Published in final edited form as:

J Phon. 2012 January 1; 40(1): 37–45. doi:10.1016/j.wocn.2011.08.004.

The role of gestural phasing in Western Andalusian Spanish aspiration

Benjamin Parrell

Department of Linguistics, University of Southern California

Abstract

This study examines the phenomenon of post-aspirated voiceless stops in Western Andalusian Spanish in /s/ + voiceless stop sequences. Previous analyses have proposed that the post-aspiration arises through a reorganization of the glottal spreading gesture for /s/ and the oral constriction gesture for the stop. This theory is tested by steadily increasing speakers' production rate, which has been shown to trigger spontaneous changes in gestural organization in speech and other motor activities. Results from the study support the initial hypothesis. There is a switch from productions with with preaspiration and short VOT to those with long VOT as rate increases. Additionally, there is a tradeoff between VOT and pre-closure aspiration, indicating they may result from the same gesture. Lastly, the variability in production shows a number of hallmarks of phase transitions in human coordination. In sum, a change in gestural organization provides a simple explanation for post-aspirated stops in this dialect.

Keywords

Andalusian Spanish; aspiration; gesture; phase shift; coordination

1. Introduction

1.1 S-aspiration

Many dialects of Spanish from all over the world, including southern Spain, the Canary Islands, much of the Caribbean, and coastal South American varieties include so-called s-aspiration. In these dialects, /s/ in coda position is produced as a voiceless glottal fricative [h] rather than an oral alveolar fricative. This gives productions such as [kahta] for /kasta/ ("caste"). This is generally accepted to be the consequence of a reduction in the tongue tip constriction gesture necessary to create oral frication accompanied by retention of the laryngeal abduction gesture necessary for voicelessness (see, for example, Romero, 1995), although whether this reduction is an online phonetic process or a phonological one has not been thoroughly investigated. While no dialects that have s-aspiration maintain /s/ in coda position, the exact production varies considerably, even within the same dialect.

Torreira (2007b) examines three dialects in particular: Western Andalusian Spanish, Puerto Rican Spanish, and Porteño Spanish (from the region around Buenos Aires). That study

© 2011 Elsevier Ltd. All rights reserved.

parrell@usc.edu, Telephone: 608-957-1371, Fax: 213-740-9306, Department of Linguistics, University of Southern California, Grace Ford Salvatori 301, Los Angeles, CA 90089-1693.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

finds significant inter-dialectal differences. The Puerto Rican and Porteño dialects, on the one hand, show significant aspiration—corresponding to coda /s/—before the closure for stops in a phonological /s/ + stop sequence. Additionally, both stop closure duration and voice-onset time (VOT) in these sequences were equal to that of plain intervocalic voiceless stops. Western Andalusian Spanish, on the other hand, shows significantly less preaspiration before stop closure than the other two dialects. Additionally, Western Andalusian Spanish voiceless stops preceded by /s/ show longer stop closure and VOT durations than their intervocalic counterparts. Though this is a robust dialectal difference, the productions of /s/ in Western Andalusian Spanish are reported to be somewhat variable, ranging from a full sibilant to pre-aspiration to a breathy period at the end of the preceding vowel to post-aspiration (Torreira 2007a). Thus, /kasta/ can be realized varyingly as [kasta], [kahta], [ka_tə], or [ka^hta], with the last being the most common. Torreira (2007a, 2007b) analyzes these findings, using Articulatory Phonology (discussed below), as the result of a change in gestural organization between the glottal spreading gesture for coda /s/ and the oral constriction gesture for the stop. He hypothesizes that postaspiration occurs when the relationship between the two gestures switches from anti-phase (sequential) to in-phase (simultaneous).

A similar study on the related dialect of Eastern Andalusian Spanish found differences in closure duration for voiceless stops based on their context—intervocalic versus preceded by /s/ (Gerfen, 2002). This study also finds differences in stop closure duration for /s/ + stop tokens with long (>35 ms) versus short (<35 ms) preaspiration. Tokens with long preaspiration had a shorter stop closure, similar in duration to that of intervocalic stops, while tokens with short or no preaspiration had a longer stop closure. This finding points to a possible compensatory effect between preaspiration and stop closure durations. This study did not address VOT differences; however, evidence from Torreira (2007a) shows that Eastern Andalusian Spanish speakers also show some degree of postaspiration in /s/ + voiceless stop sequences, though to a lesser degree than in Western Andalusian Spanish. Gerfen (2002) proposes an explanation for the durational relationship between pre-closure aspiration and closure duration that relies on gestural organization and overlap, though the analysis is not specific about the particular relationships involved.

A few sociolinguistic studies have also examined post-aspirated /s/ + voiceless stop sequences (Moya Corral, 2007; Ruch, 2008). Both studies found that the production of these sequences with a long voice onset time and/or affrication on release was more prevalent among the younger generation than the older. Moya Corral also found this production to be more frequent among women and the middle class; Ruch found no differences based on gender but did find a greater frequency of productions with long VOT or frication in the university-educated population than in those with only primary or secondary education levels. Based on this evidence, both authors conclude that the novel production of these sequences is spreading, and perhaps may be a sound change in progress which would lead to a phonological distinction between unaspirated and aspirated (or perhaps affricated) stops in intervocalic position.

The general situation, then, is this: In Western Andalusian Spanish, /s/ + voiceless stop sequences are generally realized with longer VOT than intervocalic voiceless stops. The productions, though, are variable and can include a sibilant, pre-aspiration or breathy voice preceding the stop closure. Closure duration may be longer for the /s/ + voiceless stop sequences as well, though evidence for such an effect is mixed. While this paper will focus only on these sequences, a brief overview of other cluster types may be informative. Word-internally, intervocalic /s/ is generally maintained as a coronal fricative, though it is generally produced as a glottal fricative when word-final and often is deleted entirely in absolute final position. When /s/ occurs in clusters with nasals, laterals, fricatives, or voiced

stops (which are themselves lenited to approximants), it tends to lengthen the closure duration of these sounds (see Jiménez, 1999 and Ruch, 2008 for more details on the overall distribution of /s/ in Western Andalusian Spanish and Romero, 1995 for a more thorough investigation of /s/ + voiceless stop sequences).

1.2 Articulatory Phonology

In Articulatory Phonology, speech is analyzed into discrete action units called gestures that create constrictions in different areas of the vocal tract (Browman & Goldstein, 1995). In order to produce segments, such as /s/, with more than one gesture (i.e. oral constriction and laryngeal spreading, in this case), these gestures cohere and are temporally overlapped. One model for inter-gestural coordination associates each gesture with a planning oscillator, or clock, responsible for triggering the production of that gesture (Goldstein, Byrd, & Saltzman, 2006; Nam, Goldstein, & Saltzman, 2009; Saltzman, Nam, Krivokapic, & Goldstein, 2008). By coupling these clocks with different target relative phases it is possible to account for the stable patterns of overlap and sequentiality that these gestures exhibit. In this model, clocks for gestures whose onsets are synchronous (e.g. those for a syllable onset and syllable nuclei) have an in-phase relationship (0°) while sequential gestures (e.g. those for a syllable nuclei and coda consonants) are coupled to one another in an anti-phase relationship (180°). It is important to remember that these relationships are at the level of control, not necessarily of production. That is, that because two gestures are coordinated in an anti-phase relationship, it does not mean necessarily that the second occurs when the first ends. Rather, it is a relationship about when the two gestures *begin* (the second after the first); given that the first gesture is long enough, the two gestures will indeed overlap in time, such as gestures for the vowel and coda consonant within the same syllable.

Torreira (2007a) analyzed the post-aspirated voiceless stops in Western Andalusian Spanish as arising from a shift in phasing between the glottal abduction gesture for /s/ and the oral constriction gesture for the following stop. When produced as pre-aspiration (i.e. [ht]) the oral gesture is sequential to the glottal gesture, which could indicate an anti-phase relationship between their clocks. In Western Andalusian Spanish these are shifted to a concurrent, in-phase relationship, resulting in a longer period of post-aspiration (i.e. [t^h]), as shown in figure 1. Figure 1a shows anti-phase coordination between the two gestures (onset of glottal spreading precedes onset of oral closure); figure 1b, in-phase (the two gestures begin simultaneously). Notice that the gestural durations do not change between the two figures but that differences in organization alone lead to preaspiration and a short VOT in 1a, compared with no preaspiration and a long VOT in 1b.

The question remains, however, about why this phase shift would occur if both in-phase and anti-phase relationships occur freely in Spanish. The answer lies in the relative stability of these two phasing relationships. There is a substantial body of work that demonstrates that while both in- and anti-phase modes are stable, anti-phase modes are somewhat less so. In finger tapping experiments, in-phase coupling remains stable at high rates of production, while anti-phase becomes unstable and spontaneously switches to an in-phase mode (Kelso, 1984). In that study, the initial anti-phase relationship was very stable at slow rates. As rate increases, the phasing relationship between the two fingers becomes more variable. At some point, the subjects spontaneously switch to a stable in-phase relationship, and the variability decreases. While most subjects showed a clean phase transition, with one to three cycles, there was one subject who showed a longer period of vacillation between phasing modes, though even in this case the in-phase mode was eventually settled on at fast rates. In another study, participants swung two pendulums, one with each hand, in either in- or anti-phase coordination (Court, Bennett, Williams, & Davids, 2002). The results match that of the finger tapping study; the participants exhibited the same pattern of spontaneous anti- to in-phase transition with increases in rate, with a period of high instability in relative phase

between the two pendulums immediately preceding the phase transition. Additionally, this study examined the relative phase between the two pendulums at both anti- and in-phase modes. The results indicate that there is more variability in relative phase in the anti-phase mode than in the in-phase mode, which provides more evidence for unequal stability between the two modes. Similar results have been found in a number of other studies (Kelso, Scholz, & Schöner, 1986; Swinnen, Dounskaia, Verschueren, Serrien, & Daelman, 1995; Wimmers, Beek, & Van Wieringen, 1992).

This large body of work in the motor control literature shows conclusively that different modes of phasing have different inherent levels of stability. Recall that Articulatory Phonology hypothesizes that, in speech, gestures that start sequentially (such as syllable onset and coda) to be anti-phase, while those that start simultaneously (such as syllable onset and nucleus) are hypothesized to be in-phase. If this hypothesis is correct, we should be able to find some evidence in speech phenomena for gestural phase switching of the type observed in the general motor control literature.

One such example comes from syllable affiliation of intervocalic consonants. In speech studies, numerous experiments have shown that VC syllables, when produced at a steadily increasing rate, will spontaneously switch to CV syllables (Tuller & Kelso, 1990; de Jong, Lim, & Nagao, 2004). In the study by Tuller & Kelso, subjects produced repetitive sequences of /ip/ and sequences of /pi/, steadily increasing their rate of speech in six steps. Analysis focused on the relative timing of lip and laryngeal movement. The former was collected via optical trackers on the upper and lower lips and the latter measured through glottal transillumination. The authors found that at slow rates of speech the glottal spreading gesture occurs later relative to the lip closing gesture in CV syllables than in VC syllables. However, as rate increased, the phase relation of VC syllables shifted to be indistinguishable from that of CV syllables, while the phasing in CV syllables remained constant. The authors take this result to be evidence for a bistable coordinative pattern, where the VC coordination pattern becomes unstable at high rates and switches to the more stable CV pattern. This data presents another case of a gestural realignment from anti- to in-phase coordination at high rates, similar to the bimanual motor data presented above. Additionally, in a follow-up perception study conducted on the same data, subjects heard only CV syllables at fast rates regardless on which syllable (/pi/ or /ip/) was intended in production.

In the study by de Jong et al., glottal opening was again measured by transillumination but its timing was measured relative to acoustic closure release rather than articulatory lip movement. This study in the main replicated the findings of Tuller & Kelso. This study additionally found that the same rate-induced shift in coordination pattern that was found for VC syllables can be seen in *voiced* stops occurring in CV syllables in American English. These consonants are generally produced as voiceless unaspirated stops, which have a measurable glottal opening. At slow rates, the glottal spreading gesture for syllable-initial voiced stops resembled that for syllable final voiceless stops (relatively earlier glottal opening); as rate increased, subjects spontaneously switched to the same pattern found for syllable-initial voiceless stops (relatively later glottal opening). The authors take this as evidence, along with some changes in magnitude of the glottal spreading gesture, that the more-stable modes of gestural coordination are due not only to shifts in syllabic affiliation but affect many types of coordinated gestures.

An articulatory study on Portuguese also examined gestural re-phasing as a product of rate (Oliveira & Marin, 2005). In Brazilian Portuguese, a nasal + homorganic oral stop sequence is often perceived as a nasal only (e.g. *partindo* is heard as *partino*). The authors hypothesize that this perception can be attributed to a change in production from a sequential, anti-phase relationship between the velum closure gesture for the coda nasal and the oral closure

gesture for the following oral stop to a simultaneous, in-phase one. The study examined the coordination between the oral closing gesture and velum lowering gesture for /n/ using Electromagnetic Articulography (EMA). EMA sensors were placed on the tongue tip and on the velum. The phasing relationship between the two movements was measured by target-to-target lag, the difference in time between achievement of the velum target and the achievement of the target for the tongue tip. Results show that, when producing words with /nd/ and /nt/ sequences at a slow rate, the lag is negative (velum movement precedes tongue tip movement). As rate increases, the lag between the two gestures approaches zero, close to that seen for words with intervocalic /n/ only. These results indicate that speakers do indeed switch from sequential to simultaneous coordination as rate increases.

1.3 Predictions

Given the relative stabilities of the two coupling modes, we can explain the post-aspirated stops in Western Andalusian Spanish as the result of a simplifying phase change between the glottal abduction and oral constriction gestures involved. If that is indeed the case, and taking into account the instability of anti-phase coupling at fast rates that has been demonstrated in both manual and speech motor control, we hypothesize that during production speakers of Western Andalusian Spanish will shift from stops with preceding aspiration (whether [s], [h], or [•]) at slow rates to post-aspirated stops as rate increases. This main hypothesis makes a number of predictions. First, that subjects will show a shift from relatively unstable anti-phase organization (preaspiration and short VOT) to a relatively more stable in-phase organization (no preaspiration and long VOT). There will be a period of increased instability and variation in production between the two stable phases, though the duration of this period may vary between subjects. Second, subjects will never switch in the reverse direction, from in- to anti-phase. Third, since they are hypothesized to belong to the same glottal spreading gesture, decreases in preaspiration duration should be accompanied by increases in VOT. Last, since the phase shift is independent of the duration of the gestures involved, there should be no changes in the oral constriction gesture itself (such as the lengthening effects found in Torreira (2007b) and Gerfen (2002)), just in the phasing between that gesture and the glottal abduction. The following experiment was designed to test these predictions.

2. Material and Methods

2.1 Subjects

20 college-aged subjects recruited from Sevilla and Cádiz provinces (two of the three provinces within western Andalusia) participated in this study. All subjects were paid €7 (approximately \$10) for their participation.

2.2 Stimuli

The target stimulus was a word containing a word-internal /s/ + voiceless dental stop sequence in pretonic position (*pastándola*, /pas_tandola/¹, “grazing-it”). This was combined with 26 other targets and fillers collected as part of a separate study.

2.3 Procedure²

Subjects were shown one stimulus at a time, presented in conventional spelling, on a computer monitor. One second after the stimulus appeared, a series of clicks began to play. The click increased in rate in seven steps with four repetitions at each step, with a 100 ms

¹/s/ is used throughout this paper in the transcription. While not a common production in casual speech, many subjects did produce a coronal fricative in this position for at least some repetitions during this study.

²Adapted from Oliveira & Marin (2005).

reduction in the time between clicks at each step. The inter-click interval at the first step was 1100 ms and at the final, 500 ms. 1500 ms after the last click a new stimulus was presented on the screen and a new trial began. The subjects were instructed to say the stimulus aloud synchronously with each click. After a brief training session at the beginning of the experiment session, all subjects reliably synchronized their production to the clicks.

The experiment consisted of four blocks of stimuli. Each block contained all 27 stimuli (1 target and fillers) presented in a random order, with a small break between blocks. The total duration of the experiment was roughly an hour.

The subjects were digitally recorded onto a Sony Mini Disc MZ-R3 recorder using a Shure SM58 microphone. These digital files were then transferred to a PC using Sound Forge 8.0 at 44,100 Hz for analysis with Praat.

2.4 Data Analysis

Using the acoustic waveform and spectrogram, the following measurements were made:

1. Pre-closure aspiration (/s/): measured from the point where a clear decrease in formant amplitudes was seen in the spectrogram until neither the formants nor any turbulence were visible (stop closure). This method was chosen in order to consistently measure tokens with breathy voice at the end of the vowel rather than complete voiceless aspiration. Differences in realization of the pre-closure aspiration (including [s], [•], or [h]) were coded in the data but will not be discussed in this paper.
2. Closure duration: measured from the end of aspiration noise or, if none existed, the decrease in formant intensity at the end of the vowel associated with closure to the release burst.
3. Post-closure aspiration (VOT): measured from the release burst, or first release burst if there were more than one, to the last downward zero crossing before a full glottal cycle in the waveform.
4. Total token duration

In order to assure the accuracy of these measurements, a second transcriber annotated one subject. Inter-transcriber reliability was 94% within 5 ms and 99% within 10 ms. On only four tokens did the transcribers disagree on the presence or absence of preaspiration. In one case transcriber 2 identified 8 ms of [h] where transcriber 1 did not identify any preaspiration. In the other 3 cases transcriber 2 identified a vowel occurring before the /s/ + stop sequence as entirely breathy while transcriber 1 did not identify any breathy voice.

3. Results

3.1 VOT patterns

Short and long VOT categories were designated. This is justified by the binomial distribution of VOT as shown in figure 2a. All tokens of VOT were fit with gaussian distributions using the Netlab toolbox in Matlab (Nabney, 2002). In the figure it can be seen that fitting the data with two gaussian distributions (centered on 18 and 49 ms) reduces the error from the fit with only one (centered on 41ms). Figure 2b shows that the gains in error from fitting additional gaussian distributions are negligible. In fact, the centers for three fitter curves are 18, 49 and 49ms, indicating the third curve more-or-less duplicates the second. In sum, then, VOT in the data falls into two categories, one distributed around 18 ms and the other around 49 ms. The central tendencies of these distributions are in line with

those reported for languages with a distinction between stops with short- and long-lag voice onset time such as English and Thai (e.g. Kessinger & Blumstein, 1997).

The boundary between short and long VOT was set at 40 ms at step 1 (corresponding to the slowest rate of production), decreasing in equal intervals to 30 ms at step 7.³ In the current study, speakers could produce one of four logically possible patterns: consistent short VOT across all rates (SS), consistent long VOT across all rates (LL), a short VOT at slower rates and a long VOT at faster rates (SL), or a long VOT at slower rates and a short VOT at faster rates (LS). Supposing that, because they arise from the same gesture, there is a tradeoff between preaspiration and VOT⁴ (as shown in figure 1), the first pattern is consistent with a constant anti-phase organization between the glottal abduction gesture for /s/ and the oral closure gesture for /t/ (1a); the second, with constant anti-phase organization (1b); the third, with a switch from anti- to in-phase organization (switch from 1a to 1b); the last, with a switch from in- to anti-phase (switch from 1b to 1a). The last, notably, is a pattern that is predicted to not occur based on the findings from Kelso (1984) that in-phase relationships are more stable than anti-phase.

The data from all trials (4 trials per subject, 20 subjects) were examined algorithmically to determine which VOT pattern they exhibited. The algorithm calculated the mean VOT produced at each step and compared that against the VOT boundary value for that particular step. Trials where all means fell below the boundary were categorized as SS; those where all VOT fell above the boundary, as LL; those where VOT first fell below the boundary then switched to above the boundary *without switching back*, as SL; those where VOT first fell above the boundary then switched to below without switching back, as LS. The algorithm failed to successfully categorize only one trial (subject 20, trial 4). For this trial, the mean VOT for steps 1, 3, and 5 were slightly above the short-long VOT boundary (less than 2 ms), while that for steps 2 and 4 were slightly below. Since the averages were close to the boundary for steps 1–5, and since the majority of the individual repetitions during these steps also fell below the boundary, those steps can be seen as best fitting the short VOT pattern. As the VOT at steps 6 and 7 were clearly above the boundary (more than 10 ms), this trial was included as part of the SL pattern. The results are presented in table 1. The vast majority of the trials fall into one of the three predicted categories; the null hypothesis that initial and final VOT are independent can be rejected at the $p = 0.001$ level using a chi-square test.

Notice, however, that there are two trials that show the unpredicted pattern (long initial VOT and short final VOT). These two trials are shown in figure 3.

Notice that the switch from long to short VOT for both these trials occurs only at the final step, and the mean VOT at this step is only slightly below the boundary. Additionally, the decrease does not come as a sudden jump but reflects a pattern of a steadily decreasing VOT at increased rates. Lastly, these trials do not show any increase in preaspiration duration where the VOT drops below the boundary, as would be expected for a change from in- to anti-phase coordination, and which we do see for those trials where there is a consistent switch in VOT. For all these reasons, it seems more probable that these trials do not reflect a categorical change from long to short VOT (and subsequent phase shift) but rather simply tokens at the lowest end of the long VOT category. For the remainder of this paper, these trials will be included in the LL pattern.

³These category designations are based on evidence from Kessinger & Blumstein (1997), which found that the boundary between these two categories (which they called “short lag” and “long lag” stops) decreases with rate for both Thai and English. Additionally, data from that study showed that 40 ms at a slow rate and 30 ms at a fast rate in both languages unambiguously differentiates between the two categories.

⁴Results from this study support this claim. See section 3.2.3.

Data from subject S2 is presented in figure 4 as an illustration of the three attested patterns. We can see that trial 1 shows a shift from short to long VOT as step increases, trial 2 shows consistently long VOT, and trials 3 and 4 show consistently short VOT.

In addition to step, as shown in Figure 4, total token duration was also used as a measure of rate. This was done to ensure accurate quantification of rate, as rate might not necessarily increase at each step (subjects could, for example, speak quickly at low steps and simply have a longer duration between repetition, or might speak slowly at high steps and skip every other click).

In order to test whether the step manipulation had a significant effect on speech rate, a correlation was calculated between step and the total duration of each token. The results of this test show a significant relationship between the two variables, with shorter productions at higher steps ($r = -0.66$, $p < 0.0001$). As total token duration is a more direct measure of speech rate, that measure will be used in the remainder of the analysis. As an example for comparison, the relationship between total token duration and VOT is shown in figure 5, again for subject S2 as an example.

We can see that each trial in figure 5 shows the same pattern as in figure 4. Trials 3 and 4 have consistently short VOT (SS pattern), with only a handful of tokens for trial 3 above 40 ms. Trial 2 has a consistently long VOT (LL pattern), with only four tokens below 40 ms, nearly all of which come at very fast rates (low total token duration). Trial 1 shows short VOT when total token duration is long (i.e. at a slow speaking rate) and long VOT when total token duration is short (SL pattern). This is the same pattern we saw with step as the independent variable.

3.2 Effects on VOT

If, as proposed, the aspiration after and that before the stop come from the same glottal spreading gesture, we expect a number of relationships with VOT to occur in the data. First, the amount of preaspiration and postaspiration should vary inversely. Because gestural reorganization is essentially moving the alignment of the glottal and supralaryngeal gestures, a decrease in the amount of the glottal gesture that occurs before the stop closure will result in an increase in the amount of that same gesture that lasts after the closure has been released. Second, the duration of the closure should vary inversely with VOT. Given an invariant duration for the glottal spreading gesture, any possible decrease in the stop duration will result in less overlap between the two gestures, and longer aspiration (Maddieson, 1997). Similarly, an increase in stop duration will result in *more* overlap and a concomitant *decrease* in VOT. Lastly, we should expect that rate will have a negative impact on VOT, especially for LL trials. As has been shown in previous work, increased rate of speech causes a compression of VOT for stops with long VOT, though these effects aren't seen for stops with short VOT (Kessinger & Blumstein, 1997; Pind, 1995). These predictions are illustrated in figure 6.

In order to test these predictions, the data for each group (SS, SL, LL) were fitted with a linear mixed model using the lme4 package in R (Bates, 2005). Because each group may show different patterns (as discussed in the previous paragraph for the effects of rate on VOT), they were analyzed separately rather than grouping all the data together. For all models, VOT was used as the response, subject and trial as random intercepts, and total token duration, consonant closure duration, and preaspiration duration as fixed factors. Significance reported in the following results for these factors was calculated with the same package using Markov chain Monte Carlo simulations with 10,000 samples. An advantage of using a mixed model is that, besides correctly accounting for by-subject variation, they allow us to examine the effect of each independent variable more-or-less independently. For

the current study, this means that we can examine the effects of closure duration and preaspiration duration on VOT after accounting for any co-variation due to overall compression of all gestures in faster speech.

In the SS group, a model with preaspiration duration as the only fixed factor (i.e. without closure duration or total token duration) is not significantly different from the full model ($\chi^2(2) = 4.48, p = 0.108$). In this model, preaspiration duration has a significant inverse effect on VOT ($\beta = -0.065, p < 0.001$). For the SL group, the effects of all the fixed factors are significant. VOT covaries inversely with preaspiration duration ($\beta = -0.321, p < 0.0001$) and closure duration ($\beta = -0.155, p < 0.0001$), while it covaries positively with token duration ($\beta = 0.022, p < 0.01$). The effects for the LL model are also all significant in the same direction (preaspiration: $\beta = -0.218, p < 0.0001$; closure: $\beta = -0.122, p < 0.0001$; token: $\beta = 0.055, p < 0.0001$).

The results from these models agree with the predictions derived from the hypothesis that a single glottal gesture underlies both pre- and post-aspiration in Western Andalusian Spanish. The two vary inversely in all conditions. Additionally, in the SL and LL conditions, VOT is shorter at faster rates of speech, and also when the closure duration for the stop is shorter. These two factors are not significant in the SS condition, however. This may be due to the general lower variability of VOT in this condition (standard deviation of 8 ms compared to 18 and 13 in the SL and LL conditions, respectively). The results here also mirror those found in Kessinger & Blumstein (1997), where stops with long VOT in English and Thai show a reduction in VOT duration at fast rate while those with short VOT in the same languages do not.

Additional evidence for this trade between pre-aspiration and post-aspiration comes when we examine the number of tokens with both long VOT and preaspiration (which, because they are hypothesized to come from the same gestural source, should not co-occur). Of the 1239 tokens with long VOT, 936 (75%) have no preaspiration and 1067 (86%) have less than 15 ms of preaspiration. These are roughly the same proportions as found in the LL group, where 78% of tokens show no preaspiration and 89% less than 15ms.

3.3 Closure duration

The hypothesis that relies on gestural re-phasing to explain aspirated stops in Western Andalusian Spanish does not predict any changes in the closure duration for the stop. However, there is some evidence for so-called compensatory lengthening of the stop accompanying aspiration/deletion of coda /s/. One previous study found stop closure duration to be longer in tokens with shorter preaspiration duration for the closely related dialect of Eastern Andalusian Spanish (Gerfen, 2002). Another study on Western Andalusian Spanish found stop closure duration to be greater in words where there is no preceding coda than in those where there is a phonological /s/ (Torreira, 2007a). This latter study also examined the relationship between closure duration and VOT for words with /s/ + voiceless stop sequence, and found no evidence for any effect.

If compensatory lengthening effects were visible in Western Andalusian Spanish, we should expect there to be differences in stop closure duration where there are differences in preaspiration duration (increased stop duration should be taking the place of—*compensating for*—decreases in preaspiration duration). In fact, we have just such a condition readily available, as there is a significant difference in preaspiration duration between the SS and the LL groups ($M_{SS} = 46.89, M_{LL} = 4.42, t(1289) = 44.76, p < .0001$). Despite this difference, however, there is no difference in stop closure duration ($M_{SS} = 78.57, M_{LL} = 80.94, t(1289) = -1.37, ns$).

As mentioned in section 3.2, the results from this study do show a significant inverse relationship between closure duration and VOT, contradictory to the results from Torreira (2007a). We know, however, that previous work shows a trade-off between precisely those two variables (e.g. Maddieson, 1997). Thus, Western Andalusian Spanish does not seem to be in any case unusual in this regard. It seems safe to follow the conclusions reached by Torreira, attributing this relationship in Western Andalusian Spanish to articulatory universals (that is, a shortening of the closure gesture causes less overlap with the glottal spreading gesture). In summary, no evidence is found in the current data for any sort of compensatory lengthening.

3.4 Variability in phasing

3.4.1 Background and predictions—In order to examine consistency of production, variability in production was calculated at each step. Many trials that show a switch from short to long VOT do not switch at one point but rather show a period of consistent short VOT, followed by fluctuation between short and long VOT, and lastly a period of consistent long VOT. Though this period of instability in the transition from anti- to in-phase is not unexpected, and may in fact be likely given the competing forces on subjects' speech (discussed below in section 4), it is nonetheless useful to attempt to quantify the effects of this instability.

If the gestural re-phasing hypothesis is true, we should see periods of relative stability in gestural relationships when they are anti- or in-phase. That is, variability in VOT and preaspiration should be low both *before* subjects switch from short to long VOT and *after* the switch has been completed. During a phase transition, however, variability should go up as subjects fluctuate between the two production possibilities. Additionally, variability should be lower in the in-phase than the anti-phase relationship (Court et al., 2002): subjects should show lower variability after they switch from low to high VOT productions.

3.4.2 Methods for calculation—First, an algorithm was devised to automatically identify those trials with a phase transition by examining the means at each step for both preaspiration duration and VOT. The onset of the phase transition was identified at the step where the means were equal, and the end of the phase transition was identified as the step where the mean for VOT was twice the mean for preaspiration. These locations were compared with a calculation for the phase transition that located the step where VOT crossed the boundary dividing short and long VOT (which ranged from 40-30 ms, as discussed in 3.1). In all but two cases, this point fell within the algorithmically defined phase transition window. In these two cases, the end point of the phase window was changed to the point identified by the VOT switch, which always came at a later step than the point defined algorithmically.

Variability was measured in the following manner: the ratio of VOT over total aspiration (VOT and any preaspiration) was calculated for each token (aspiration ratio). The mean ratio was then calculated at each step (four tokens per step). For each token, the absolute value of the deviation of that token's ratio from step mean was taken as a measure of variability. This approach mirrors the use of standard deviation of relative phase as a measure of variability in previous work (Court et al., 2002).

Once these deviations were calculated, these values were sorted into three groups based on the pre-identified phase transition boundaries: pre-phase transition (PRE), during phase transition (SWITCH), and post-phase transition (POST). This was done for all subjects together, but only included trials where a phase transition was identified.

3.4.3 Results—Because the distribution of the deviation data was not normal (see figure 7), the nonparametric Kruskal–Wallis one-way analysis of variance was used. This test, with group as the independent variable, was conducted on the deviation data, shown in figure 7. The test revealed significant differences between the three phase groups (Mean rank_{PRE} = 463.4, Mean rank_{SWITCH} = 527.9, Mean rank_{POST} = 289.6, $\chi^2(2,831) = 155.5$, $p < 0.0001$). A Tukey–Kramer post-hoc test showed there to be significant differences between all three groups (PRE vs. SWITCH, $p < 0.01$; other comparisons, $p < 0.0001$). We can see that the variability in production is highest around the area near the switch from short to long VOT, and lower both before and after this region. Additionally, variability is higher before the switch than after it. These results are consistent with a transition from anti-phase to in-phase coordination. Variability is highest around the transition between the two stable modes, and higher for anti-phase coordination than for in-phase.

4. Discussion

The data presented above support the hypothesis that long VOT stops in Western Andalusian Spanish occur through a process of gestural phase realignment. Several predictions were made based on this phase shift hypothesis, which are briefly reviewed here:

1. Subjects will show a shift from anti-phase organization (preaspiration and short VOT) to in-phase organization (no preaspiration and long VOT).
2. They will not switch in the reverse direction, from in- to anti-phase.
3. Decreases in preaspiration will be accompanied by increases in VOT.
4. There should be no changes in the oral constriction gesture itself, just in the phasing between that gesture and the glottal abduction gesture.

As predicted, subjects show a transition from short to long VOT as rate increases. This occurred on more than a third of all trials (28/79), and more than half of those that began with short VOT (28/46). The overall pattern of significant correlations between both VOT/preaspiration and both measures of rate for these trials are consistent with the hypothesis that rate increases are causing the change in production. Additionally, variance measurements show that overall variance for VOT and preaspiration is lower after a phase switch than before, which is predicted given that the gestural relationship in the former category is more stable than in the latter.

Importantly, a switch from long to short VOT did not occur except in two cases, both of which can be explained better by an attested decrease in VOT at fast rates due to an overall compression of speech than by a phase change. This is exactly the pattern that was predicted to be absent.

Contrary to predictions, however, an increase in speech rate did not always lead to a shift in mode. Some trials maintained a consistently short VOT at high rates, and some trials had a long VOT even at slow rates. These divergent results are important, in that they show that attractiveness of coupling phase between two gestures is not the only force at work.

To explore this notion further let us first examine the trials in which no change took place. These trials show that there are forces acting to keep gestures in their original phasing relationships. The most obvious of the factors is the fact that phasing relations between gestures are specified in the grammar; the sequence is phonologically specified as /s/ + stop, and not the reverse. This will resist gestural reorganization. Additionally, the normative variety of Spanish does not allow for gestural realignment in these sequences, which may influence the speakers. This is especially true given the presence of an unfamiliar (and nonnative speaker) experimenter and the laboratory setting.

Next, let us turn to the examples with long VOT at slow rates. These trials may reflect the growing standardization, at least in casual speech, of the post-aspirated production. It is possible that the phenomenon was originally caused by a gestural re-phasing to a simpler mode of production but that the in-phase coupling has now become the standard. These forces interact in complex ways and even affect individual speakers differently at different times (take for instance the example in 3.1 of subject S2, who showed all three patterns).

There is strong evidence that preaspiration varies with changes in VOT. First, for trials where there is a switch from short to long VOT, there is a significant negative relationship in all groups between VOT and preaspiration, indicating that increases in VOT are accompanied by decreases in preaspiration duration. Second, comparisons of trials with only short VOT and only long VOT show significant differences in preaspiration duration. Short VOT trials have a mean preaspiration duration over 40 ms longer than trials with long VOT, and the mean preaspiration duration for the latter is less than 5 ms. This, again, is the predicted pattern as both preaspiration and long VOT are consequences of one glottal gesture.

Lastly, the current study found no evidence for compensatory changes in stop closure duration due to changes in VOT pattern (short vs. long). While there was a significant inverse relationship between preaspiration and stop closure duration for some groups, this is attributable to the general fact that decreases in the duration of one of two overlapping gestures without concomitant reduction in the duration of the other will result in less overlap between the two. This lack of evidence for compensatory lengthening is the pattern predicted by the phase-change hypothesis, though contradictory to some previous studies on Eastern Andalusian Spanish (Gerfen, 2002; Torreira, 2007a).

5. Conclusions

This study provides evidence that post-aspirated voiceless stops in Western Andalusian Spanish are due to a shift in the phasing relationship, from anti- to in-phase, between the glottal abduction gesture for coda /s/ and the oral closure gesture for the voiceless stop. By steadily increasing rate, which has been shown in other motor and speech activities to trigger such a phase shift, it was possible to trigger speakers to change from preaspiration and short VOT, the expected anti-phase production, to little to no preaspiration and long VOT, the expected in-phase production. As predicted, though contrary to previous studies, no evidence was found for any consonant lengthening of post-aspirated stops. However, distinct from findings in simple finger tapping studies, this phase shift did not occur in all trials, nor did it occur in one clean break. Rather, many speakers passed through a period of instability where productions wavered between short and long VOT (corresponding to a waver between anti- and in-phase).

While the current study provides evidence that a change in gestural coordination from an anti- to an in-phase relationship can help explain these postaspirated stops, the question of why such a change occurs only in Western Andalusian Spanish remains outstanding. As is the case with explaining any sound change, whether historical or ongoing, it is difficult if not impossible to explain *why* a particular change spreads throughout one dialect or group and not another. The current paper attempts only to provide an explanation for how a production of [t^h] for phonological /st/ may have originally arisen. This novel production, caused by inherent differences in stability of coordination patterns, must originally have occurred sporadically in Western Andalusian Spanish. Such productions may in fact occur in a similar sporadic way in other dialects, but only in Western Andalusian Spanish has it been (to a greater or lesser degree) phonologized. What phonologization means, in this case, is an open question. One possibility is that speakers of Western Andalusian Spanish have both

coordination patterns available for these sequences. This would explain the fact that many trials in the current study showed only short VOT or only long VOT productions, though perhaps not why it was possible to trigger a switch between the two modes in so many cases. An alternative proposal that might explain this triggering is that Western Andalusian Spanish speakers have a reduced *strength* of the anti-phase coupling between the gestures for /s/ and those for the voiceless stop. A reduction in strength of coupling would lead to greater instability due to random noise in the system and a higher probability of exhibiting a phase shift. This agrees well with what we see in Western Andalusian Spanish—productions of these sequences are highly variable (Ruch 2008, for example, finds at least eight possible distinct productions of these sequences), and often show a complete shift in phasing to the in-phase mode ([t^h]). Alternatively, both coordination patterns may be present, though with unequal coupling strengths. While further examination of these hypotheses remains an issue for future work, this study has shown that a change in relative phasing between the gestures involved in these sequences can explain the novel post-aspirated production. These results provide evidence that speech, like motor control in general, can be accurately modeled with coupled-oscillatory dynamics, and that using this model can provide insight into otherwise unexplained language variation and change.

Acknowledgments

This work was funded by NIH grant DC03172 and a grant from the Del Amo foundation. The author would like to thank Mary-Hely Cuenca and Marina Barrio at the Universidad de Sevilla for the use of their equipment and recording space, the phonetics lab group at the University of Southern California for their invaluable input, and Dani Byrd, Louis Goldstein, and Mario Saltarelli for help in guiding this project to completion.

References

- Bates DM. Fitting linear mixed models in R. *R News*. 2005; 5:27–30.
- Browman, CP.; Goldstein, L. Dynamics and articulatory phonology. In: Port, R.; van Gelder, T., editors. *Mind as motion: dynamics, behavior, and cognition*. Cambridge, MA: MIT Press; 1995. p. 175–193.
- Court ML, Bennett SJ, Williams AM, Davids K. Local stability in coordinated rhythmic movements: fluctuations and relaxation times. *Human Movement Science*. 2002; 21(1):39–60. [PubMed: 11983433]
- de Jong, KJ.; Lim, B-J.; Nagao, K. Phase transitions in a repetitive speech task as gestural recomposition. *IULC Working Papers*[On-line serial]. 2002. URL <https://www.indiana.edu/~iulcwp/pdfs/02-dejong06.pdf>
- Gerfen C. Andalusian codas. *Probus*. 2002; 14:247–277.
- Goldstein, L.; Byrd, D.; Saltzman, E. The role of vocal tract gestural action units in understanding the evolution of phonology. In: Arbib, M., editor. *From Action to Language: The Mirror Neuron System*. Cambridge, UK: Cambridge University Press; 2006. p. 215–248.
- Jiménez, R. *El Andaluz*. Madrid: Arco Libros; 1999.
- Kelso J. Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*. 1984; 246(6):R1000–R1004.
- Kelso JAS, Scholz JP, Schöner G. Non equilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters A*. 1986; 118(6):279–284.
- Kessinger R, Blumstein S. Effects of speaking rate on voice-onset time in Thai, French and English. *Journal of Phonetics*. 1997; 25:143–168.
- Maddieson, I. Phonetic Universals. In: Laver, J.; Hardcastle, WJ., editors. *The handbook of phonetic sciences*. Oxford: Blackwell; 1997. p. 619–639.
- Moya Corral JA. Noticia de un sonido emergente: La africada dental procedente del grupo -st- en andalucía. *Revista De Filología De La Universidad De La Laguna*. 2007; 25:457–465.
- Nabney, IT. *Advances in Pattern Recognition*. Londong: Springer; 2002. Netlab: Algorithms for Pattern Recognition. Code available from <http://www.ncrg.aston.ac.uk/netlab/>

- Nam, H.; Goldstein, L.; Saltzman, E. Self-organization of syllable structure: a coupled oscillator model. In: Pellegrino, F.; Marisco, E.; Chitoran, I., editors. *Approaches to phonological complexity*. Berlin/New York: Mouton de Gruyter; 2009. p. 299-328.
- Oliveira, L.; Marin, S. Patterns of velum coordination in Brazilian Portuguese. Paper presented at the Phonetics and Phonology in Iberia Conference; Barcelona, Spain. 2005.
- Pind J. Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues. *Perception & Psychophysics*. 1995; 57:291–304. [PubMed: 7770321]
- Romero, J. Unpublished PhD Dissertation. The University of Connecticut; 1995. Gestural Organization in Spanish. An experimental Study of Spirantization and Aspiration.
- Ruch, H. MA Thesis. Universität Zürich; 2008. La variante [ts] en el español de la Ciudad de Sevilla: Aspectos fonético-fonológicos y sociolingüísticos de un sonido innovador.
- Saltzman, E.; Nam, H.; Krivokapic, J.; Goldstein, L. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In: Barbosa, PA.; Madureira, S.; Reis, C., editors. *Proceedings of the 4th International Conference on Speech Prosody (Speech Prosody 2008)*; Campinas, Brazil. 2008.
- Swinnen SP, Dounskaia N, Verschueren S, Serrien DJ, Daelman A. Relative phase destabilization during interlimb coordination: The disruptive role of kinesthetic afferences induced by passive movement. *Experimental Brain Research*. 1995; 105(3):439–454.
- Torreira, F. Pre- and postaspirated stops in Andalusian Spanish. In: Prieto, P.; Mascaró, J.; Solé, MJ., editors. *Segmental and prosodic issues in Romance phonology*. Amsterdam/Philadelphia: John Benjamins; 2007a. p. 67-82.
- Torreira, F. Coarticulation between aspirated-s and voiceless stops in Spanish: an interdialectal comparison. In: Sagarra, N.; Toribio, AJ., editors. *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Somerville, MA: Cascadia Press; 2007b. p. 113-120.
- Tuller, B.; Kelso, JAS. Phase transitions in speech productions and their perceptual consequences. In: Jeannerod, M., editor. *Attention and performance XII*. Hillsdale, NJ: Erlbaum; 1990.
- Wimmers RH, Beek PJ, van Wieringen PCW. Phase transitions in rhythmic tracking movements: A case of unilateral coupling. *Human Movement Science*. 1992; 11 (1–2):217–226.

- We analyze the shift from pre- to post-aspiration in Western Andalusian Spanish voiceless stops.
- A change in the coordination of speech gestures is hypothesized to underlie this shift.
- We can trigger this shift by increasing the rate of production, as in other motor control studies.
- Variability in production is also consistent with the gestural-rephasing hypothesis.

Fig. 1. Schematic gestural organization for /VsCV/ sequence in Western Andalusian Spanish. In the upper scenario (a), the tongue tip closure gesture is subsequent (anti-phase) to the glottal spreading gesture, resulting in a period of aspiration both before and after the stop. In the lower scenario (b), the two gestures are simultaneous (in-phase), resulting in no aspiration before the stop and a longer period of VOT. These differences arise solely from the organization of the invariant gestures.

Fig. 2.

Fig. 2a (on left): Gaussian distributions fitted to VOT distribution across all tokens. The fit is better for two Gaussians (bottom panel) than for one (upper panel), which can be seen in the lower error for the former compared to the latter. Figure 2b (on right): Fitting additional Gaussian distributions beyond two to the data does not improve error.

Fig. 3. VOT and pre-closure aspiration for the two trials misidentified by the sorting algorithm. The shaded area represents productions classified as short VOT, while the white area represents those identified as long VOT.

Fig. 4.

Graph showing average VOT at each step for four trials from subject S2. As in fig. 2, the shaded area represents productions classified as short VOT, while the white area represents those identified as long VOT. Trial 2 shows consistently long VOT, trials 3 and 4 show consistently short VOT, and trial 1 shows a sudden and near categorical switch from short to long VOT.

Fig. 5.

Graph plotting VOT against total token duration for four trials from subject S2, with regression lines shown for each trial.

Fig. 6.

Fig. 6a (top): Shifting alignment of tongue tip closure gesture with regards to glottal spreading gesture causes preaspiration to vary inversely with VOT. On right, closure duration has shifted to the left, causing a decrease in preaspiration and an increase in VOT. 6b (middle): An increase in closure duration (right) causes a shorter VOT when the duration of the glottal spreading gesture remains constant. 6c (bottom): A decrease in the duration of both tongue tip and glottal gestures (perhaps due to speech rate) causes decrease in both preaspiration and VOT durations.

Fig. 7. Deviation from mean aspiration ratio, $VOT/(VOT + \text{preaspiration})$, before, around, and after the switch from short to long VOT.

Table 1

Results of a chi-squared test examining the different patterns of aspiration produced at the beginning and end of each trial.

		Final VOT	
		short	long
Initial VOT	short	18	29
	long	2	31

$$\chi^2 = 10.746, p = 0.001$$

Upper left box: number of trials with short VOT on first and last steps. Upper right: number of trials with short VOT on first step and long VOT on last. Lower left: number of trials with long VOT on first step and short on last. Lower right: number of trials with long VOT at both first and last steps