# A New Perspective on Silbo Gomero

**Annie Rialland**
Laboratoire de Phonétique et Phonologie (UMR 7018)
CNRS / Sorbonne-Nouvelle, Paris, France
E-mail: rialland@idf.ext.jussieu.fr

## ABSTRACT

Silbo Gomero is a whistled encoding of the variety of Spanish spoken on the island of La Gomera in the Canarian archipelago. The purpose of the present study is to determine how and to what extent the whistled language encodes the phonemic structure of spoken Spanish. It shows that vowel qualities and consonantal place of articulation distinctions are transmitted primarily by F2-to-H0 transposition, while consonantal manner distinctions and rhythmic information are conveyed by modulations in the signal envelope.

## 1 INTRODUCTION

Although whistled languages, including Silbo Gomero, attracted some attention in the 1970s ([1], [2], [6]), they have been generally ignored since then. The present study is based on new data collected on the island of La Gomera in 1998.

The purpose of the present study is to show how and to what extent this whistled language encodes spoken Spanish. Silbo Gomero is based upon acoustic cues extracted from speech. The main acoustic features of spoken Spanish as encoded in Silbo are :

- F2 trajectories, which form the basis of the whistled contour, as previously observed by Brusis [1], Busnel and Classe [2]. It will be shown here that Silbo does not simply mimic second formant (F2) movements, but transposes consonant loci.
- The signal envelope, which conveys further phonemic information. This part of the study is the newest one, as the role of the signal envelope in whistled languages has not yet been investigated systematically.

We shall show how acoustic cues taken from speech (signal envelope modulations, F2 contours of vowels, consonant loci) are transposed into a global structure.

Until about forty years ago, Silbo Gomero was used as a means of long-distance communication by most of the inhabitants of the mountainous interior to convey a wide variety of messages (indeed any message at all, according to La Gomera residents). Subsequently the language became dramatically endangered, dwindling to less than 50 fluent practitioners five years ago. Recently action has been taken to preserve it as part of the communal patrimony of La Gomera. It is currently undergoing a revival and has been introduced into the primary school curriculum, where it has been a required subject since 1999.

The primary aim of this paper is to improve our understanding of the relationship between spoken Spanish and Silbo. We will first consider the realization of vowels and consonants, and then draw general conclusions on the structure of Silbo. Focus will be placed on the transposition of F2 trajectories and modulations of the signal envelope.

## 2 DATA

The data discussed in this paper is taken from recordings and films of whistled exchanges in various situations, including classroom instruction in Silbo, interviews with several fluent whistlers (primarily three), and a perception test involving two subjects. The materials were collected in La Gomera in November 1998. Some of them were shown in a 50-minute TV documentary "The Last Whistlers of La Gomera" broadcast on France's TV 5 [3]. The interviews were designed in part to determine which Spanish phonological oppositions are considered easy, difficult, or impossible to transpose into Silbo. Teaching strategies were also taken into account in determining the scale of difficulty. It was found that whistlers are quite conscious of this scale and refer to it to avoid ambiguities in composing messages.

## 3 VOWELS IN SILBO

Let us first consider the realizations of vowels. The whistled signal is poorer in information than the speech signal as it consists mainly of a single sinusoid whose frequency fluctuates between 1000 Hz and 3000 Hz, with or without harmonics. The challenge is to encode vowels as well as consonants in this signal with minimal loss of distinctivity. First, let us consider the realization of the five Spanish vowels as produced in sequence.

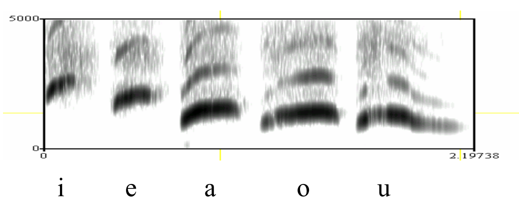


**Figure 1:** Spectrogram of whistled *i, e, a, o, u* produced in sequence

 ISBN 1-876346-48-5 

At first glance, we notice that the first harmonic (H0) of Silbo presents a descending pattern from *i* to *u*, paralleling that of F2 in the spoken language. This relationship between H0 and F2 has been observed previously [1, 2]. The values of Silbo H0 and Spanish F2 are not identical for all vowels, however.

We find a good match between Silbo H0 and the F2 value of spoken *a*, both located around 1500 Hz. In Spanish the vowel *a* falls in the midst of the vowel system in terms of its F2 value. In Silbo it provides a reference point for whistlers, who begin most of their messages with an *a* which regularly precedes the name of the addressee: *a Bernardo, a Maria, a Sebastian, a Domingo.* The *a* functions rather like the *la* to which the members of an orchestra tune their instruments..

Consider next the two front vowels *i* and *e*. The H0 of Silbo *i* tends to have a higher value than the F2 of its spoken counterpart, while the corresponding values of *e* are roughly the same. The rounded vowels *o* and *u* have much higher H0 values than the F2 values of their spoken counterparts; however these vowels are not distinguishable in Silbo. Their merger can be explained by articulatory and acoustic constraints on whistling: the F2 value of spoken *u* cannot be reproduced as a harmonic in Silbo because its value is below the range of whistled speech, and of whistling in general (between 800 Hz and 3300 Hz).

Up to this point, we have commented only on the first harmonic (H0), though other harmonics can be seen in Figure 1. Considering the full spectrum of these vowels (which is not reproduced here), we conclude that the other harmonics do not play a role in distinguishing vowels, since their amplitude progressively decreases, only that of H0 being prominent enough to hear under normal conditions of transmission.

Recall also that Silbo Gomero is a whistled language designed for long-distance communication. It was commonly used in the past to convey messages over distances of 2 or 3 kms and could be heard as far as 8 kms away. Our recordings made at a distance of 1.5 kms did not show any harmonic other than the first one (H0). Considering recordings over various distances, Busnel and Classe [2] also concluded that only the first harmonic can be perceived across the distances at which Silbo Gomero was normally used.

A comparison with other whistled languages, and more specifically with those based on tone languages such as Mazateco [2] or Moba [4], is relevant. This second type of whistled language transposes F0 contours, and there is no relationship between whistling and vowel formants. In these languages, however, the acoustic nature of the whistled signal is the same as in Silbo, with the same type of harmonics.

# 4 REALIZATION OF CONSONANTS

The realization of consonants in Silbo Gomero is based on the transposition of F2 as in vowels, but also involves a transposition of signal envelope modulations. As a first illustration, let us compare spoken and whistled sequences involving *p*. An example is shown in Figure 2.

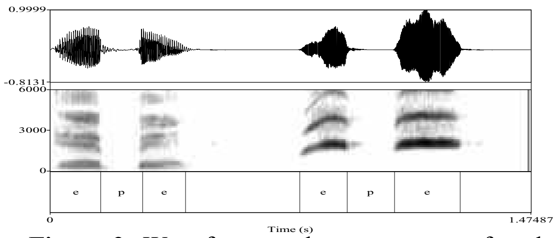


**Figure 2:** Waveforms and spectrograms of spoken and whistled *epe.*

We see that the F2 movements of speech shown in the spectrogram on the left and the H0 contour of the whistled utterance shown in the one on the right are similar: both show a low-frequency locus at each edge of the consonant. Intensity modulations are also parallel in the spoken and whistled utterances: the vowels are intense sounds while the stop occlusion is represented by silence. We can also see that some details of the intensity modulations are transposed with considerable precision, such as the low-intensity periods just at the beginning of the p closure.

We will next see that the relationship between the spoken F2 and the whistled H0 does not involve a simple copy, as suggested by previous authors [1], [2], but a transposition of consonant loci. We will then examine how envelope modulations are transposed from spoken to whistled language and how they are used to distinguish one consonant from another.

## 4.1. CONSONANT LOCI IN SILBO

Let us consider realizations involving the coronal consonant *t.*

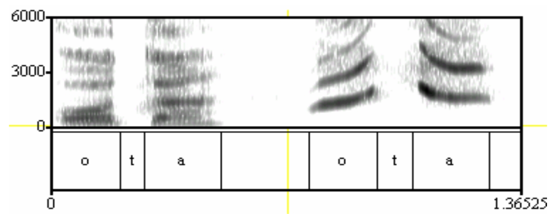


**Figure 3:** Spectrograms of spoken and whistled *ota.*

We see that the transition movements are roughly similar in speech and in whistling. Both rise towards a high-frequency locus, which is typical of a coronal sound. However, we do not always find such a close parallel between F2 and H0 movements, as is shown by the differences in the following tracings:

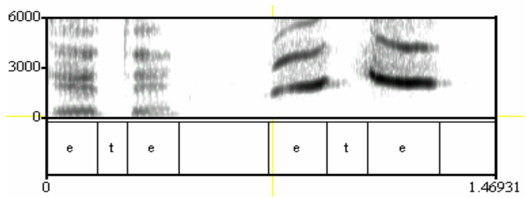


**Figure 4:** Spectrograms of spoken and whistled *ete.*

The F2 contour falls in the spoken utterance shown on the left, while the H0 contour rises in its Silbo counterpart as shown on the right. The same type of discrepancy can also be observed when *t* is flanked by *i* in *iti* (see [4] ).

These discrepancies make sense on the assumption that the Silbo utterance does not copy the F2 movement but has its own coronal locus. This locus, located at 3000 Hz, is slightly higher than the H0 value of *i* and is much higher than the spoken coronal locus at 1800 Hz. This locus has the same effects as a consonant locus in speech: transitions move toward and away from this point, and their length depends on the H0 values of the surrounding vowels. In early work on speech synthesis, the coronal locus was modeled as a unique point located around 1800 Hz, defining the points towards and away from which the which transitions moved. In natural speech, F2 movements are influenced by the vocalic context, and transitions usually undershoot the presumed locus point. In Silbo, the same type of undershoot can be observed; for example, transitions between round back vowels and coronal consonants as in *oto* converge to a lower point than do the transitions between front vowels and the same consonants (as in *ete* or *iti*).

We also observe a low-frequency locus at the edges of labial consonants, as shown in the following spectrogram:

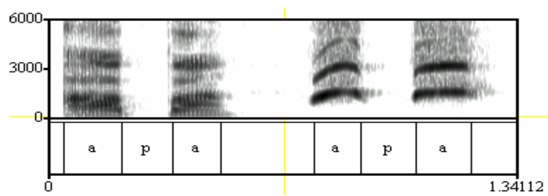


**Figure 5:** Spectrograms of spoken and whistled *apa.*

In this example, as well as in Figure 2 showing realizations of *epe*, we see that transitional movements are parallel in the spoken and whistled utterances. We find a pattern of fall towards a grave locus typical of labial consonants. Let us consider another case, however, that does not show this parallelism.

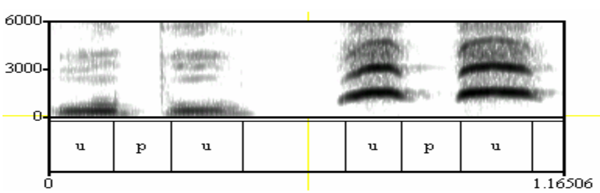


**Figure 6:** Spectrograms of spoken and whistled *upu*.

Here we see a discrepancy between the level F2 of the spoken utterance on the left and the falling H0 movement of the whistled utterance on the right. We see that the whistled transitions fall towards a locus at around 800 Hz which is lower than the H0 value of the lowest vowel. Thus the whistled labial locus is pitched below that of any vowel of the system. It occupies a symmetrical position to the coronal locus in that it is situated below the range of vowel realizations while the coronal locus is located above it. Thus, the coronal and labial consonant loci mark upper and lower bounds on vowel realizations.

As velar consonants are realized in the same way as labial consonants, there is are no further loci for stop consonants. The same coronal and labial/velar loci are found in the realization of other consonants (nasals, fricatives, liquids and glides). An extra-high locus is used for the palatal consonants *(nasal, liquid, glide)* by some whistlers [4].

Thus we conclude that in reproducing consonants, Silbo does not copy F2 trajectories, but transposes consonant loci. We end up with a simple acoustic “geometry” in that consonant loci frame the range of vowel realizations.

### 4.2 THE VOICED/VOICELESS CONTRAST IN SILBO

Now let us consider the realization of the consonants ( *b g p k* ). Figure 7 compares *apa* [apa] and *aca* [aka] with *aba* [aßa] and *aga* [aγa] in which the intervocalic consonants are realized as voiced continuants, following the well-known spirantization rule in Spanish. (It will be recalled that labials and velars are realized similarly.)

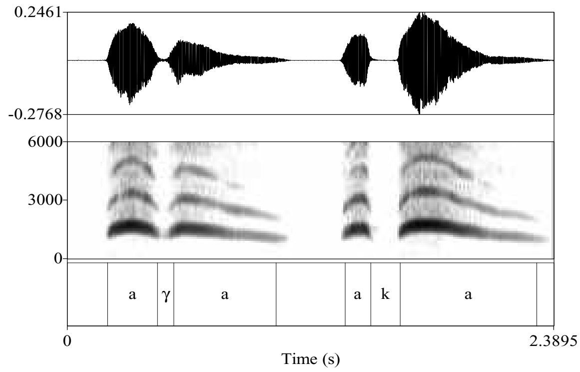


**Figure 7:** Waveforms and spectrograms of whistled [aßa] / [aγa] and [apa] / [aka].

We observe that the H0 is not interrupted in the realization of whistled [aßa] and [aγa] as it is in [apa] and [aka]; instead it is continuous and characterized by a dip in the signal envelope. This dip transposes the shape of the intensity through characterizing the constriction of the voiced continuant. The voiceless stop vs. voiced continuant distinction is encoded by the reproduction of the envelope modulations, as interruption in the voiceless stop and as a dip in the voiced continuant. This distinction is realized in the same way in coronal consonants, the


ISBN 1-876346-48-5 

[aδa] exhibiting the same type of intensity dip as the [aßa] / [aɣa]. Note that, according to our data, it is unclear whether a voiced stop can be distinguished from its continuant counterpart.

Like the voiceless stops, the voiceless fricatives *f, s, χ* and the palatoalveolar affricate show an interruption in the whistled contour. Whistlers consider them difficult to distinguish from voiceless stops as they are all realized with an interruption of H0 and show friction noise only irregularly (see Rialland, ms ). Note that in these cases, the signal envelope is different in speech and Silbo, as spoken fricatives and affricates involve friction noise, triggering a peak in the envelope. This difference suggests that whistlers reproduce the speech signal envelope eliminating its high frequencies.

### 4.3. TAPS, TRILLS AND SOME DETAILED TRANSPOSITIONS OF THE ENVELOPE

Envelope modulations do not only carry over the stop vs. continuant distinction in Silbo. They also transpose tapped *r* and trilled *rr* quite well, reproducing their typical alternation of peaks and valleys. The following figure illustrates the realization of the trilled *rr* in Silbo :

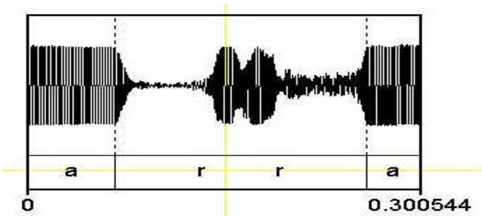


**Figure 8:** Waveform of whistled *rr*.

The envelope also contains cues to the distinction between nasal and voiced stops/continuants, but these cues are unstable and the existence of this contrast in Silbo is debatable. Whistlers consider it a difficult distinction to realize. Laterals also seem to be indistinguishable from nasals. (See [4] for a detailed study of stable vs. unstable oppositions, showing that stable oppositions form the core system of Silbo.)

## 5 CONCLUSION: AN OVERVIEW OF SILBO STRUCTURE

We have shown that envelope modulations transmit important distinctions between classes of phonemes. Several authors have investigated the information-bearing capacity of envelope modulations in speech. On the basis of experiments involving stimuli which preserve the envelope but remove spectral information, it has been shown that while English-speaking listeners do not recognize individual phonemes on the basis of envelope modulations, they do use this information to identify contrasts between classes of phonemes (e.g. voiceless vs. voiced stops, liquids vs. glides, etc.) [7].

Moreover, envelope modulations not only carry cues to phoneme classes, they also carry rhythmic information, which is important not only in the identification of phonemes but also in the perception of entire linguistic messages. Using everyday sentences, Shannon et al. (1995) showed that listeners are able to reach a level of 50% correct sentence recognition using only long-term envelope modulations. Thus, Silbo has selected very efficient cues for transposing spoken language to whistling.

Noteworthy too is the fact that the manner-of-articulation information that can be extracted from envelope modulations is complementary to the information carried by F2 trajectories, since these refer basically to consonant places of articulation and vowel qualities. H0 in Silbo encodes just the latter characteristics. Moreover, in Silbo there is an interesting complementarity between the encoding of vowels and consonants by H0, since as we have observed, consonant loci "frame" the range of vowel realizations.

Additional work on other F2-based whistled languages is needed to determine whether they make use of the same principles of envelope modulation and transposition of F2 trajectories and consonant loci.

## ACKNOWLEDGMENTS

I give special thanks to Marc Jampolsky for giving me the opportunity to discover Silbo.

## REFERENCES

[1] T. Brusis, "Uber die phonetische Struktur der Pfeisprache 'Silbo Gomero' dargestellt an sonagraphischen Untersuchungen," Zeitschrift für Laryngologie, 52.4, 292-300, 1973.

[2] Busnel R.G. and Classe A, Whistled Languages, Springer-Verlag, Berlin, 1976.

[3] Jampolsky M., Les derniers siffleurs de la Gomera , 50mn documentary, broadcast on TV5 in France, 11/11/1999 (with an English and a Spanish version).

[4] Rialland A., "Silbo Gomero revisited," ms, CNRS, Paris, 2003.

[5] Shannon F., Zeng F., Kamath V., Wigonski and Elekid M., "Speech recognition with primarily temporal cues," Science 270, 303-304, 1995.

[6] Trujillo R, El Silbo Gomero : análisis lingüístico, Editorial I. Canaria - I. Andrés Bello, Tenerife, 1978.

[7] Van Tasell D., Soli S., Kirby V. and Widin G., "Speech waveform envelope for consonant recognition," J.A.S.A. 82.4, pp. 1152-1161, 1987.