

Utterance final /S/ in Andalusian Spanish. The phonetic neutralization of a phonological contrast

PAUL O'NEILL
Universidad de Barcelona

1. Introduction

This paper focuses on the phonetic realization and the morphological implications of the sibilant element when in utterance final position in the Spanish spoken in the Western part of Andalusia. One of the most salient differences between this dialect of Spanish and the standard dialect, Castilian, is the treatment of the implosive sibilant. Whilst in Castilian the phonetic resolution is the alveolar fricative [s], Andalusian Spanish is considered to undergo a phonological process whereby an underlying sibilant is “aspirated” when it occurs in the coda of the syllable. In word internal position before obstruents this aspiration is thought to be common to all varieties of Spanish spoken in Andalusia. In utterance final position, however, its resolution differs between what has been termed Eastern Andalusian Spanish (EAS) (Gerfen, C. 2001) and Western Andalusian Spanish (WAS).

In EAS no aspiration takes place and, in its stead, a change occurs in the timbre of the final vowel, which is transmitted to the rest of the vowels producing vowel harmony (Martínez Melgar 1998). However, in WAS it is reported that the aspiration in utterance final position can either be maintained as [h], voiced [] produced very slightly [^h] and even lost completely, without any change in the timbre of the final vowel (Jiménez Fernández 1999). In this position, the underlying /s/ is of great importance for the morphological system of Spanish due to the fact that the sibilant is the marker of plurality on the nominal and that of the second person singular on the verb. In this way the preservation of an aspirated segment provides a means of maintaining these morphological contrasts.

The problem, however, is that, to the author's knowledge, there have been no phonetic studies as to the actual phonetic realisation in this particular position. The term *aspiration* and the symbols for a glottal fricative are used in a perceptive sense and based on the impressionistic judgements of the authors, not on solid phonetic data. As such, to date, it

has neither been proved that there exists any audible friction in this position and, in the case it does exist, that it is being produced at the glottis.

The purpose of this paper is two-fold: firstly, through acoustic analysis, I wish to clarify what the phonetic surface realisation is of this underlying sibilant, in this position, in the Spanish spoken in the Western part of Andalusia; more specifically the cities of Málaga and Seville. Secondly, through a perception test, I aim to discover if this phonetic realisation has come to be distinctive as the marker of plurality in nouns and the second person plural of the verb.

2. Experimental Method

In order to achieve my objective, two experiments were necessary. The first was an acoustic analysis of recorded sounds in which words ending in the sequence Vowel + Sibilant (VS) were spectrographically analysed to determine their acoustic nature and compared with their minimal pairs which contrasted only with the lack of the final sibilant element, that is they ended in a Vowel (V) only. The second experiment was a perception test, in which informants listened to each of the aforementioned words and were asked to make a judgement about whether they could perceive the sequence VS or V at the end of the word. A total number of six informants took part in both experiments, three from the city of Seville and three from the city of Málaga. All informants were adults who were born and had lived all their lives in the respective cities and had not studied at university¹. The problems with the formulation of the two experiments and the experimental method are explained below.

2.1 Acoustic Experiment

The elaboration of the first experiment was a complex matter for both sociolinguistic and technical reasons. With regards to the former reason, it must be stated that one cannot undertake a study of the sibilant element in Andalusian Spanish without taking into account the highly sociolinguistic factors involved with this element. Both in Andalusia and the rest of Spain, there exists the opinion that the Spanish spoken in this region is a “bad” or “degenerated” form of the standard, and thus, when in formal contexts, native speakers very often modify their pronunciation and adopt a phonology which is not their own, treating the sibilants in the coda of

¹ The sex and age of the informants was not a controlled variable.

the syllable in the same way as those in the onset, that is, giving them a fricative pronunciation [s] in order to give the impression of speaking “well”. This tendency is also reinforced by Spanish orthography in which there exists a tight correspondence between phoneme and grapheme and the generally accepted popular principle is that a pronunciation is more prestigious the more it adheres to the orthography.

On account of these reasons, it was not possible both (a) to carry out the recordings in a phonetics laboratory since the formality of the occasion would undoubtedly interfere with the naturalness of the informant’s pronunciation and, (b) to use a list of words to be read in order to achieve the sounds in the appropriate context, as the presence of an orthographic <<s>> could effect the outcome.

The method used therefore was the following. The recordings were carried out in a quiet room in the house of the informant, the only people present being the informant and the interviewer. The informant was presented with two series of images. The first were photographs of 14 objects which were to be named in the singular². The second set consisted of the same objects but this time in plural. The informants were also presented with four verbs in the first person singular of the present tense and asked to conjugate them in the following way: *yo hablo, pero tú*____, *y nosotros*____, *y él*____ “I speak, but you____, and we____, and he____”. The words were recorded using a digital recorder (Sony MiniDisc MZ-NH700) and analysed using a speech analyser (SIL Speech Analyzer 1.5). The number of words submitted to spectrographic analysis for each informant was 44: 18 nouns in the singular, 18 in the plural, 4 verbs in the second person singular and 4 in the third person singular. Since six informants took part in the experiment the total number of words analysed was 264.

The analysis and identification of the sounds was also a problematic matter for two reasons: (a) the nature of the sound being analysed (aspiration) and, (b) the nature of the position being studied in the phrase (utterance final). Regarding the difficulty of pinpointing a segment of aspiration acoustically, more will be commented upon this in the discussion (cf.4). Here, suffice to say that, unlike other segments, aspiration appears at no constant frequency but rather as noise around the formants/frequencies of the preceding sound or in the higher frequencies. This causes it to be easily confused with background noise or a sound

² It must be noted that none of these words contained an implosive sibilant inside the word.

produced less intensely. This problem was overcome by establishing criteria for the identification of a possible segment of aspiration³, based on the characteristics of aspiration in other languages, which possess this sound, and also the acoustic characteristics of aspiration found in other contexts in Andalusian Spanish.

With regards to the problem of the spectrographic analysis of sounds which occur in utterance final position, this position, as various phoneticians have pointed out, “...*aunque de validez teórica, no es adecuado para un estudio sonográfico, porque la pérdida de la intensidad hace que los sonidos no queden bien definidos en el sonograma*” (Martínez Melgar 1986). This difficulty was complicated to overcome since, in experimental phonetics, the usual methodology to avoid such circumstances, is to place the word to be analysed in a carrier phrase, thus ensuring it is pronounced with sufficient intensity. This method, however, was considered unsuitable for the present experiment since it would not fulfil the objectives of the study: the analysis of the sibilant in utterance final position in normal speech. If a carrier phrase were used, this would place the sibilant, through word coalescence, either in a prevocalic or preconsonantal position, which in turn would have phonological consequences on its phonetic outcome. Therefore, to ensure good quality spectrograms within the aims of the experiment, optimal recording conditions were sought and the informants were required to produce each word separately. In addition, as a control, a limited number of words were solicited from the informants with an interrogative intonation, in which the words were naturally pronounced with greater intensity, in order to appreciate any differences therein.

2.2 Perception Test

For the perception test, the data recorded for the first experiment, that is the nouns of the 18 objects both in singular and plural plus the four verbs in the 2nd and 3rd person singular, were isolated, separated for any determiner or pronoun, arranged randomly and presented to the very same informants who had pronounced them. This was carried out in a group, in a quiet room of one of the informants, using a digital player (Sony Mini-

³ These criteria were: a prolonged continuation of noise after the last well-defined vocalic pulse, in which the following segment has noise in the high frequencies and noise around the formants of the preceding vowel.

Disc MZ-NH700) and good quality speakers⁴. The number of words was 44 for each informant and since the informants listened to all the recorded words for the three informants of their respective city, the total number of words presented to the informants was 176.

On hearing these words the informants were asked to write on a piece of paper, in the case of the nouns, whether they understood the noun as in the singular (in which case they wrote *1*) or as in the plural (in which case they wrote *2*); in the case of the verbs, they had to note whether they understood the second (in which case they wrote *tú* 'you') or the third person of the verb (in which case they wrote *él* 'he'). In cases in which they were unsure, this was indicated with a question mark.

To avoid the situation whereby the informants, on hearing a series of repeated words, when listening to a word for the second time answered the opposite of what was noted on the first hearing; they were forewarned that many of the words were repeated.

3. Results

3.1 Experiment 1

The results of the acoustic analysis of the last syllable of the nouns and verbs which varied only in the presence or lack of a final underlying sibilant are presented here in (1) as a percentage of the words which for acoustic reasons could be considered to possess a segment of aspiration, in adherence to the criteria for the identification of such a sound. That is the percentage represents the frequency with which utterance final words ending in (a) and vowel only (V) and (b) a vowel and an underlying sibilant (VS) contained traces of aspiration at the end of the word. The results are presented for each informant, via their initials, and each category:

⁴ It ought to be noted that the recorded words were placed over again a number of times to the satisfaction of the informants.

(1) Results of the acoustic analysis of the final syllable of the words

MÁLAGA	MRR		MGR		MMG	
	/VS/	/V/	/VS/	/V/	/VS/	/V/
Percentage	66%	70%	60%	66%	79%	57%

SEVILLE	BRR		CAR		J	
	/VS/	/V/	/VS/	/V/	/VS/	/V/
Percentage	50%	47%	75%	75%	63%	59%

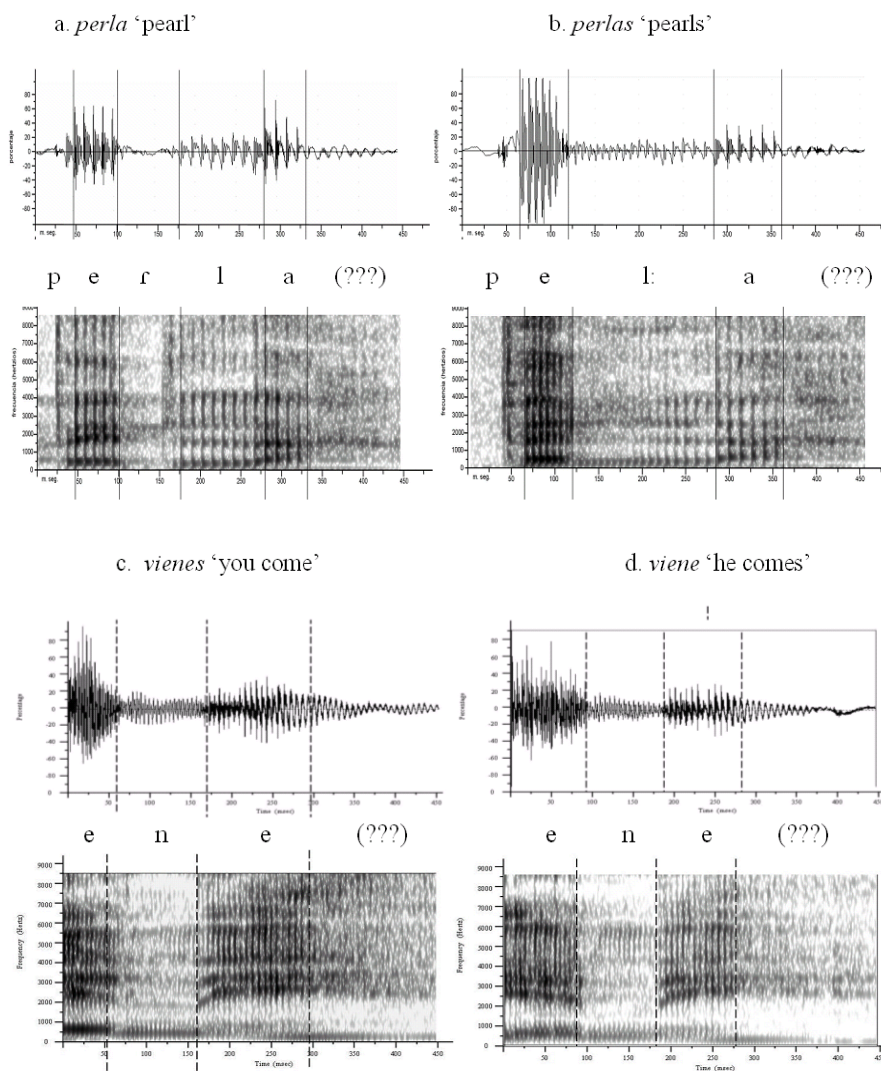
These results show an incidence of “aspiration” in utterance final position which is similar both for those words with the underlying sequence VS and for those which lack the final sibilant element. In none of the cases is the presence of aspiration systematic⁵.

The conclusion to be drawn from these results is that, from an acoustic point of view, there is a neutralisation between the sequences VS and V in utterance final position, brought about by the fact that in this position, for both segments, it is possible that some type of aspiration is produced.

This conclusion can be illustrated by the waveforms and spectrograms in (2). The first ones are of the minimal pair *perla/perlas* ‘pearl/pearls’, the second *viene/vienes* ‘he comes/you come’. In both cases, at the end of the spectrogram there are formants which seem to correspond to those of the preceding vowel, but are weaker and somewhat faded. Also, a considerable loss in intensity can be seen in the waveform. Acoustically, these two realisations are similar to what is considered, in many languages, as a portion of aspirated phonation.

⁵ Although for the informant MMG there seems to be a tendency of a greater occurrence of aspiration in the sequences VS-this is not corroborated by the other informants. In fact, for two of the informants (MRR and MGR) there was a greater incidence of a possible aspiration in the forms without the underlying final sibilant.

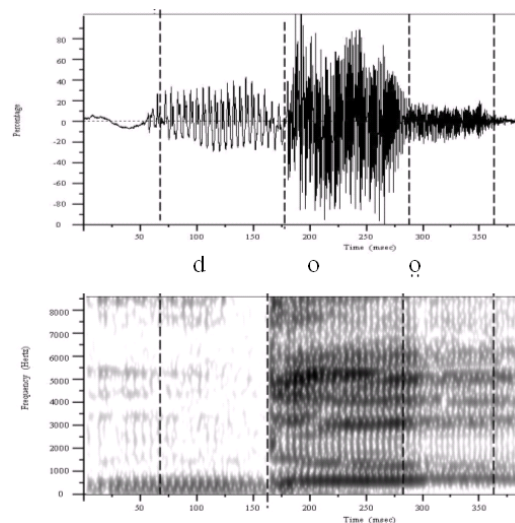
(2) Waveform and spectrograph of the minimal pairs *perla/perlas, viene/vienes* showing that in both cases there seems to be a period of breathy phonation at the end of both words.



In comparison with the results of the words pronounced with a normal intonation are the results of the control words pronounced with an interrogative intonation and thus with greater intensity. In these words (which

contained an underlying final sibilant) all the results pointed to the same realisation. This realisation, which can be seen in the spectrogram in (3), is a clearly identifiable segment on which seems to be, basically, a copy of the preceding vowel, but produced with less intensity and a different type of phonation.

(3) Waveform and spectrogram of the word *dos* “two” produced with an interrogative intonation and showing two clearly distinct vocalic segments.



In the light of this evidence, the conclusion must be that there is a phonological process at work whereby an utterance final sibilant is given a phonetic realisation. However, this can only be detected acoustically when the word is pronounced with greater intensity. When pronounced normally there is no acoustic difference between the underlying sequences VS and V. That is to say, the phonology of the speakers has a phonetic output for a word final sibilant; however, this is only acoustically salient when produced with an interrogative intonation. With normal intonation, although there is acoustic evidence of a possible final segment, this segment also present, for reasons which shall be addressed below, for words ending only in a vowel.

3.2 Experiment 2

The results of the second experiment in which the informants listened to words pronounced by themselves and by their fellow citizens and were asked to note if they understood the sequence VS or V are given in the table below as a percentage of the answers which were correct.⁶

(4) Results of the perceptive test.

a.	Málaga	Correct answers
	MRR	40%
	MGR	46%
	MMG	56%

b.	Seville	Correct answers
	BRR	40%
	CAR	49%
	J	46%

These results are consistent with the acoustic analysis given above. That is, in final position, in normal speech, there is no distinction between the sequence VS and V and therefore, the morphological distinctions which rely on this final sibilant element are lost in this position.

⁶ The answers in which the informants were unsure, signalled by a question mark, were counted as wrong answers, since they represented a case of a lack of distinction between VS and V.

4. Discussion

The results of the experiments are both interesting and perplexing in that, on the one hand, the words produced with an interrogative intonation show acoustic evidence of the existence of a segment which could be classed as some type of aspiration. This suggests that we are dealing with a phonological process upon an underlying sibilant element, otherwise there is no explanation as to why this element would systematically appear in this context and its similarity to aspiration in other contexts in WAS⁷. However, on the other hand, both perceptive and acoustic evidence indicate that, in normal speech, this is not perceived and possibly not phonologically produced since the incidences of aspiration are not only not conclusive but they are also present in contexts which lack an underlying sibilant, i.e. the sequences ending in V. The questions that arise, then, are basically the following:

Why is there aspiration in utterance final position when a word ends only in V?

Why is there consistent aspiration when the sequence VS is stressed but lack of distinction between the final sequences V and VS in normal speech?

The answer to these questions and the key to understanding the results of the experiments, in my opinion, lies in the phonetic nature of aspiration.

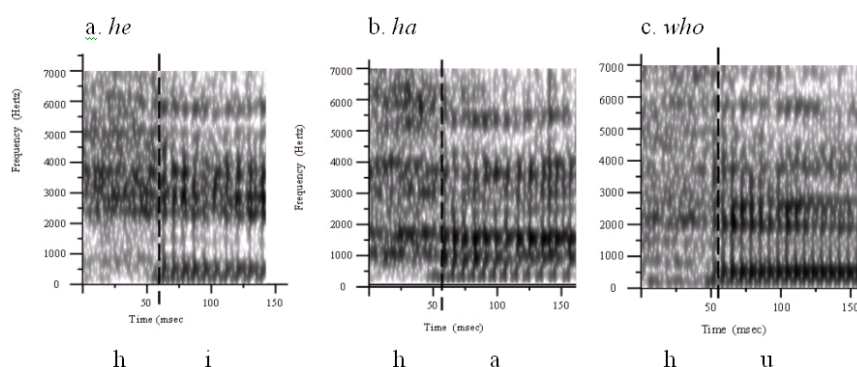
4.1 Aspiration Defined

In Phonetics aspiration refers to the audible breath of air which accompanies the onset or offset of an obstruent gesture and is usually symbolized by the diacritic [^h] in the IPA. However, in the literature, the term aspiration is often used in the absence of obstruents to signify the perception of an audible breath of air and is transcribed using either the symbol for the glottal fricative [h], or [^h] if the aspiration is considered to be less audibly prominent. The tendency seems to be a definition of aspiration based on the perception of a current of air with no prolonged

⁷ By this reference is made to the portion of “aspiration”, as observed by the author, present word internally before voiceless occlusive consonants in words such as *pasta* ‘pasta’, *caspa* ‘dandruff’, *bizco* ‘cross-eyed’.

frication; however, the usage of the symbols suggests that friction is being produced at the glottis. Many phoneticians have raised doubts over the adequateness of the symbol [h] to represent what has been termed aspiration, since in those languages which possess aspiration as a distinctive segment such as English⁸, it is not clear that the audible friction is produced at the glottis. This can be illustrated by the Spectrograms in (5), of various segments of aspiration preceding different vowels in English, in which the visual representation of the sound changes depending upon the following vowel.

(5) Spectrographs of the English words *he*, *ha*, *who* in which the spectrographic representation of the segment of aspiration is different.

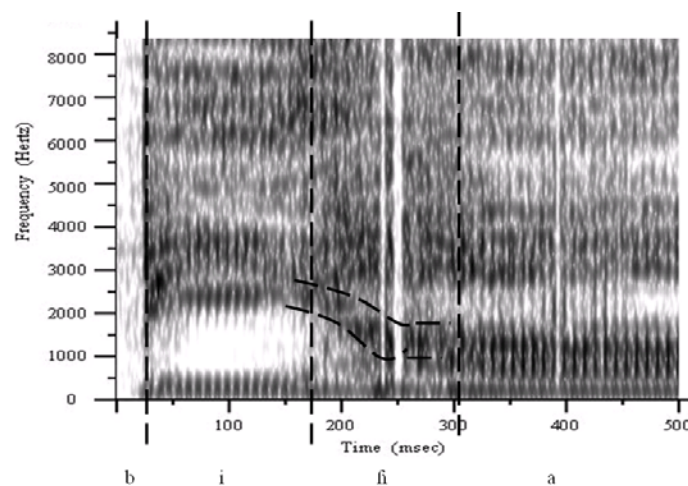


Here in each spectrogram the friction is represented differently. In the case of *he* the noise is concentrated around F_2 and F_3 of the vowel and there is little noise around F_1 ; however, in *ha* the noise concentrated around this formant is rather intense; which, in turn, contrasts with *who*, where the noise is less intense than in the other two cases. Common to all, however, is the presence of noise in the high frequencies and around the vocalic formants. Such data has led to aspiration being characterised as merely a silent vowel or one produced with a different type of phonation (Ladefoged 1971).

⁸ In English the presence of a segment of aspiration is phonemically distinctive as shown by the words *heat*, *heal*, *high*, *hill* and their minimal pairs *eat*, *eel*, *eye*, *ill*.

However, Keating (1998) has shown that what seems to be the defining factor of the English cases of aspiration is that although it is specified with regards to the action of the vocal folds in that a large amount of air ought to be able to pass through them, it is not specified for any other aspect, that is, it poses no oral feature. Therefore, the shape of the oral cavity during the production of this sound is underspecified and, as such, the organs adopt the position of the contiguous sounds (which explains why some considered it to be a merely silent vowel). This does not mean that it assimilates to the point of articulation of a particular surrounding sound but rather its oral articulation is dynamic, changing from the position of the preceding sound to that of the following sound. This is evident in intervocalic position where throughout the duration of the segment the articulators move from the position of the preceding vowel to that of the following one. This is the case in the English word *behind*, given here as a spectrogram in (6) in which one can appreciate how the vocalic formants move from the position of the first vowel to that of the second during the segment of aspiration.

(6) Spectrogram of the English word *behind*, showing the dynamic nature of aspiration.



Understood in this way, the IPA symbol [h] commonly used for aspiration is not adequate since it denotes a fricative noise produced at the glottis

and, although it is true that friction is produced at this point, it is a mere secondary phonetic effect. Thus, a distinction ought to be made between aspiration which is specifically articulated at the glottis and aspiration in which the glottis is concerned but has a subsidiary role; such is the case with the distinction between a labial consonant (where the primary articulators are the lips) and a labialised consonant (in which the place of articulation can be independent of the lips but the lips influence the pronunciation as a secondary articulation). A more exact definition of the latter type of aspiration would be that of an open glottis with no stipulation as to a point of articulation. In these terms, two factors are necessary for a period of aspiration: (a) an abduction of the vocal folds, and (b) the lack of any type of constriction or obstruction in the supraglottal region. This definition of aspiration, however, comes very close to the physiological characteristics of the production of a type of phonation termed *breathy voice*.

4.2 *Breathy Voice and Aspiration.*

Traditionally the glottis was considered to adopt two settings during speech: the vocal folds together or the vocal folds apart. Thence come the two classes of sounds; the result of the former being the vibration of the vocal folds and voiced sounds, and the result of the latter, the lack of vibration and voiceless sounds. However, recent research into various languages has revealed that the vibration of the vocal folds is not limited to this binary option, but rather there can be different types of phonation corresponding to the different rates of proximity of the vocal folds; and that these differences in tension can constitute phonemic differences.

Peter Ladefoged, who has studied phonation types in various languages, defines, in Ladefoged and Maddieson (1996), five types of phonation; however, in more recent works he seems to reduce these five types to three: Breathy Voice → Modal Voice → Creaky Voice.

The first of these, *breathy voice*, represented in the IPA by the diacritic [.] is attested in many languages in the North of India to distinguish between occlusive consonants and, in some cases, vowels⁹

⁹ This is the case in Gujarati, as attested by the following minimal pairs (Ladefoged 2001:126):

(Ladefoged 2001:123). From an articulatory point of view breathy voice is produced with a greater aperture of the vocal folds than that which occurs in modal voice. That is to say, in breathy voice the vocal folds are vibrating, but not as strongly and less regularly due to the fact that the arytenoid cartilages are more separated than in modal voice. This means that in breathy voice, as opposed to modal voice, the vocal folds do not close completely and, also, the point at which they are most abducted in breathy voice is greater than that in modal voice. A direct consequence of this is that breathy voice is produced with less intensity and with a greater glottal airflow¹⁰; due to the fact that the greater the distance between the vocal folds not only prevents them from making significant contact when they come together, but also, at the same time, allows more air to flow through the glottis.

In this way breathy voice is similar in what has been termed aspiration in that both constitute the setting of a wide glottis which is superimposed on the form of the oral cavity. The difference between the two lies in the amount of separation between the vocal folds: if they are very far apart (wide glottis) a great quantity of air will pass through them and no vibration will occur; this, in lack of any constriction in the vocal tract, could be understood as aspiration. However if, on the other hand, the distance is not enough to prevent them from vibrating, a segment of breathy voice will be produced. In this way, aspiration and breathy voice are basically contiguous terms on a scale ranging from a very wide glottis to a closed glottis. Thus, the symbol used to represent a voiced aspiration [*h*], which denotes a vibration of the vocal folds coincides totally with the diacritic for breathy voice and so, on phonetic grounds, the transcription for the English interjection *aha* could be both [a *h*a] or [aa-a].

Modal Voice	Breathy Voice
[pɔɪ] 'Last year'	[pɔɪ̤] 'the dawn'
[bɑɪ] 'twelve'	[bɑɪ̤] 'outside'
[kɑɪ] 'ear'	[kɑɪ̤] 'Krishna'

¹⁰ According to Ladefoged y Maddieson (1996:50), for a male with a subglottal pressure of 8 cm. H₂O, during a voiceless sound the rate of airflow through the glottis is a maximum of 1000 ml./s, whilst for Breathy Voice it is around 500-250 ml/s and 120 ml/s for Modal Voice.

5. Conclusion

In light of this, the results from the tests carried out on WAS can be better understood. If we assume that there exists a phonological process which aspirates an underlying implosive sibilant, then this, in articulatory terms, translates into an abduction of the vocal chords; this abduction can either be wide giving aspiration, or less so giving breathy voiced phonation of different types. Now, in utterance final position this process may also be carried out, however, with less intensity, as is true of all sounds produced in this position. As to why in this position the distinction is lost between the underlying sequences VS and V; this can be explained through the physiological mechanisms of speech production. That is, upon speaking a phrase which ends in a vowel, the end of the phrase is often produced less intensely since, at this point in speech, the subglottal pressure is low as the speaker is running out of air. The articulatory and perceptive consequences of this phenomenon are that the reduced pulmonary pressure is not sufficient enough to be able to produce a significant vibration of the vocal folds which would constitute a modal voice phonation. What it frequently does produce, however, is a weak vibration of the vocal folds which can only be described as breathy voice phonation or even, if the subglottal pressure is very low, aspiration. In the case of high subglottal pressure at the end of the utterance, in these cases, the final vowel, as opposed to when it occurs inside an utterance does not suddenly end in order for another gesture to begin. In this particular position, the glottis relaxes in order to assume its rest position. This is not achieved through a brusque abduction of the vocal folds but rather a gradual one in which air continues escaping through a slowly relaxing glottis producing slight vibrations which in turn must be phonetically characterised as incidents of breathy voice phonation.

Therefore, in utterance final position, interesting circumstances arise in which the result of a phonological process for an underlying sibilant element coincides with something purely phonetic, related to the physiology of speech. That is, in utterance final position, final vowels tend to be pronounced, at the periphery, with a breathy voice phonation which is the same phonetic realisation as an underlying implosive sibilant in WAS. Once again, phonetics is impinging upon and mixing with phonology.

REFERENCES

- GERFEN, C. 2001. A critical view of licensing by cue: codas and obstruents in Eastern Andalusian Spanish. In L. Lombardi, ed. *Segmental Phonology in Optimality Theory*. New York: Cambridge UP.
- JIMÉNEZ FERNÁNDEZ, R. 1999. *El andaluz*. Madrid: Arco-Libros
- KEATING, P. 1988. Underspecification in phonetics. *Phonology* 5: 275-292.
- LADEFOGED, P. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- LADEFOGED, P. & MADDIESON, I. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- LADEFOGED, P. 2001. *Vowels and Consonants: An Introduction to the Sounds of Language*. Malden, MA: Blackwell.
- MARTÍNEZ MELGAR, A. 1986. Estudio experimental sobre un muestreo de vocalismo andaluz. *Estudios de Fonética Experimental* 2: 195-248.