Journal of Phonetics (1996) 24, 99–111

998

Articulatory blending of lingual gestures

Joaquín Romero

University of Connecticut and Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, U.S.A.

Received 21st September 1994, and in revised form 27th September 1995

This paper presents some data on the articulatory nature of assimilatory processes in Spanish consonant clusters. The nature of those assimilations is explored in the light of some of the theoretical assumptions concerning assimilation suggested in the theory of Articulatory Phonology. Differences in articulatory configurations will be assessed qualitatively with respect to the different articulators associated with the production of a particular consonant or group of consonants. Also, the relationship between a consonant's specifications for constriction location and constriction degree will be studied for these Spanish data. The results point toward a view of assimilation that operates at the level of the whole gesture, not its spatial and temporal specifications, such that constriction location and constriction degree do not assimilate independently. This proposal is consistent with a view of articulatory gestures as the minimal units of phonological organization.

1. Introduction

1.1. *Views on assimilation*

Catford (1977) defines assimilation as "... the process by which one or both of two successive segments become more like the other" (p. 225). Given the continuous nature of speech, it is not surprising to find that assimilation is one of the most common phonological/phonetic processes in languages of the world. Phonological studies of assimilation abound for a large number of languages and from a variety of theoretical perspectives. In phonetic terms, too, assimilation has been investigated extensively, with or without reference to a particular phonological approach. Many of the existing accounts of assimilation propose ways to explain the mechanics of the phenomenon, that is, how to get from some theoretical underlying representation to the actual phonetic representation. Other studies have investigated the reasons why the result of a particular assimilation process should be one way or another.

Different explanations have been proposed as to why a given assimilation process should progress in one direction and not another; for example, in a sequence of two segments XY, why should X become more like Y, instead of Y becoming more like X? Grammont (1933) answered this question by appealing to the general principle of "articulatory strength": "Le phonème qui command est celui qui a plus de force ou de résistance ou de stabilité ou de faveur ... L'assimilation obéit à une seule loi:

0095-4470/96/010099 + 13 $12.00/0

to similar principles, often referred to as "articulatory constraints" (Recasens, 1990, 1991). In this view, the requirements or demands—articulatory, aerodynamic, etc.—of a segment determine whether assimilation takes place and what the resulting output is. In a different approach, Ohala (1990) argues for an acoustic/auditory basis for the explanation of the observed fact that, in a sequence of two consonants C1 and C2, assimilation of C1 to C2 is almost universally preferred to that of C2 to C1.

Another aspect of assimilation that has received significant attention in the phonetics/phonology literature is the issue of whether assimilation is a categorical process—as most phonological approaches imply (Chomsky & Halle, 1968; Clements, 1985)—or a gradient process, which seems to be more in accordance with the nature of speech. Recent work (Nolan, 1992; Farnetani & Busà, 1994) shows experimental evidence that assimilatory processes are hardly ever categorical substitutions of one unit—be it segment, autosegmental node, or feature—for another, even in cases of assimilation involving nasals, which have often been claimed to assimilate categorically (Harris, 1984.)

1.2. *Assimilation in articulatory phonology: gestural overlap*

The notion that most phonetic/phonological processes are indeed not categorical but gradient is at the core of the theory of articulatory phonology (Browman & Goldstein, 1989, 1990, 1992). This theory regards articulatory gestures as the primitives of phonological organization. Gestures are dynamically defined movements of the articulators in the vocal tract during the production of speech. One crucial implication of this view of sound structure is that gestures are specified both spatially and temporally. This is one of the major differences between this and other phonological theories, namely, the inclusion of time (duration) as an intrinsic property of gestures. In Browman & Goldstein's gestural model, gestures are abstract entities that are defined in terms of tract variables—dimensions for the articulator involved in the realization of a particular gesture, as specified in a task dynamics model (Saltzman & Kelso, 1987). Thus, an alveolar gesture, for example, is defined in terms of the tract variables TTCL (tongue-tip constriction location), TTCD (tongue-tip constriction degree) and stiffness, which can be seen as an indicator of the temporal characteristics—duration—of a gesture.

Based on these assumptions, commonly discussed phonological phenomena such as assimilation, substitution, deletion, or epenthesis can be easily explained as the result of two simple processes: overlap and blending of articulatory gestures over time, and reduction in gestural magnitude. Thus, in their account of the common nasal assimilation process whereby an alveolar nasal /n/ surfaces as a labial /m/ in front of a labial stop /p/, Browman & Goldstein claim that the resulting [mp] sequence is caused by the overlap of the labial gesture for the /p/ over the velum lowering gesture associated with the /n/. Consistent with this overlap explanation, it is also claimed that the lingual gesture for the /n/ does not disappear, but is merely hidden by the labial one. Some experimental evidence to support this claim is provided in Browman & Goldstein (1990).

A very similar account of precisely this kind of assimilation in Spanish is proposed in Navarro Tomás (1957). In his discussion of the behavior of [n] in front of other consonants, Navarro points out that, in an [nm] sequence—as in *inmenso* [imménso] "immense"—the tongue-tip gesture for the [n] is present but its acoustic output is hidden by the overlapping labial [m]. He also mentions the rate dependency of this type of assimilatory process, which is another indication of its non-categorical nature.

1.3. *Assimilation as a function of gestural organization*

In their discussion of the different levels of gestural organization, Browman & Goldstein make some predictions regarding the articulatory characteristics of some of the most common assimilation processes in languages of the world. Based on the specifications of the model, different articulatory configurations are predicted depending on the tier affiliation of the gestures involved, that is, the grouping of gestures according to their articulatory characteristics or function.

Browman & Goldstein make a distinction between articulatory tiers and functional tiers. Articulatory tiers are defined on the basis of articulatory independence and correspond to the major articulator that is activated during the production of a particular gesture. Major articulators include the lips, the tongue tip, and the tongue dorsum. Functional tiers, on the other hand, separate consonants and vowels into two different levels. This allows for a more accurate specification of the particular phasing relationships between consonants and vowels.

When there is overlap in time of two gestures that are on different articulatory tiers, it is predicted that they do so without perturbing each other's trajectories, since they concern different tract variables. Such is the case in the Spanish [nm] example mentioned above, where the two overlapping gestures, the tongue-tip gesture for the [n] and the lip gesture for the [m], are executed with two articulators that are to a large extent independent of each other: the tongue tip and the lips. When the overlapping gestures are on the same articulatory tier, however, they cannot overlap without somehow affecting each other's trajectories. In such cases, the result of the overlap will show blending of the characteristics of the two gestures. Thus, we can use the particular articulatory organization of gestures (their tier affiliation) as a frame of reference to study the articulatory characteristics of assimilatory processes that take place when gestures overlap in time.

Browman & Goldstein (1990) mention a potential case of blending of the constriction location of two same-tier gestures that overlap in time, as in the English sequence *ten themes*. In this case, it is predicted that the result of the overlap of the constriction location for alveolar [t] and dental [θ] would be a location somewhere in between those of the individual gestures. No explicit mention is made, however, of what the result of such a case of gestural overlap would be for the gestures' constriction degree specifications. It is to be expected, nevertheless, that in such a case both constriction location and constriction degree would blend. This is so precisely because of the assumption in articulatory phonology that the gesture is the minimal unit of contrast. Such an assumption necessarily implies that constriction location and constriction degree—two attributes of the gesture—cannot be manipulated independently (Browman & Goldstein, 1989.)

Support for this in autosegmental terms is presented in Padgett (1991). His account of the role of stricture in feature geometry suggests that the autosegmental equivalent of constriction degree, that is, the feature [continuant], is more accurately placed as dependent of an "articulator node", rather than at a higher level in the hierarchy, as other previous proposals had suggested (Sagey, 1986; McCarthy, 1988). Evidence for such a move comes in part from the observation that place assimilation in a sequence of nasal plus fricative is, at best, a rare phenomenon.

For those cases in which such assimilation seems indeed to take place, Padgett proposes a [+continuant] specification for the nasal. Such a case would be, for example, Spanish *triunfo* tri[uɱf]o "triumph", where alveolar [n] assimilates to labiodental [f]. In other words, stricture assimilates also in this case, so that the nasal takes on some of the characteristics of the fricative. In contrast, English *infamous* [Inf]amous shows no assimilation of either place or stricture. Even though Padgett does not provide any experimental evidence, this is precisely what articulatory phonology would predict: either both place (CL) and stricture (CD) assimilate, as in Spanish *triunfo,* or neither does, as in English *infamous.* A similar situation would likely obtain in the *ten themes* case mentioned above.

Thus, the goal of the experiment described below is to provide some preliminary data on the articulatory mechanisms that are at play when two consonants with the same or different specifications for constriction location and constriction degree appear together in a cluster. Consistent with articulatory phonology's assumption that the gesture is the minimal unit of contrast and that, therefore, constriction location and constriction degree cannot be manipulated independently, it is predicted here that blending of the characteristics of the two gestures will take place at the gestural level or not at all. Moreover, it is hypothesized that the indivisibility of the gesture is not limited to constriction location and constriction degree, but possibly includes other aspects of the gesture, in particular, stiffness.

In order to investigate these points, an experiment was designed that looked at the articulatory characteristics of some lingual consonants in Castilian Spanish, in particular, those involving a tongue-tip gesture. Some articulatory studies of tongue-tip consonants have been reported in the literature, mostly in English (Bladon & Nolan, 1977; Cohen & Perkell, 1986, among others). Recasens (1993), however, presents articulatory data for Catalan that are similar to those analyzed here for Spanish. The results obtained in the present study are expected to match those of Recasens (1993) to a large extent.

2. Data

Castilian Spanish tongue-tip consonants can be separated into two categories according to their constriction location: dental/interdental and alveolar. Dental/interdental consonants are stops, /t/, /d/, and fricative /θ/; alveolar consonants are nasal /n/, lateral /l/, fricative /s/, and the rhotics /r/ and /ɾ/—a trill and a tap, respectively. This situation results in potentially interesting cases of assimilation when consonants from the two groups are in contact within a cluster, given the various constriction degrees.

Thus, the following consonants and consonant clusters were included in the design: /l/, /d/, /t/, /s/, /ɾ/, /ld/, /sd/, and /ɾd/. The experimental material consisted of a list of sentences containing a target word inside the carrier phrase

"Diga____cada vez" ("Say____each time"). The target words were always trisyllabic nonsense sequences of the form CV(C)–CV–CV, where the hyphens indicate syllable boundaries. The consonants under study occupied the highlighted positions. Two vowel contexts were used, /e/ and /o/, a mid front and a mid back vowel, respectively. Also, stress was consistently placed on the second syllable. One native speaker of Castilian Spanish read each list of sentences five times. Before proceeding with the experimental details, however, it is necessary to point out two language-particular issues regarding some of these consonants, which will be useful in the ensuing discussion of the articulatory patterns obtained in the experiment.

On the one hand, Spanish voiced stop /d/—like /b/ and /g/—is characteristically reduced (spirantized) in most contexts except in absolute initial position and when it is preceded by a homorganic consonant with a complete closure. In all other contexts it surfaces as a continuant, characteristically as a voiced approximant (Martínez Celdrán, 1984). In addition to a change in constriction degree, this reduction process results also in a shorter duration and a slight change in constriction location. Even though the shift in constriction location affects all three consonants (Romero, 1995), it has been remarked most often for the dental case (Navarro Tomás, 1957), for which a slightly more advanced constriction is commonly acknowledged. Thus, in the contexts included in this experiment, /d/ is realized as an approximant in intervocalic position and following /s/ and /ɾ/—[eð̞e], [ezð̞e][1] and [erð̞e]—but as a stop after /l/—[elde]. Thus, the stop allophone [d] does not occur naturally in intervocalic position and, therefore, it could not be included in the design. Instead, [t] was included, as an example of a non-reduced dental stop, with the assumption that the constriction locations of [t] and [d] are the same.[2]

On the other hand, the two rhotic consonants—the trill /r/ and the tap /ɾ/—are contrastive only in intervocalic position; in all other positions they neutralize: only the trill occurs in absolute or word-initial position, while in coda position, it is the tap that occurs almost exclusively. Thus, in the contexts involving rhotics analyzed in this study, only the tap occurs. For that reason, the symbol /ɾ/ is used throughout, even though in the [ɾð̞] cluster, there is no contrast between the tap and the trill.

2.1. *Experimental procedures*

The experiment was performed using the EMMA (ElectroMagnetic Midsagittal Articulometry) technique for tracking articulatory movements (Perkell, Cohen, Svirsky, Matthies, Garabieta & Jackson, 1992; Löfqvist, Gracco & Nye, 1993). Coils were placed, along the midsagittal line, on the subject's upper and lower lips, lower incisors—as an estimate of jaw movement—upper incisors and bridge of the nose—as references for head movement correction—and at two points on the surface of the tongue, identified as tongue tip (TT) and tongue body (TB). The

[1] An independent process of anticipatory voicing assimilation is responsible for the surfacing of /s/ as [z] in this case.

[2] For an extensive discussion of numerous issues regarding spirantization in Spanish, including those concerning the nature of the canonical or underlying forms, see (Romero, 1995).

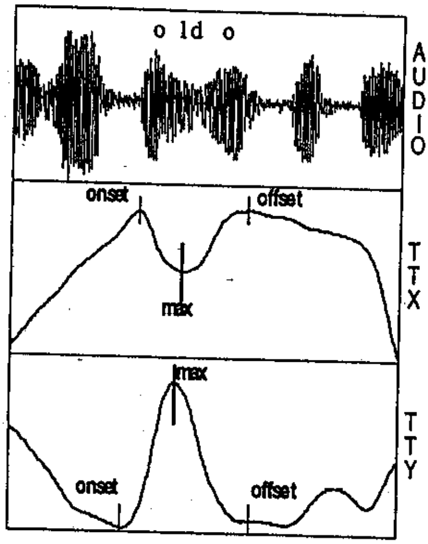

Figure 1. Example of the event identification procedure. The top panel shows the audio signal, with the identified target sequence [oldo]. The middle panel shows the synchronized section in the tongue-tip horizontal (TTX) movement signal; here, a more advanced position in the horizontal dimension corresponds to a lower point in the signal. The bottom panel shows the corresponding section of the tongue-tip vertical (TTY) movement signal; here, a higher point in the horizontal dimension corresponds to a higher point in the signal. The labels in the movement signals indicate the maximum position associated with the target consonant(s)—max—and the points associated with the beginning and end of the consonantal movement—onset and offset, respectively.

tongue-tip coil was placed at approximately 0.5 cm from the very tip of the tongue; the tongue-body coil was placed approximately 2 cm behind the tongue-tip one.

The movement data were digitized at a sampling rate of 625 Hz, while the simultaneous audio signal was digitized at 20 000 Hz. Following input, the signal from each coil was separated into a horizontal (X) and a vertical (Y) component. The ensuing discussion will make reference to the articulatory movements corresponding to the consonant(s) in the target words, which were identified with the help of the acoustic signal and the zero crossings of the corresponding velocity profiles. Fig. 1 illustrates how the articulatory events corresponding to the target consonants were identified in the movement signals. For the purposes of this paper, only the data from the dentoalveolar sequences will be considered, that is, the data obtained from the tongue-tip (TT) coil. Comparisons between tongue-tip consonants in isolation and in consonant clusters will allow us to observe the behavior of same-tier gestural overlap.

2.2. *Articulatory patterns*

In this section, the different articulatory patterns shown in the tongue-tip signals will be examined for the contexts included in the experimental design. Qualitative observations will be made with reference to the two-dimensional displays of tongue

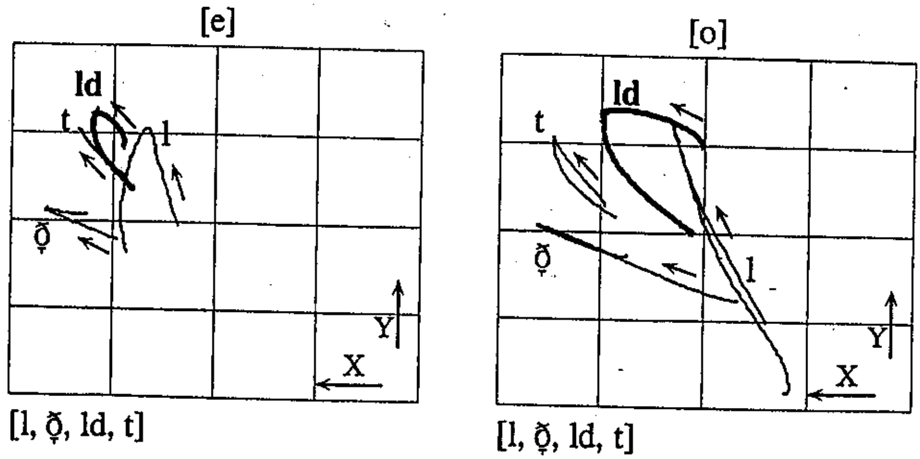

Figure 2. Two-dimensional display of tongue-tip movement paths for intervocalic [l], [ð̞], [ld], and [t] in the context of vowel [e] on the left and in the context of vowel [o] on the right.

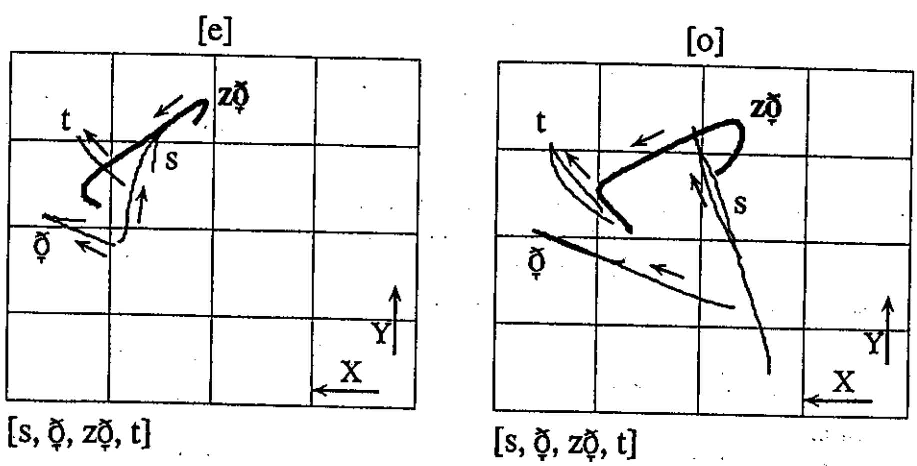

Figure 3. Two-dimensional display of tongue-tip movement paths for intervocalic [s], [ð̞], [zð̞], and [t] in the context of vowel [e] on the left and in the context of vowel [o] on the right.

movement paths shown in Figs. 2–4. The XY plots of tongue movement paths displayed in each figure correspond to the V(C)–CV sequences in the target words. All figures show the same coordinate space, where every division in the grid represents 0.5 cm. The vertical and horizontal arrows in the lower right corner of each figure point to a higher and a more fronted position, respectively. The smaller arrows next to the paths indicate the direction of the movement for each individual context, also identified by the corresponding phonetic symbols. Thus, most paths start out from a rather back position for the first V, move forward and/or upward to achieve a maximum point for the (C)–C and return to a more retracted and/or

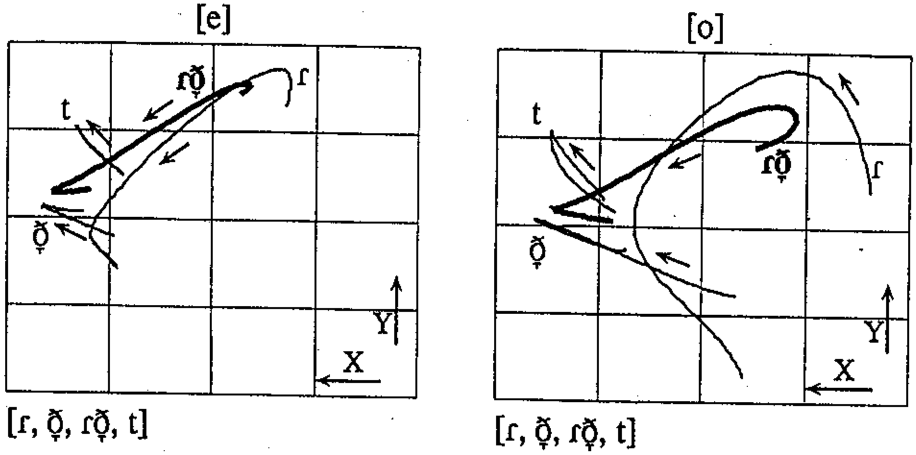

Figure 4. Two-dimensional display of tongue-tip movement paths for intervocalic [ɾ], [ð̞], [ɾð̞], and [t] in the context of vowel [e] on the left and in the context of vowel [o] on the right.

lower position for the second V. In each figure, thin lines are used to represent single consonants; clusters are identified by a thicker line.

In all cases, each movement path represents an average of all five tokens of the same utterance. Averages were obtained using the following procedure. A lineup point was identified in the movement signal ("max" point in Fig. 1) corresponding to the maximum point of movement associated with the target consonant(s). The temporal intervals between the lineup point and the points identified as onset and offset of the movement for the consonant (Fig. 1) were calculated. The mean of those intervals for the five tokens was used as the temporal constant in the averaging procedure. Variability across tokens was, in general, very small; the largest amount of variability corresponded to the approximant [ð̞], which is not surprising given the spirantization facts discussed above. Figs. 2–4 show consonants in the context of vowel [e] on the lefthand side graph and in the context of vowel [o] in the righthand side graph.

Tongue-tip movement paths for [l], [ð̞], [ld], and [t] are represented in Fig. 2. Leaving aside for now the variation due to the vowel contexts, a clear difference can be observed in the shape of the movement for the single consonants. The paths for [l] show a mostly vertical movement, with constriction locations that are less advanced than those of the stop [t]. This corresponds with the above mentioned alveolar constriction location of [l], as opposed to dental [t]. Even more advanced are the constriction locations of [ð̞], which is to be expected given the spirantization issue explained earlier. The movements corresponding to the [ld] clusters show evidence of a single constriction, with locations that are more advanced than single [l], but less than single [t]. Thus, the movements observed for the [ld] clusters indicate the existence of overlap of the gestures for the two consonants. As a consequence, the characteristics of the two gestures blend in an intermediate constriction location, which confirms the predictions of articulatory phonology.

The main characteristics of the movements of both single consonants and clusters are preserved across the two vowel contexts. Vowel context seems to have little effect on the general constriction location of tongue-tip consonants. The only obvious differences across vowels have to do with the extent of the movement. Because of the back character of [o], the tongue tip is more retracted during the production of this vowel than for [e], which means that the tongue tip has to travel further to reach the dentoalveolar region.

Fig. 3 displays movement paths for [s] and [zð̞],[3] in addition to those for [ð̞] and [t], which are the same as the ones in Fig. 2. Movements for single [s] are quite similar to those seen in the previous figure for [l]. Clearly, [s] also shows an alveolar constriction degree distinct from dental [t] and the more advanced approximant [ð̞]. The [zð̞] cluster, however, presents a very different picture. Here the tongue tip moves from a position in the vicinity of single [s] to a position approaching the dental constriction locations of [t] or [ð̞]. This sliding movement reflects the existence of little or no overlap—no assimilation—between the two tongue-tip gestures, which seem to be articulated sequentially instead.

Despite the lack of overlap between the gestures, the constriction locations of the two consonants within the clusters are slightly different from those of the consonants in isolation. The constriction location of [s] in the cluster is considerably more retracted than in isolation. It is not clear from these graphs why that should be so, but it is conceivable that it is simply a mechanical consequence, a slight overshoot in preparation for the sliding movement. As for the second element, in spite of its spirantized character, the point of articulation fluctuates between those of [t] and [ð̞]. Perhaps this means that what is relevant in the realization of spirantized [ð̞]—or any other spirantized stop—is not so much the exact constriction location, but its continuant character.

As in the previous figure, the effects of vowel context are small, mostly again having to do with the extent of the movement and, in the case of the cluster, the constriction location of spirantized [ð̞], which is between that of single [t] and [ð̞] in the [e] context, but almost identical to [t] in the [o] context.

Fig. 4 displays movement paths for [ɾ] and [ɾð̞], as well as for [t] and [ð̞] just as in the two previous figures. The movements corresponding to the single tap [ɾ] show wide paths reminiscent of the sliding movements of the [zð̞] clusters in Fig. 3.[4] Precisely because of the shape of the movement, it is hard to identify the exact constriction location of [ɾ]. However, in spite of its traditional classification as an alveolar, it seems that the constriction location of this consonant is more retracted than that of either [l] or [s], at least judging by the location at the highest point in the path. The movement of the tongue tip for the [ɾð̞] cluster starts at a similar point as for single [ɾ], and then slides forward until it achieves a constriction location similar to that of spirantized [ð̞]. Thus, just as in [zð̞] above, it seems that the [ɾð̞] cluster is realized as a sequence, with minimal overlap and no blending of the individual gestures' constriction locations.

In summary, the articulatory patterns for the clusters shown in Figs. 2–4 exhibit two clearly distinct behaviors. On the one hand, [ld] shows one constriction, which is interpreted as a sign of overlap of the individual [l] and [d] gestures. Consequently, blending of the individual gestures' constriction locations takes place. On the other

[3]As mentioned earlier, the reason for the voicing of /s/ in this context is an anticipatory voicing assimilation process.

[4]This seems to imply that, in spite of the traditional classification of this consonant as a tap (Ladefoged, 1971), its articulatory characteristics suggest a "flapping" rather than a "tapping" motion.

hand, the [zð̞] and [ɾð̞] clusters seem to be articulated sequentially, with very little temporal overlap and no blending of the characteristics of the individual gestures. These two types of behavior are reminiscent of Catford's distinction between "shifting" and "sliding" articulatory "accommodation", a distinction also invoked by Holst & Nolan (1993) in their description of /s/ to /ʃ/ assimilation in English. The fact that the articulatory patterns obtained here are in many respects comparable to those found in Recasens (1993) for Catalan, provides some further support for the methodology employed in this experiment.

3. Discussion

The previous section examined the articulatory behavior of some Spanish lingual consonants in a variety of contexts. In this section, the results from the observation of the articulatory behavior of these consonants will be discussed in the light of the theoretical issues discussed in the introduction concerning articulatory phonology's view of assimilation.

The articulatory patterns exhibited by the same-tier clusters examined in this paper show a clear split between the cases where there seems to be temporal overlap of the two gestures and, therefore, blending of their respective constriction locations, as in [ld], and those where the overlap seems to be minimal and no blending of the characteristics of the individual gestures can be observed, as in [zð̞] and [ɾð̞]. Thus, assimilation seems to take place in the first case but not in the other two. Given that all three cases are instances of clusters involving two same-tier gestures, why then do they behave differently? Evidently, there is something about the articulatory nature of [l] that is different from that of [s] and [ɾ] and that causes the observed difference.

Figs. 2–4 allow for observation of the constriction locations associated with each of the consonants and consonant clusters. Unfortunately, this kind of display does not provide information about constriction degree. It is probably safe, however, to assume what is commonly known about these segments, that is, that they differ in their specifications for constriction degree. The tongue-tip gesture for [l]—leaving aside the issue of laterality, which might or might not involve a separate tongue-body gesture—has a complete closure; in articulatory phonology terms, a [closed] constriction degree (Browman & Goldstein, 1989). In contrast, [s] is specified as having a [critical] constriction degree, which reflects its fricative nature. The constriction degree of [ɾ], though to the author's knowledge not explicitly stated by Browman & Goldstein, is likely also [closed], since there is at some point a complete closure. The issue of the constriction degree of /d/ is controversial. However, following the view of spirantization as a non-categorical reduction process (Romero, 1995 and references therein), it is assumed here that the canonical constriction degree specification for /d/ is that of a stop, namely, [closed]. The continuant nature of [ð̞] is the result of a reduction in the magnitude of the gesture.

Given this scenario, it is straightforward to see that assimilation—resulting in blending of constriction location—takes place in the cluster whose components share the same specification for constriction degree. In contrast, assimilation does not take place in the clusters whose components do not share the same constriction degree.

At this point, it is necessary to remember articulatory phonology's view of the gesture as a unit and the assumption that constriction location and degree cannot be manipulated independently. Consistent with that position, these data show that assimilation takes place at the gestural level, not the tract variable level.

In articulatory phonology notation, the tract variables that need to be specified for the individual gestures involved are as follows, where TTCL and TTCD stand for tongue-tip constriction location and tongue-tip constriction degree, respectively; in parentheses are the particular specifications (gestural descriptors) for each tract variable. In the case of /d/, a voiced dental stop, the tract variables involved are TTCL(dental) and TTCD(closed). In /l/, a voiced alveolar lateral, the tract variables are TTCL(alveolar) and TTCD(closed). In /s/, the tract variables are TTCL(alveolar) and TTCD(critical), while for /ɾ/ they are TTCL(alveolar) and TTCD(closed).

Thus, looking first at the [ld] cluster, where blending indeed takes place, we have two gestures with the same specification for constriction degree—TTCD(closed)—but different specifications for location. The two can overlap in time because the specifications for CD are compatible; there is blending of the two gestures, which in this case results in a constriction location somewhere between the dental and alveolar regions. In the [zð̞] case, on the other hand, the specifications for CD are not compatible—critical *vs.* closed. This seems to prevent any large amount of overlap from taking place in this kind of cluster. Because the two CDs cannot blend, the CLs do not either.

We are left, then, with the last case, [ɾð̞], where, following the same reasoning as with [ld] and [zð̞] one should expect assimilation, since the TTCDs are compatible; as we know from the data, however, this is not the case. Again, there has to be something else about /ɾ/ that prevents it from behaving like /l/. A possible explanation might have to do with the consonant's specifications for duration. In the gestural model, duration is typically characterized by "stiffness", which, like CL and CD, is also an integral part of a gesture. Recasens (1991) shows palatographic evidence that taps are indeed very short consonants. Navarro Tomás (1918) provides some general duration values for all Spanish consonants: in his account, the average duration of [ɾ] is about 25 ms, compared to approximately 72 ms for [d], 65 ms for [l], and 92 ms for [s]. Thus, there is a large difference in duration between [ɾ] and [d], which can be seen as an indication that the two gestures' specifications for stiffness are not compatible, which in turn seems to prevent assimilation from taking place in a [ɾð̞] cluster.

This supports the hypothesis that assimilation takes places at the gestural level, for even when CDs are compatible, blending does not take place if the gestures' specifications for stiffness are not. A comparison of the duration values for [l] and [d], on the one hand, and for [s] and [d], on the other hand, provide further evidence for the validity of the assimilation patterns discussed above for [ld] and [zð̞]. Duration values for [l] and [d]—65 and 72 ms, respectively—are reasonably similar; assimilation takes place in [ld]. The difference in duration between [s] and [d]—72 and 92 ms, respectively—is much larger; assimilation does not take place in [zð̞]. Since these are not measures of articulatory duration, it is hard to make definitive conclusions regarding stiffness. It does not seem impossible, however, to speculate that these duration specifications play a role in the assimilatory patterns observed for [ld] and [zð̞] as well.

4. Conclusions

The goal of this paper has been to test some of the implications of a theory of phonology that regards assimilation as a result of articulatory gestures overlapping in time. Knowing that spoken language is a continuous process and that speech units are typically articulated in a coarticulated—overlapping—fashion, such a view of assimilation promises to be more in accordance with underlying principles of speech production than other approaches that do not consider the temporal aspect of speech in their formulation. Within this general framework, the precise points to be investigated were concerned with differences in assimilatory processes that can be related to the role of the gesture as the minimal unit of phonological organization.

Despite the preliminary nature of the results presented here (data for only one subject were analyzed), it has been shown that the different tract variables that specify a gesture cannot be manipulated separately. This confirms Browman & Goldstein's prediction that assimilation processes cannot affect constriction location or constriction degree separately and that one cannot assimilate without the other, as in the Spanish /sd/ cases studied here. This idea is reinforced by the behavior observed in /ɾð̞/, where we saw that the claimed cohesiveness of gestures might not be restricted to the specifications of constriction location and constriction degree, but also includes other inherent gestural attributes such as stiffness.

The particular gestural specifications that need to be compatible in order for assimilation processes to take place are probably related to the inventory of contrastive units that a particular language has. In the Spanish case, the stiffness specification seems to play a decisive role in assimilations involving /ɾ/ and /d/, and perhaps also /l/ and /s/, but it might not be a relevant descriptor in the gestural specifications for other segments. Also, assimilation will be expected to differ across languages insofar as the specifications for gestures—tract variables and/or gestural descriptors—differ across languages. Finally, it might also be up to the particular language whether assimilation—gestural overlap—is blocked or not by incompatible gestural specifications; such seems to be the case in Spanish. Other languages might choose to assimilate whenever two same-tier gestures appear in a cluster, resulting in blending of the entire gesture; languages that favor the proliferation of geminates might fit that pattern. The data presented in this paper suggest, however, that, regardless of the language and its gestural specifications, gestures behave as whole, cohesive units.

The interpretations presented here regarding the nature of assimilation in same-articulator clusters are not incompatible with existing accounts that invoke articulatory constraints or requirements as the explanation for assimilatory patterns. It is believed, however, that the gestural approach investigated here makes more explicit and formalized predictions as to when assimilation occurs and when it does not within a a particular language. In any case, it seems that this type of assimilation has a clear articulatory base, which could be supported by the acknowledgement made in Ohala (1990) that the acoustic/auditory explanation for assimilation might not work for same-articulator cases.

I am grateful to Daniel Recasens, Ignatius Mattingly, Louis Goldstein, and three anonymous reviewers for their comments on earlier versions of this paper. Clearly, any possible inadequacies are exclusively the author's responsibility. This work was supported by NIH Grants DC-00121 and HD-01994 to Haskins Laboratories.

References

Bladon, R. & Nolan, F. (1977) A video-fluorographic investigation of tip and blade alveolars in English, *Journal of Phonetics*, **5**, 185–193.

Browman, C. & Goldstein, L. (1989) Articulatory gestures as phonological units, *Phonology*, **6**, 201–251.

Browman, C. & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology I* (J. Kingston & M. Beckman, editors), pp. 341–376. London: Cambridge University Press.

Browman, C. & Goldstein, L. (1992) Articulatory Phonology: an overview, *Phonetica*, **49**, 155–180.

Catford, J. C. (1977) *Fundamental problems in phonetics*. Bloomington, IN: University of Indiana Press.

Cohen, M. & Perkell, J. (1986) Palatographic and acoustic measurements of the fricative consonant pair /s/ and /ʃ/. In *Proceedings of the 12th international congress on acoustics*. Toronto.

Chomsky, N. & Halle, M. (1968) *The sound pattern of English*. Cambridge, MA: MIT Press.

Clements, G. (1985) The geometry of phonological features, *Phonology yearbook*, **2**, 223–250.

Farnetani, E. & Busà, M. G. (1994) Consonant-to-consonant interactions in clusters: Categorical or continuous processes? *Quaderni del Centro di Studio per le Ricerche di Fonetica*, **XIII**, 220–245.

Grammont, M. (1933) *Traité de phonétique*. Paris: Librairie Delagrave.

Harris, J. W. (1984) Autosegmental phonology, lexical phonology and Spanish nasals. In *Language sound structure* (M. Aronoff & R. Oehrle, editors), pp. 67–82. Cambridge, MA: MIT Press.

Holst, N. & Nolan, F. (1993) The influence of syntactic structure on [s] to [ʃ] assimilation. Paper presented at *LabPhon IV*, Oxford.

Ladefoged, P. (1971) *Preliminaries to linguistic phonetics*. Chicago, IL: University of Chicago Press.

Löfqvist, A., Gracco, V. L. & Nye, P. W. (1993) Recording speech movements using magnetometry: one laboratory's experience. *Forschungsberichte des instituts für phonetik und sprachliche Kommunikation der Universität München*, **31**, 143–161.

Martínez Caldrán, E. (1984) Cantidad e intensidad en los sonidos obstruyentes del castellano: hacia una caracterizatión acústica de los sonidos aproximantes. In *Estudios de fonética experimental*, pp. 71–129. Barcelona; Promociones Publicaciones Universitarias.

McCarthy, J. (1988) Feature geometry and dependency: a review, *Phonetica*, **45**, 84–108.

Navarro Tomás, T. (1918) Diferencia de duración entra las consonantes españolas, *Revista de Filología Española V*, 367–393.

Navarro Tomás, T. (1957) *Manual de Pronunciación Española*. New York: Hafner.

Nolan, F. (1992) The descriptive role of segments. Evidence from assimilation. In *Papers in laboratory phonology II* (G. Docherty & D. R. Ladd, editors), pp. 261–280. Cambridge: Cambridge University Press.

Ohala, J. (1990) The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology I* (J. Kingston & M. Beckman, editors), pp. 258–271. London: Cambridge University Press.

Padgett, J. (1991) *Stricture in feature geometry*. University of Massachusetts, Amherst.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I. & Jackson, M. (1992) Electro-magnetic midsagittal articulometer systems for transducing speech articulatory movements, *Journal of the Acoustical Society of America*, **92**, 3078–3096.

Recasens, D. (1990) The articulatory characteristics of palatal consonants, *Journal of Phonetics*, **18**, 267–280.

Recasens, D. (1991) On the production characteristics of apicoalveolar taps and trills, *Journal of Phonetics*, **19**, 267–280.

Recasens, D. (1993) Coarticulació i assimilació en fonologia. Dades articulatòries sobre els grups consonàntics amb C2 = /d/ en català. *Unpublished manuscript*.

Romero, J. (1995) *Gestural organization in Spanish. An experimental study of spirantization and aspiration*. Ph.D. Dissertation, University of Connecticut.

Sagey, E. (1986) *The representation of features and relations in nonlinear phonology*. Ph.D. Dissertation, M.I.T.

Saltzman, E. & Kelso, J. A. S. (1987) Skilled actions: a task dynamic approach, *Psychological Review*, **94**, 84–106.