

THE EFFECT OF VOICING ASSIMILATION ON GESTURAL COORDINATION

Joaquín Romero

Universitat Rovira i Virgili, Tarragona, Spain, and Haskins Laboratories, New Haven, USA

ABSTRACT

The present study investigates the phonetic realization of voicing assimilation in consonant sequences in Spanish. The effects of voicing assimilation are investigated in terms of both the magnitude of the laryngeal gesture and the interarticulator timing between laryngeal and supralaryngeal gestures. We looked at the effects at the gestural level of a phonological process by which two segments with opposite specifications for voicing are placed side by side. The data have been collected using an EMMA system with simultaneous observation of the laryngeal behavior by means of laryngeal transillumination. The results show that the patterns of interarticulatory timing are affected by the assimilatory process. This supports an interpretation of assimilation as gestural blending, which is a more realistic explanation for voicing assimilation than is provided by traditional phonological accounts. It also provides an interpretation of gestural blending that is consistent with other similar types of assimilatory processes.

1. INTRODUCTION

1.1 Objectives

Studies have shown the effects of blending between the laryngeal gestures of two adjacent consonants [3]. Also, a tight link between oral and laryngeal gestures has been observed for a variety of consonants [2], which suggests a certain level of coordination between laryngeal and supralaryngeal gestures.

In this study, we look at the effects at the gestural level of a phonological process by which two segments with opposite laryngeal specifications are placed side by side. The data come from the standard variety of peninsular (Castilian) Spanish spoken in Spain.

1.2 The phenomenon

Voicing assimilation in Castilian Spanish has been described as a straightforward anticipatory process by which a voiceless consonant becomes voiced when it is followed by a voiced consonant.

Spanish has a very restricted number of possible consonants in syllable coda position. Thus, voicing assimilation is most common with syllable-final /s/. Table 1 illustrates this point

<i>esperar</i> 'wait': /esperar/	>	[esperar]
<i>resbalar</i> 'slip': /resbalar/	>	[resβalar]

Table 1. Examples of voicing assimilation.

1.3 Predictions

Following this description, one should expect a simple spreading of the laryngeal specifications of the second consonant to the first. Thus, in the *resbalar* example we

anticipate little or no change in the laryngeal settings from V to CC to V, since the result of the assimilation process should be a sequence with uninterrupted laryngeal pulsing.

Alternatively, the result of the assimilation could be the blending of the laryngeal gestures of the two consonants, resulting in a laryngeal gesture that is a compromise between those of unassimilated /s/ and of the following voiced consonant. Such a result would be predicted by Articulatory Phonology [1].

2. EXPERIMENTAL PROCEDURES

2.1 Design

In order to test the hypothesis, an experiment was designed that investigated the laryngeal and supralaryngeal characteristics of Spanish consonant sequences. The sequences consisted of /s/ plus either a voiced or a voiceless stop. Labial, dental and velar stops were included.

The consonant(s) appeared within words and across word boundaries. In addition, each consonant appeared as part of a cluster and in isolation. Thus, there were four variables in the design:

- voicing: voiced vs. voiceless
- context: cluster vs. single
- environment: within word vs. across word boundary
- point of articulation: labial, dental and velar

The contexts and words that were included in the design appear in Table 2 below. All the words appeared in the carrier sentence:

Diga _____ muchas veces

Each utterance was repeated six times.

2.2 Data collection

The experiment used two techniques for articulatory data acquisition simultaneously: electromagnetic articulometry —EMMA— and laryngeal transillumination. Data were collected from one native speaker of Castilian Spanish.

The EMMA tracking technique consists of three transmitters that generate a magnetic field, and a number of receiver transducer coils that are placed on the subject's vocal tract. In the present experiment, coils were placed on the subject's upper and lower lips, jaw, on four positions on the tongue (tongue tip, tongue blade, tongue dorsum and tongue rear) and as references for subsequent error correction.

The transillumination technique consists of two parts: a flexible endoscope with a light source at its tip that is introduced through the subject's nose and hangs in the pharynx, and a photoelectric cell which is placed on the subject's neck right below the thyroid cartilage. The cell captures the amount of light that passes through the glottis. During the recording session, a simultaneous video recording of the larynx was obtained.

The experiment was performed in one session at Haskins Laboratories.

	voiceless		voiced		single
	within	across	within	across	
labial	despacio	dos pasos	resbala	dos balas	paso/bala
dental	destierra	dos tierras	desdeña	dos dueñas	tierra/dueña
velar	descara	dos caras	desgana	dos ganas	cara/gana

Table 2. Experimental design.

2.3 Data processing

The data were digitized simultaneously with the audio signal. All EMMA signals, as well as the glottal signal from the transillumination, were filtered, calibrated, and corrected for any possible experimental error. For each token, a number of synchronized signals were low-pass filtered and extracted to include only the target CVC(CVC) sequence.

The following measurements were obtained with the help of the HADES package from Haskins Laboratories:

Amplitude (or magnitude) of the laryngeal signal at the peak associated with the target consonant(s)

Timing interval between laryngeal and supralaryngeal peaks

The laryngeal amplitude peaks were identified by using the zero crossing of its first derivative as landmark points. Similarly, minima in the tangential velocity signals were used to identify peaks in the movement signals.

2.4 Statistical Analysis

The data were analyzed statistically using separate factorial analyses of variance with each of the measurements above—amplitude (or magnitude) and timing interval—as the dependent variable. The independent variables were VOICING (voiced vs. voiceless), CONTEXT (cluster vs. single) ENVIRONMENT, (within word vs. across word boundary) and POINT OF ARTICULATION (labial vs. dental vs. velar). Thus, three separate factorial analyses were performed for each dependent variable: VOICING*CONTEXT, VOICING*ENVIRONMENT and VOICING*CONTEXT*POINT OF ARTICULATION*.

The analyses were performed with the Statview statistics package for the Macintosh.

3. RESULTS

3.1 Magnitude

Table 3 indicates which of the analyses of variance produced significant main effects and interactions for the magnitude

(amplitude) of the laryngeal gesture.

The VOICING*CONTEXT ANOVA shows significant results at all three points of articulation. These relationships are also illustrated in the bar graphs in Figure 1. In these graphs larger bars indicate a smaller glottal opening, that is, more voicing.

The highest levels of voicing correspond to the *single voiced* condition, while the *cluster voiceless* condition shows the lowest. The *cluster voiced* condition is always lower than the *single voiceless* condition.

The variable ENVIRONMENT does not appear to be significant with respect to voicing at any point of articulation.

Results for POINT OF ARTICULATION do show significant results: the level of voicing is consistently higher in labials than in dentals or velars.

3.2 Timing

Table 4 gives the results of the analyses of variance for the timing measure (interval between laryngeal and supralaryngeal peaks). These analyses only include two levels of the POINT OF ARTICULATION factor—namely, labial and velar. The reason for this is that the data for the dental consonants were confounded by the fact that, in the *cluster* condition, the peaks in the tongue-tip movement signals reflected the movement of both /s/ and the target consonant. This caused great difficulty in isolating the point in time associated with the maximum movement of the tongue tip for the target consonant exclusively.

In the VOICING*CONTEXT ANOVA the main effect for CONTEXT is highly significant for both labial and velar points of articulation. However, in the velar point of articulation, there is a significant interaction. The VOICING effect is also highly significant for both labials and velars. For the VOICING*ENVIRONMENT ANOVA, there seems to be no clear pattern of significance, with non-significant results for both individual effects but a significant interaction in the case of

	VOICING*CONTEXT		VOICING*ENVIRONMENT		VOICING*CONTEXT*POA	
labial	context	.01	environment	not sig	context	.01
	voicing	.01	voicing	.01	voicing	.01
	interaction	not sig	interaction	not sig	cont*voic	not sig
dental	context	.01	environment	not sig	POA	.01
	voicing	.02	voicing	.02	voic*POA	not sig
	interaction	not sig	interaction	not sig	cont*POA	not sig
velar	context	.01	environment	not sig	cont*voic*POA	not sig
	voicing	not sig	voicing	not sig		
	interaction	not sig	interaction	not sig		

Table 3. ANOVA results for magnitude

	CONTEXT*VOICING		ENVIRONMENT*VOICING		CONTEXT* VOICING*POA	
labial	context	.01	environment	not sig	context	.01
	voicing	.01	voicing	not sig	voicing	.01
	interaction	not sig	interaction	.01	cont*voic	.01
velar	context	.01	environment	not sig	POA	not sig
	voicing	.01	voicing	.01	voic*POA	not sig
	interaction	.01	interaction	not sig	cont*POA	not sig
					cont*voic*POA	not sig

Table 4. ANOVA results for timing

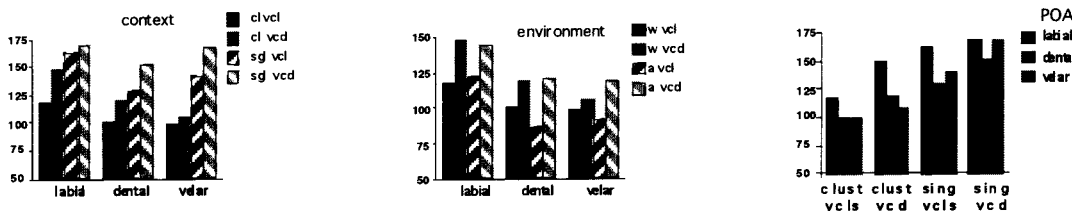


Figure 1. Bar graphs for amplitude (magnitude) measurements

the labials, and only a significant effect for VOICING in the velars; in either case the ENVIRONMENT effect was non-significant.

The relationships between the different treatments are illustrated as bar graphs in Figure 2. In these graphs 0 corresponds to the laryngeal peak. Thus, positive values indicate a lag between the laryngeal and supralaryngeal peaks, that is, the laryngeal peak occurs first and then the supralaryngeal; negative values indicate the opposite: the supralaryngeal peak occurs first and is followed by the laryngeal peak. As mentioned above, these graphs do not include the dental point of articulation due to the difficulties in identifying the separate gestures corresponding to the two consecutive tongue-tip gestures.

The graph for CONTEXT shows *single voiceless* as the only condition with negative values. Thus, for /p/, /t/ and /k/ the supralaryngeal constriction takes place prior to the laryngeal opening peak.

4. DISCUSSION

4.1 Magnitude

Results for CONTEXT show a gradation of voicing that goes from most voiced for the intervocalic *voiced stops*, to least voiced for the */s/+voiceless stop* sequence. Interestingly, the *single voiceless* category shows levels of voicing that are higher than those of the *cluster voiced* one, which can imply two things: a) that Spanish voiceless stops are not completely voiceless in colloquial speech, or b) that there actually is no assimilation of voicing to the expected degree.

Thus, the assimilated */s/+voiced stop* sequence shows laryngeal gestures whose magnitude is somewhere in between those of the *single voiced* category and the unassimilated */s/+voiceless stop* sequence. This can be taken as an indication that Spanish voicing assimilation is not a categorical phenomenon.

As the bar graphs show, the 'environment' —whether the consonant(s) appear within words or across word boundaries— does not seem to have an effect on the size of the laryngeal opening.

In terms of point of articulation, as mentioned above, labials seem to have consistently higher degrees of voicing than either dentals or velars. Velars have smallest degrees of voicing in the cluster, which could be an indication of their more 'stop-like' nature, as opposed to the single consonants, which are typically spirantized in Spanish.

4.2 Timing

The results for the 'timing interval' variable show some interesting patterns. With respect to CONTEXT, both labials and velars present positive lags between the laryngeal and supralaryngeal peaks in the clusters, while the single consonants are either negative or slightly positive. Thus, in single consonants the supralaryngeal constriction occurs before the laryngeal peak —for the voiceless ones— or slightly afterwards —for the voiced ones—, while in clusters, the laryngeal peak always precedes the supralaryngeal constriction.

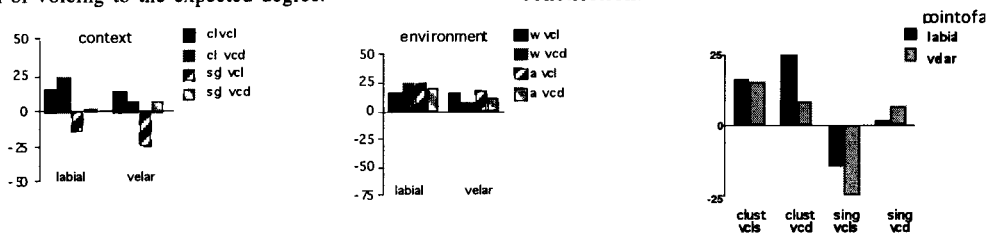


Figure 2. Bar graphs for timing interval measurements (in ms).

The fact that the laryngeal gesture occurs consistently before the supralaryngeal gesture of C2 in the cluster, must be clearly caused by the presence of C1, i.e. /s/. The issue then becomes whether the observed laryngeal gesture is in fact synchronized with /s/ or if it is timed with both C1 and C2 and thus, occurs somewhere between the supralaryngeal peaks of C1 and C2.

It was pointed out earlier that *single voiceless* was the only category that showed a lag between the supralaryngeal constriction and the glottal opening peak (Fig. 2). Therefore, it is conceivable that, similarly, the above mentioned laryngeal gesture in the clusters is indeed synchronized with the preceding /s/.

However, as Figure 3 below illustrates, the relationship between the laryngeal peak and the supralaryngeal peak for /s/ differs quantitatively from that of the single voiceless category in Figure 2. Just as in Figure 3 above, here 0 also corresponds to the laryngeal peak. The lag between supralaryngeal and laryngeal peaks in /s/ is much larger than in single voiceless consonants. This is an indication that, in fact, the laryngeal peak in the clusters is not actually synchronized with the supralaryngeal peak of /s/ but, rather, it occurs somewhere between the supralaryngeal peaks of /s/ and C2.

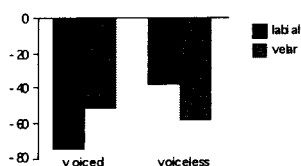


Figure 3. Timing interval (in ms) between laryngeal peak and supralaryngeal peak for /s/.

5. CONCLUSIONS

The results of this experiment show that the phonetic patterns observed do not correspond with the traditional phonological account. Such an account would predict a simple spreading of the laryngeal configurations from a voiced consonant to a preceding /s/ and, thus, we would expect to get a single, uninterrupted stretch of laryngeal pulsing in /sb/ for example.

This is clearly not what these data show. On the one hand, voicing assimilation appears as a gradual process that does not affect all sequences equally. The magnitude of the laryngeal gesture varies depending on context —single consonants present smaller glottal amplitudes than clusters, regardless, to a certain extent, of their voicing specifications. The reason for this is, presumably, the existence of a relationship between magnitude and duration. The longer the gesture, the greater its amplitude. This relationship has been found to be present in other assimilatory/lenition phenomena [4.]

On the other hand, we consistently see only one laryngeal gesture regardless of whether we have a single consonant or a consonant cluster. The timing relationships between this single laryngeal gesture and the two oral gestures present in the /s/+C clusters show mutual influence and, thus, blending in the resulting laryngeal configuration. Again, this situation seems to parallel cases of gestural blending in other assimilatory processes [5.]

Clearly, then, these observations warrant a more finely tuned explanation than can be provided by formal

phonological accounts. As mentioned above, the temporal dimension is crucial in order to account for these phenomena. Articulatory Phonology, which contemplates time as an inherent quality of the gesture, can incorporate the current findings more satisfactorily than other theories of phonology.

ACKNOWLEDGMENTS

I thank Vince Gracco, Louis Goldstein and Walter Naito for their assistance with the experiment. I would also like to thank Carol Fowler for allowing me to use the Haskins facilities to finish the processing of the data.

REFERENCES

- [1] Browman, C., & Goldstein, L.. 1992. Articulatory phonology: an overview. *Phonetica*, 49, 155-180.
- [2] Löfqvist, A. & Yoshioka, H. 1984. Intrasegmental timing: laryngeal-oral coordination in voiceless consonant production. *Speech Communication* 3, 279-289.
- [3] Munhall, K. & Löfqvist, A. 1992. Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics*, 20, 111-126.
- [4] Romero, J. 1995. Gestural organization in Spanish: an experimental study of spirantization and aspiration. Storrs, CT: unpublished doctoral dissertation.
- [5] Romero, J. 1996. Articulatory blending in lingual gestures. *Journal of Phonetics*, 24, 99-111.