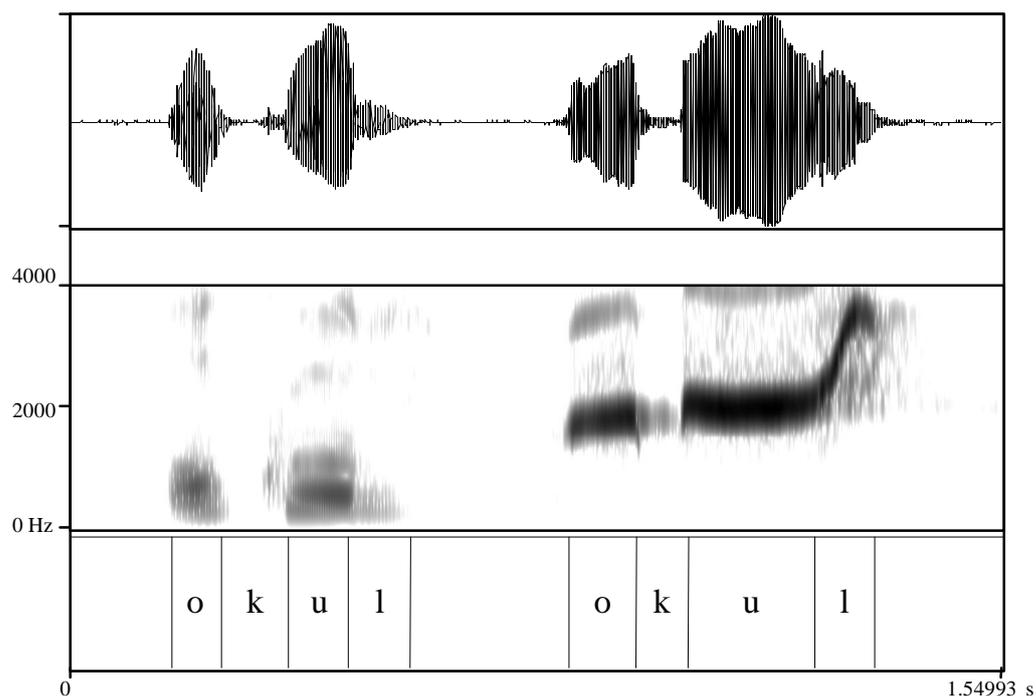


Figure 15. Waveforms and spectrograms of spoken and whistled [okul] "school" in Turkish.

In this example, the shape of the whistled signal envelope parallels that of the spoken signal. Voiceless *k* corresponds to a period of silence followed by the abrupt onset of the following vowel. *l* is less intense than the preceding vowel in both forms. The spectrograms show the same type of relation between speech and whistling as in Silbo, based in both cases on the transposition of F2.⁸ The "interrupted" envelope characterizes voiceless stops (*p, t, k*), voiceless fricatives (*s, ʃ, f*)⁹ and the voiceless affricates (*tʃ*) as in Silbo.

The second shape is "continuous", without any interruption of the signal, as illustrated in the example *hoyun* "sheep" shown in Figure 16.

⁸ Note that the locus of *k* is rather low, much more so than the velar locus found by Leroy. Turkish has two phonemically contrastive *ks*, one front and one back. The one in *okul* is back and so is expected to have a low locus. Perhaps Leroy averaged the two.

⁹ We have not confirmed this point for the fricative *f*, which does not occur in our corpus.

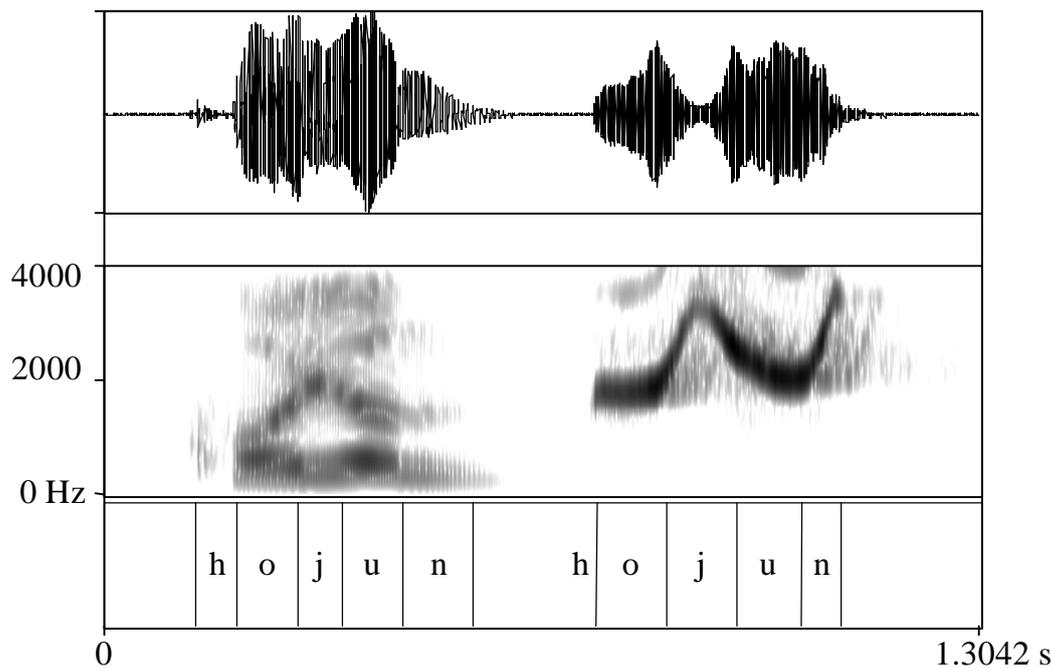


Figure 16. Waveforms and spectrograms of spoken and whistled *hoyun* [hojun] "sheep" in Turkish.

In the whistled word, the glide *j* creates an intensity dip. This type of shape is found regularly in liquids (*r* and *l*), glides (*j*) and voiced fricatives (*v*, *ʒ*).¹⁰ This type corresponds to the continuous shape in Silbo and is associated with the same type of consonants, namely glides, liquids, and voiced spirants.

As in Silbo, nasal consonants present an intermediate type, illustrated by the tracings of *inek* "cow" shown in Figure 17.

¹⁰ We had no [z] in our corpus.

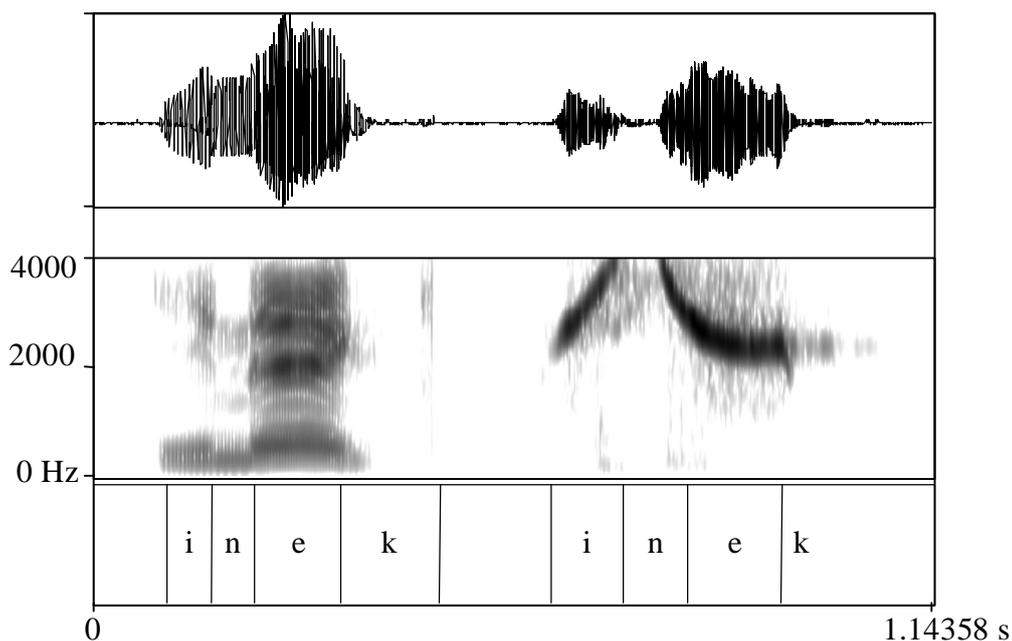


Figure 17. Waveforms and spectrograms of spoken and whistled [inek] "cow" in Turkish.

The nasal is realized with a gradual decay followed by silence. In whistled Turkish, this shape is also used to transmit intervocalic voiced stops (which only occur in Spanish as contextual variants of voiced spirants). In a few realizations of Turkish whistled nasals, we have noticed a variant with the continuous shape. However, in both languages nasals occupy an intermediate position in the system.

Thus, both Silbo and whistled Kusköy Turkish use amplitude modulations to differentiate manner of articulation in consonants. In both languages, a ternary scale can be found, with voiceless consonants occupying one end, glides, liquids, and voiced fricatives (or spirants) the other, and nasals (as well as voiced stops, when present) occupying the middle.

Combining our analysis with the earlier ones of Leroy and Moles, we propose a tentative feature chart as shown in Table 3.¹¹

	P (p)	T (t)	K (k)	S (s,ʃ)	B (b,m)	D (d,n)	G (g)	V (v)	Z (z,ʒ)	L (r,l)
continuous					+	+	+	+	+	+
gradual decay					+	+	+			
mid			+				+			
acute		+		+		+			+	+
sharp				+					+	

¹¹ Note that *f*, *v* and *h* are not included as they were not studied by Leroy (1970). Only simple consonants, and not affricates, were taken into account.

Table 3. Feature analysis of whistled Turkish consonants (shown in upper case). Spoken Turkish counterparts of each consonant are given in parentheses.

Not all these contrasts have been perceptually validated. Some of them may be produced but not perceived. Perception tests with CV and CVC syllables performed by Moles (1970) show relatively poor identification scores for individual phonemes, suggesting that some of the above classes may not be well distinguished, at least in the experimental context.

Thus, the overall system of whistled Turkish is quite similar to that of Silbo. The H1 and signal envelope shapes distinguish vowels and consonants in much the same way in both languages. The only important difference seems to be that whistled Turkish, unlike Silbo, distinguishes velar sounds from labials.

1.4. Other formant-whistled languages

The transposition of vowels and consonants in other formant-based whistled languages appears to obey the same basic principles as in Silbo and Turkish. The Bearnese dialect of Aas was vestigial when it was studied by Busnel and his co-authors (Busnel 1964, Busnel and al, 1962 a and b). Published spectrograms show rising C-to-V transitions for labial consonants and falling C-to-V transitions for coronal sounds as in Silbo and Turkish, and there may be a third locus for velar consonants as in Turkish whistled language. The signal envelope was not studied systematically, but spectrograms show interruptions for voiceless stops as in Silbo and Turkish. Busnel and his coauthors had Turkish whistlers listen to whistled words and nonsense tokens of Aas, and some of them were correctly identified. Thus, it is clear that this whistled language belongs to the same type as Silbo and Turkish.

As mentioned earlier, several formant-based whistled languages are known outside Europe. One of them is based on Tepehua, a Totonacan language spoken in the state of Hidalgo (Mexico) that has been described by Cowan (1972, 1976). This whistled language has many points in common with Silbo. According to this author, "Whistled Tepehua has minute glides in the whistled tone in the transitions into and out of consonant position, much as described by Classe for Silbo Gomero" (Cowan 1976, p. 1404) There are also striking similarities of detail. Cowan (1976, p. 1404) states: "The bilabial /m/ is voiceless [M] in the whistle" [i.e., the whistled language – AR], and again, "alveolar /n/ occurs as voiceless [N]." (Cowan uses the term "voiceless" to refer to a silent period in the whistled signal.) It will be recalled that nasal consonants involved a silent interval in both Silbo and whistled Turkish. Cowan's description of the palatal glide (transcribed /y/) is also noteworthy: "In the whistle, /y/ between vowels (also in Spanish "oye") is usually heard as a high rising glide with very little if any break in the whistling tone" (p. 1405). The same description can be applied to the realization of this glide in Silbo and Turkish (see e.g. Figure 16). Realizations of intonational features (intonation and stress) are also mentioned, but a precise phonetic description is not provided.

1.5. Interim summary

Formant-based whistled languages transpose the signal envelope of the spoken language, which forms the frame of the whistled message. To this frame is associated a H1 contour which transposes consonant loci and F2 contours of vowels. The signal envelope also encodes differences between various manner of articulation types. Vowel and consonant

systems both involve categorization into a small number of contrasting units. The phonological inventory involves unary features, organized hierarchically. These features are probably best defined in acoustic terms, as there is considerable variation in their articulation from one whistled language, or even speaker, to another.

2. Tone-based whistled languages

We first introduce some background information (§2.1), and then examine whistled languages based upon Hmong (§2.2) and Moba (§2.3).

2.1. Background information

2.1.1 Distribution

Most known tone-based whistled languages are based upon tone languages spoken in Africa, Asia, and Mesoamerica. Providing an exhaustive list of whistled languages on these continents is not possible. In the case of Africa, the situation is quite confusing. It is clear that many whistled languages remain to be discovered. The author has recorded two previously unknown tone-based whistled languages of Africa based on Moba and Gulmancema, both spoken in northern Togo and eastern Burkina-Faso. Whistled languages in this part of Africa are only used marginally as means of long-distance communication, since drummed languages assume(d) this function more generally. Other African whistled languages mentioned in the literature include Diola (Moreau 1997), Lobi, Kasem, and Ewe (Thomas 1995).¹²

There is also a great lack of documentation concerning whistled languages in Asia. The better-known examples include those based on Hmong (Southeast Asia) and Chepang (Nepal), among others, but one can suppose that many more are still undocumented. In the Americas, the presence of tone-based whistled languages was first mentioned by SIL linguists working on native American languages, especially in Mexico where the whistled language based on Mazateco is quite well known. Whistled languages are also found in Oceania. A full bibliographical review has recently been provided by Thierry (2002).

We will show here that tone-based whistled languages not only transpose the tones of the source language, as has been previously observed, but also its signal envelope. This is what is common to both types of whistled languages: both of them make use of a signal envelope-based "frame".

2.1.2. Earlier studies

Although there are many references in the literature to tone-based whistled languages in America, Asia, Africa, there are few detailed studies, and almost no phonetic work. This lack might be explained by the fact that these languages seem quite simple: it is quite obvious that they transpose tonal realizations and rhythm. Thus, our main question will be: do these languages transpose anything else?

The first study of a tone-based whistled language, Mazateco, was published by Cowan (1948). Cowan observed that "The whistle is obviously based upon the spoken language. It follows the same tonal system with regard both to registers and to lexically, morphologically,

¹² In some (mainly second-hand) sources, there is a confusion between true tone-based whistled languages and the whistling of tones commonly practiced by linguists and their language consultants in doing fieldwork. We consider only the former here.

and syntactically significant glides" (p. 1390). He found no transposition of consonants or vowels, and concluded: "In whistled languages, the absence of segmental features gives opportunity for ambiguities". The few spectrograms of this language published later by Busnel and Classe (1976) do not reveal any consonantal cues and so tend to confirm Cowan's assumption.

The detailed description of the Kickapoo whistled language by Voorhis (1971) introduces the idea that a tone-based whistled language can encode consonantal cues as well as tones. Kickapoo is a Central Algonquian language spoken in Kansas, Oklahoma, Texas and Mexico; it is the Mexican variety that has a whistled language. In Voorhis' notation, an *o* represents each mora of a note, a *q* represents a glottal transition between moras, and an *h* represents a plain transition. The symbol *q* corresponds to stop consonants or clusters containing a stop, and *h* to other consonants or clusters.¹³ This distribution is not surprising at this point in our analysis. We hypothesize that Kickapoo uses the signal envelope to convey basic consonantal contrasts in much the same way formant-based whistled languages do. (The Chepang whistled language of Nepal, studied by R. C. Caughley (1976), also presents consonantal cues, but as it has some atypical characteristics we will leave it aside in this paper.)

Other studies such as Dugast (1955) and Moreau (1997) consider only tones, and deal mainly with sociolinguistic and pragmatic aspects of whistled communication. Besides written documentation, there exist recordings of music from various parts of the world which sometimes include excerpts of whistled speech. Thus, Thierry (2002), for example, has found recordings of the whistled language of the Wayãpi (French Guyana) and the Lobi (Burkina-Faso).

2.2. Notes on the Hmong whistled language

We now consider some data from two whistled languages, one from Asia (based on Hmong) and the other from Africa (based on Moba) to illustrate how these whistled languages transpose tonal realizations and, also, amplitude modulations.

Hmong is spoken by 2.5 million people who are now widely dispersed in China, Vietnam, Laos, Thailand, Burma, continental France and French Guiana and the USA, mainly as a result of wars. Despite this dispersion, Hmong is less diversified than might be expected. Various dialects are recognized, including "White Hmong" or White Meo, the most important group in Thailand, Laos, continental France and French Guiana, upon which the whistled language is based.

White Hmong is a tone language with seven tones : Tone 1 (level-high, 55¹⁴), Tone 2 (level-mid 33), Tone 3 (level-low 22), Tone 4 (low glottalized 21?), Tone 5 (mid-rising 24), Tone 6 (high-falling 52), and Tone 7 (murmured falling 42). It has a rather simple vowel system with six vowels, /i u e ɔ a/ and five diphthongs, but a very complex consonant system with some 58 consonants,¹⁵ including 28 nasal sounds (5 voiced nasal stops, 5 voiceless nasal stops, 9 prenasalized stops, and 9 prenasalized aspirated stops). The writer participated in a preliminary study of the Hmong whistled language in 1988, published in Busnel et al. (1989). Recordings were made in France and in French Guiana, where a

¹³ The initial consonants were not transposed in the whistled language, according to Voorhis.

¹⁴ The numbers refer to Chao's scale, which is commonly used to transcribe tone levels in East Asian languages. This scale is divided within 5 levels, with 1 being the lowest one and 5 the highest one.

¹⁵ For more information see Niederer (2001). We thank B. Niederer for her help with the Hmong corpus and for her interest in this language.

community of 1600 Hmong settled at the end of the Vietnam war. These communities knew the whistled language but none were using it. The corpora comprised 20 spoken and whistled messages.

Busnel et al. (1989) showed that the whistled language reproduced the tone melody, but the existence of phonetic cues for consonants was mentioned only in a footnote. We return to this question here.

As an example of an utterance in the Hmong whistled language, let us consider spoken and whistled versions of the following sentence:

[ko⁵² j j⁴⁵⁴] tua⁵² ti^{21?} no³³ pe⁵⁵ mu²² tua³³ ηqai⁵² o²²
 [you interjection] come away there we go kill game interjection

'[You , eh] Come, let's go hunting, eh.'

The corresponding waveforms and narrow-band spectrograms are shown in Figure 18

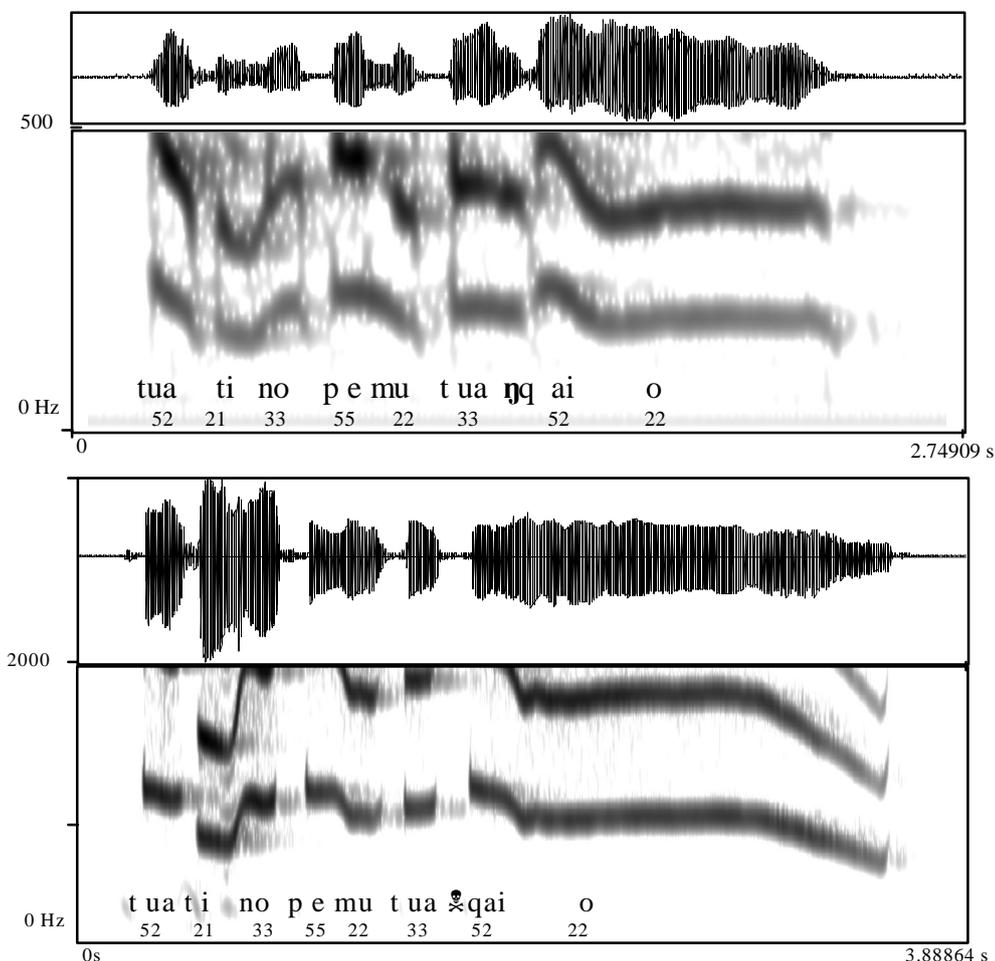


Figure 18. Waveforms and narrow-band spectrograms of a Hmong utterance. The spoken version is shown at the top and the whistled version at the bottom. Tone realizations are indicated below the segmental transcription

A comparison of the whistled and spoken versions confirms that it is the tone melody of the spoken utterance, rather than its F2 movements, that is transposed in whistled speech. Thus, the H1 movements of whistled speech (bottom) follow the F0 contour of the spoken utterance (top) and bear no relation to F2 movements. Similarities between the two waveforms can be observed as well; the general pattern of alternation between intensity peaks (corresponding to vowels) and intensity dips (corresponding to consonants) is similar in both. Thus, as in formant-based whistled languages, the amplitude modulation frame is transposed. Tonal realizations are temporally aligned with this frame. This alignment is even more precise than in the spoken version as there is no disturbance from consonant realizations. For example, the waveform and the spectrogram of the whistled syllable *ŋqai* show a sharp synchronization between the beginning of the whistled block and the onset of the F0 contour.

The reader may also have noticed finer points of agreement between the signal envelopes of the spoken and whistled consonants. In both, *n* and *m* are "continuous" sounds, while the stop consonants *t*, *p* and *q* introduce an interruption. In a set of 86 consonants¹⁶ in this corpus, we found that the voiced nasals, laterals and glides were regularly realized as continuous sounds, while all others, which form the natural classes of obstruents and voiceless sounds, were realized with a silent period. This difference between "continuous" and "interrupted" consonants is clear-cut and precisely implemented, without any slight period of silence within the continuous consonants, as we found in Silbo, for example. This accuracy might be related to the fact that the coding of consonants by a diversity of signal envelope shapes in a tone-based whistled language might interfere with the recognition of contour tones, which require precisely-marked beginnings and endings.

Thus, in the Hmong whistled language, as in the languages examined earlier, the signal envelope provides the basic dynamic frame. Tone realizations transposed from speech are aligned with this frame. The signal envelope also encodes some cues to manner of articulation, introducing, at least, a binary contrast between "interrupted" and "continuous" consonants. Additional studies would be needed to capture completely the consonant cues.

2.3. Some notes on the Moba whistled language

Moba is spoken in northern Togo. It belongs to the Gur division of Niger-Congo languages and is closely related to Gulmancema (Ncam). This language has a rich tonal system with four tones and downstep, as described in Rialland (1983). We discovered the whistled language in the course of our work on the tonal system. At the time we recorded it in Bombouaka in 1981, it was used only by children and adolescents. It was employed as a substitute for drummed language, and was also used as a secret language to communicate without being understood by the missionaries who ran the school. We don't know whether it is still in existence.

We recorded whistled exchanges, songs involving whistled language, and fifty spoken and whistled Moba proverbs taken from Reinhart (1974). Each of these whistled proverbs was followed by the spoken version, and each word was also whistled independently. A group of eight subjects between the ages of eight and fourteen was involved in the session. Many proverbs were recorded by two or three whistlers. Due to time limitations we were unable to run perception tests. As expected, we found that the Moba whistled language transposes both tones and the dynamic frame. This is illustrated by the example in Figure 19.

¹⁶ We left aside consonants whose transcription was unclear.

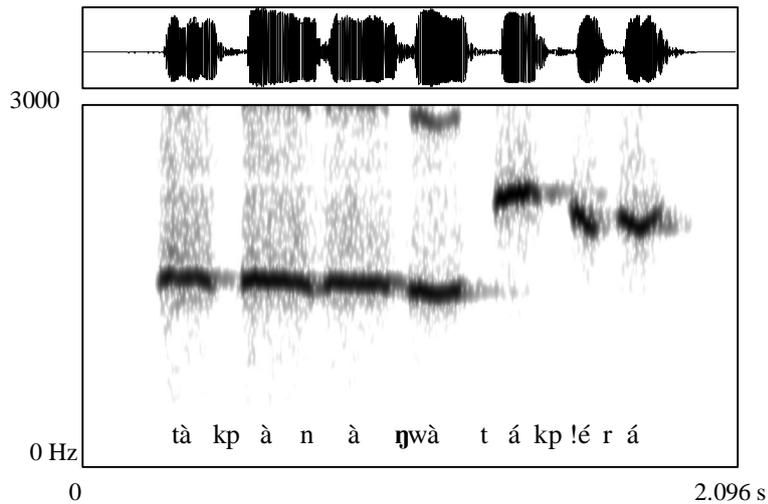
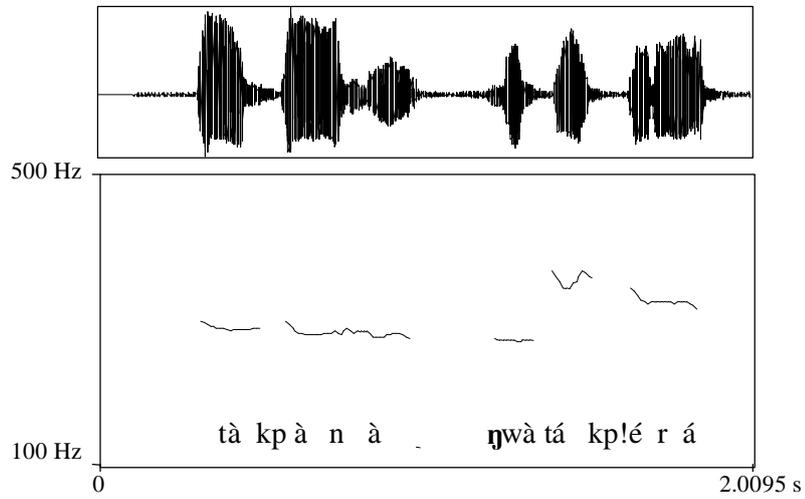


Figure 19. Waveform and melodic curves of the Moba utterance *tàkpàrà wà tákp!érá* "The young ones follow the footsteps of their elders". The spoken version is shown at the top and the whistled version at the bottom

This example shows that both tone (including downstep) and the dynamic frame are transposed into the whistled language. Some cues to manner of articulation can be seen in the envelope of the whistled signal; for example, nasals (*n* and) are continuous with a short dip (0.37s in *n*, 0.56s in), while voiceless stops (*t* and *kp*) introduce a period of silence, or near silence (between 0.10s and 0.12s). Note also that in this example, the *r*, with its sharp dip, mimicks the acoustic effect of the spoken tap.

In the fifty spoken and whistled proverbs that we studied, we found that liquids, nasals and glides were regularly realized as "continuous" or near-continuous sounds, while obstruents, either voiced or voiceless, were realized as "interrupted" sounds. These patterns seem basically the same as those of Hmong, and perhaps Kickapoo. Of course, further study with better recordings as well as perception tests would be necessary to obtain a fuller account of these consonant distinctions.

One intonational feature of whistled Moba merits closer study: question intonation. In the spoken language, question intonation is realized as a lengthening of the final vowel associated with a gradual opening of the glottis (Rialland 1985).¹⁷ Quite expectedly, we found a similar realization in the whistled language, with glottal opening encoded as a progressive decline in intensity.

Thus, the Moba whistled language has many points in common with that of Hmong, showing a similar dynamic frame and tone realizations, as well as transposing the distinctive profile of question intonation.

2.4. Tone based-whistled languages: conclusion

To summarize, tone-based whistled languages encode two prosodic components of speech: the dynamic frame, characterizing syllables and incorporating certain consonantal cues, and tones. They differ from formant-based whistled languages in the extent of segmental coding, which appears to vary somewhat from one language to another. Like formant-based whistled languages, they have an amplitude modulation frame with which the realizations of another "tier" – in this case, tonal rather than segmental -- are aligned. Their phonological system conveys the full tonal system of the language, but is very poor in terms of vowel and consonant distinctions: while vowels and consonants are well distinguished as a class, most consonant and all vowel distinctions are merged.

Note that this type of whistling is closely related to the whistled rendition of the tones, practiced by linguists and their consultants in the study of tone languages or to the whistling of a tune, which are also based on the reproduction of melodic contours, aligned with an amplitude frame. However, compared to whistled language, the amplitude frame of these whistlings is simpler with a simple alternation between intensity peaks and silences, as there is basically one whistled blow per syllable and a silence between each whistled blow (see Rialland 2003 for an analysis of the syllable in whistled renditions of tones).

3. General conclusions

This paper has offered two main contributions to the study of whistled languages. First, developing the phonological perspective initiated in the work of Trujillo (1978), this paper has shown that whistled languages have a phonological structure of their own. Whistled sounds are distinguished in terms of features. It appears that these features are best

¹⁷ The same type of question intonation pattern is found in other Gur and West African languages (Rialland 2004).

defined in acoustic terms as there is considerable variation in articulation from one speaker or language to another. In Silbo Gomero, the features [continuous], [gradual decay], [acute], and [sharp] are sufficient to account for all reliably-perceived distinctions. These features can be organized into a hierarchy deriving the main patterns of neutralizations, with robust distinctions appearing at the top and more fragile distinctions at the bottom.

Second, from a phonetic point of view, it has been shown that in addition to melodic variation, amplitude modulations play a central role in all whistled languages. These modulations provide crucial landmarks for the alignment of information from the formant tier (in formant-based languages) and the tonal tier (in tone-based languages). They function as the basic "carriers" of the structure.

REFERENCES

- Busnel R. G. (1964). *Documents sur une langue sifflée pyrénéenne*, film, S.F.R.S., Paris
 - (1967). *Documents sur une langue sifflée turque*, film S.F.R.S., Paris
 - (1968). *Etude radiocinématographique d'un siffleur turc de Kuskoy*, S.F.R.S.
 - (1970). Historical briefing of the whistling language of Kusköy, *Etudes sur la langue sifflée de Kusköy en Turquie, Revue de Phonétique appliquée* **14-15** reprinted in Sebeok and Sebeok (eds.) 1029-1030
- Busnel R. G., Moles A. & Vallancien B. (1962) Sur l'aspect phonétique d'une langue sifflée des Pyrénées Françaises. In *Proceedings of the Fourth International Congress of Phonetic Sciences Helsinki (1961)*. 533-546, Mouton & Co. reprinted in Sebeok and Sebeok (eds.) 900-913
 - (1962 b) Un cas de langue sifflée utilisée dans les Pyrénées Françaises, *Logos* **5. 2.** 76-91. reprinted in Sebeok and Sebeok (eds.) 881-899
- Busnel René-Guy & Classe A. (1976) *Whistled Languages*. (Communication and Cybernetics 13). Springer-Verlag, Berlin.
- Busnel René-Guy, Alcuri G., Gautheron Bernard & Rialland Annie (1989) Sur quelques aspects physiques de la langue à ton sifflée du peuple H'mong, *Cahiers de l'Asie du Sud-Est* **26** 39-52
- Carreras M, Lopez J., Rivero F & Corina D. (2005) Neural processing of a Whistled language, *Nature* **433** 31-32
- Caughley R. C. (1976) Chepang whistle talk. In Sebeok and Sebeok (eds.) 997-1022
- Charalambakis C. (1994) A case of whistled speech from Greece. In I. Philippaki-Warbuton et al (eds.) *Themes in Greek linguistics : papers from the first international conference on Greek linguistics*, J. Benjamins, Amsterdam/Philadelphia, 389-396,
- Classe A. (1957) Phonetics of the Silbo Gomero. *Archivum Linguisticum* **IX.1** 44-61
 - (1963) Les langues sifflées, squelettes informatifs du langage. In Moles A. and Vallancien B. (eds.). *Communications et Langues Paris Gauthiers-Villars* 129-139
 - (1999) *Le langage sifflé de Gomera* in "Les langues du monde". *Pour la science* 127-129
- Cowan G. M. (1948) Mazateco whistle speech *Language* **24** 280-286
- Cowan G. M (1971) Segmental features of Tepehua whistle speech. *Proceedings of the Fourth International Congress of Phonetic Sciences*, Montréal (1971), Mouton & Co Montréal 695-698
- Cowan G. M. (1976) Whistled Tepehua. In Sebeok T.A. and Umiker-Sebeok D. J. (eds.), *Speech surrogates : drum and whistle systems* The Hague-Paris Mouton 1400-1409.
- Dugast Idelette. (1955) Le langage sifflé ou tambouriné chez les Banen, excerpted from *Monographie de la tribu des Ndiki (Banen du Cameroun), Travaux et mémoires de*

- l'Institut d'Ethnologie* **50** 567-602. Appendix I reprinted in Sebeok and Sebeok (eds.) 708-741
- Jampolsky Marc (1999) *Les derniers siffleurs de Gomera*, documentary film. Broadcast on TV5 (France) on 11/11/1999.
- Lajard M. (1891) Le langage sifflé des Canaries *Bulletin de la Société d'Anthropologie de Paris* **4. 2** 469-483, reprinted in Sebeok T.A. and Umiker-Sebeok D. J (eds.) 954-965
- Leroy Christine (1970) Etude de phonétique comparative de la langue turque sifflée et parlée. In *Etudes sur la langue sifflée de Kusköy en Turquie, Revue de Phonétique appliquée* **14-15**, reprinted in Sebeok T.A. and Umiker-Sebeok D. J (eds.) 1128-1167
- Moles A. (1970) Etude sociolinguistique de la langue sifflée de Kusköy. In *Etudes sur la langue sifflée de Kusköy en Turquie, Revue de Phonétique appliquée* **14-15** reprinted in Sebeok and Sebeok (eds.) 1090-1127
- Moreau M.L (1997) La communication sifflée chez les Diola (Casamance Sénégal). *Divers Cité Langues*, www.quebec.ca/diverscite
- Navarro Tomás T. (1970) *Spanische Aussprache-Lehre* München. Max Hueber Verlag
- Niederer Barbara (2001) La langue Hmong, in *Langues de Guyane, Amerindia* **26/27** (special issue)
- Rialland Annie (1983) Le système tonal du moba comparé à celui du gurma. In Jonathan Kaye, Hilda Koopman, Dominique Sportiche & André Dugas (eds.) *Current Approaches to African Linguistics* 2.5. Dordrecht. Foris 217-234
- (1985) Le fini/ l'infini ou l'affirmation/ l'interrogation en moba (langue du Nord Togo), supplement of *Studies in African Linguistics*, suppl 9 258-261
- (2001) «Une «ponctuation vocalique» en gulmancema (langue gurma) et les conséquences de sa perte dans une langue proche : le moba». In N. Nicolaï (ed.), *Leçons d'Afrique : Filiations, ruptures et reconstitutions des langues, en Hommage à Gabriel Manessy* Louvain Peeters.91-102.
- (2001) Silbo Gomero, Coloquio Nacional, Associação Portuguesa de Linguística, Lisboa
- (2002) Typologie des langages sifflés, paper presented at *Typology and sound universals*, Nantes, Mars 2002
- (2003 a) La syllabe ou la more en tonologie africaine ou comment se fait l'interface entre segments et tons, in Jean-Pierre Angoujard & Sophie Wauquier-Gravelines (eds.) *Phonologie : champs et perspectives*, ENS Editions, Lyon
- (2003 b) What types of whistled languages are there?, invited paper presented at the *The 1st International Conference on Whistled Languages*, La Gomera, April 2003
- (2003 c) "A New Perspective on Silbo Gomero", in M.J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona Rundle Mall : Causal Productions. 2131-2134
- (2004) Question prosody: an African perspective, *Prosodic Typology*, European Science Foundation Workshop, Lisboa
- Reinhart P. (1974) *Proverbes moba* Bombouaka Togo
- Sebeok T.A. & Umiker-Sebeok D. J. (1976) *Speech surrogates : drum and whistle systems*, The Hague-Paris. Mouton
- Shannon R. V, Zeng F.G, Kamath V, Wygonski J. & Ekelid M. (1995) Speech recognition with primary temporal cues, *Science* **270** 303-304
- Stevens K. (1994) Phonetic evidence for hierarchies of features. In P. Keating (ed.) *Phonological Structure and Phonetic Form, Papers in Laboratory Phonology III*, Cambridge (Ma.). Cambridge University Press. 235-242
- Thierry Elise. (2002) *Les langages sifflés* mémoire de DEA en philologie et linguistique. Ecole Pratique des Hautes Etudes

- Thierry Elise. and Rialland Annie., to appear, Variété des usages et typologie des langages sifflés. In S. Wauquier (ed.). *Typologie des sons du langage* Presses Universitaires de Rennes
- Thomas A. (1995) Whistled speech *Linguist List* 6.1319, September 27, 1995
<http://www.linguistlist.org/issues/6/6-1319.html>
- Trujillo Ramón (1978) *El Silbo Gomero : análisis lingüístico* Tenerife Editorial I. Canaria - I. Andrés Bello
- Van Tasell D., Soli S., Kirby V. & Widin G. (1987) Speech waveform envelope for consonant recognition, *J.A.S.A.* **82.4** 1152-1161
- Voorhis P.H. (1971) Notes on Kickapoo whistle speech *International Journal of American Linguistics* **13** reprinted in Sebeok T.A. and Umiker-Sebeok D. J (eds.). 1426-1433
- Wilken G.C (1979) Whistle speech in Tlaxcala (Mexico) *Anthropos* **74** .881-888
- Xiromeritis N.-H. & C.Spyridis (1988-1989) A whistling language in the village Antias in the Greek island of Evia *Glossologia* **7-8** 219-224

Appendix: Results of a preliminary perception test for consonant identification (whistlers L and A)

Silbo is based on Spanish, in its Canarian and standard varieties. A perception test was designed to test consonant identification. The consonant phonemes which occur in the test were /p, t, k, ch, b, d, g, f, s, m, n, ñ, l, r, rr, j/, representing an important subset of Spanish phonemes. A few others such as /z/ and /χ/, were not included in the test, for various reasons (mainly skipping or unexpected readings of some of the cards which were provided).

The test items were nonsense VCV tokens. Nonsense tokens were used to control for possible interference from lexical frequency effects. The VCV context was chosen because of its high frequency and because it provides transitional cues on both sides of the consonant, providing favorable circumstances for recognition. Two whistlers, L and A, were selected for this test on the basis of their skill. Both were considered by other whistlers in the community as very good whistlers. These subjects were asked to stand fifteen meters apart on a hillside. During the testing, each subject was recorded independently, L with a DAT Tascam and A with a minidisk recorder. Parts of the test were filmed and can be viewed in the documentary film mentioned earlier (Jampolsky 1999).

The test tokens, such as *ota, iti* were written on cards. The cards were presented one at a time in random order to L, who whistled them as each card was presented. After each production, A was asked to repeat the whistled utterance and to say aloud what Spanish VCV sequence she had heard. At this point, L stated whether the intended utterance had been correctly identified or not. The test itself was preceded by a short training session. The test could not last longer, as Silbo whistling is exhausting. Due to time and budget limitation, it could not be done a second time with another couple of whistlers with the same level of skill.

Table 4 summarizes the results of the test.

	p	β	f	m	t	ð	n	s	t	l	r	rr	j		k	γ
p	7															
nt														1		
nd					1									1		
					1											
N		1			1											2
A/ UN																

Table 5 Confusion matrix of the consonants in our perceptive test. "Produced" consonants are displayed horizontally and "perceived" consonants vertically. (Numbers in bold correspond to correct identifications.)

Let us consider the results for consonant identification. The overall percentage of correct identification is 57% (39 correct answers out of 68 answers or non-responses). This score is a fairly good one for a whistled language, as far as we can judge from previous tests on other whistled languages: between 37% to 50% correct identification for CV tokens in Aas, between 10 and 25 % in Kusköy, depending upon the whistlers. However, the tests differed in the nature of the stimuli (which included CV tokens and some CVC tokens), and the overall scores included vowel identification, so that a direct comparison is not possible. Table 4 shows also that identification varies considerably depending upon the phonemes: /t/ and /p/, for example, are well recognized while /r/, /rr/ or /l/ did not get correctly identified even once. A more detailed analysis of the depending upon the phonological features is given in the body of this article (see section 1.2.4.).

This test was primarily designed to test consonant identification. However, it also provides instructive data concerning the identification of the vowels *a*, *o* and *i*, which we present in Table 6.

	i	e	a	o	u
i	15	1			
e			1		
a	3	1	79	4	
o			5	15	3
u				1	2

Table 6. Confusion matrix of the vowels in our perceptive test. Produced vowels are displayed horizontally and perceived vowels vertically. (Numbers in bold correspond to correct identifications)

i is a well recognized vowel, which was expected. The interesting point here is the number of correct identifications of *a* and *o* and the relatively small number of confusions between them (see section 1.2.3. , for further comments).