A CROSS-LINGUISTIC STUDY OF BLENDING PROCESSES IN CONNECTED SPEECH

M.J. Solé and E. Estebas
Laboratori de Fonètica, Universitat Autònoma de Barcelona, Spain

ABSTRACT

Assimilatory and blending processes in connected speech in English and Catalan in a variety of conditions (slow vs fast speech, functional vs lexical words and oral vs nasal segments) were studied. The intergestural adaptation of alveolar+dental/palatal clusters were analyzed with simultaneous EPG, acoustic and EGG data. The results show gradient blending processes rather than discrete substitution of features. Sliding trajectories of two strongly coarticulated gestures and adaptation of the constriction location of C1 to that of C2 were found. The interactions betwen the two consonants in the cluster were significantly affected by the oral/nasal nature of C1 and to a lesser extent by rate and word category. The observed crosslinguistic differences require an explanation of blending in terms of language dependent structural factors.

INTRODUCTION

This study investigates assimilatory and blending processes in connected speech in English and Catalan in the light of current phonological theories. The modelling of connected speech processes has particular relevance for speech technology and for developing models of speech production.

Blending has traditionally been represented in phonological analyses as a categorical process (expressed in terms of phonological rules or feature linking processes) by which some feature of the segment is replaced by another feature. Such view assumes a discrete change in the input to speech production and in the neuromuscular commands to articulate the segment. Models of speech production and gestural phonology, on the other hand, claim that blending processes involve gradient modifications resulting from overlap of two competing contiguous gestures involving the same articulator, with no change in the input to speech production. Such overlap will show intergestural accomodation giving rise to a shift in the articulatory configuration of C1 only, a sliding movement of two strongly coarticulated gestures or one single gesture for the cluster with an intermediate target. Articulatory and acoustic data on blending in connected speech [1, 2] show residual alveolar gestures for the alveolar provide support for the gradient nature of assimilatory processes. In this paper the articulatory behaviour of blending processes -- i.e., those involving two successive consonantal gestures produced with the same articulator -- was examined in English and Catalan across different speech rates. The aim of the study was to determine 1) whether blending processes involve a planned reorganization of articulatory movements or the modification, due to overlapped instructions and simultaneous articulation, of the individual trajectories of C1 and C2, and 2) the gradient vs categorical behaviour of blending processes in a variety of conditions: slow vs fast speech, oral vs nasal alveolars, functional (e.g. *on thoughts*) vs lexical (e.g. *John thought*) words and English vs Catalan.

METHOD

Simultaneous EPG, acoustic and electroglottographic data were obtained for six repetitions of oral/nasal alveolar consonants that can occur in word-final position in these languages followed by a word-initial consonant involving the same articulator (VC1#C2V). Comparable VC1#V and V#C2V utterances were also analyzed. Low vowel contexts were chosen to avoid anterior contacts due to high vowels. The interactions between C1 and C2 were studied in the following clusters: /n/+/tʃ/ in English and Catalan, /t/+/tʃ/ in Catalan and /d/+/tʃ/ in English (in Catalan only voiceless obstruents are possible in word final position; in English word final /d/ was chosen to avoid (pre)glottalization of final /t/), /n/ + /t/ in Catalan (apicoalveolar + apicodental stop) and /n, d/+/θ/ in English. The test sequences involved meaningful phrases spoken by two native speakers each of English (speakers AR and MM) and Catalan (speakers MJ and AF) at slow and fast rate. Overall 1056 tokens were analyzed. Variation in the degree and distribution of linguopalatal contact in the two consonants in the cluster was analyzed using the EPG contact index method [3] and compared to that of canonical VC1#V and V#C2V utterances. For each consonant the constriction location and constriction degree was assessed by means of the contact anteriority (CA) index (since all consonants involved are produced in the dento-palatal zone), and the contact centrality (CC) index. CA reflects variation in the degree of contact fronting and CC reflects variation in the degree of contact from both sides of the palate towards its median line. The value of the two indices increases as the linguopalatal contact area becomes larger or more anterior (CA), or approaches the central zone or shows a larger central contact area (CC). The contact indices were calculated for the EPG frames showing maximum contact (PMC).

To examine whether the modification of the articulatory configuration of consonants in the cluster is the result of blending of the two gestures, or rather an assimilatory process we compared the CA and CC indeces at the PMC in C1 (and C2) in the cluster with those of intervocalic C1 (and C2), and the CA and CC indeces of C1 in the cluster with those of C2 in the cluster. Single/cluster X Rate (and Word type) ANOVAS were performed for each speaker and each cluster type separately in order to find out whether single consonants and consonants in clusters showed significant differences in contact indeces, and whether significant effects were rate and word category dependent. C1/C2 X Rate (and Word category) ANOVAS were performed in order to determine whether the differences in contact indeces between C1 and C2 in the cluster were significant, and dependent on rate or word category.

RESULTS

First we will comment on differences in linguopalatal contact indeces for C1 (and C2) in isolation and in clusters, and then we will analyze whether such differences lead to blending of the the two individual gestures in clusters. Finally rate and word category effects will be analyzed for each cluster type.

English /nθ/ and /dθ/ clusters

For the two speakers, /n, d/ in the cluster show a significant increase in CA values vis-à-vis intervocalic position, which indicates that the alveolar segment becomes more anterior due to the effect of dental /θ/. Speaker AR also shows a significant decrease in central contacts (CC) for /n/ in the cluster (only in function words) making the constriction degree of this segment more similar to that of fricative /θ/, which shows virtually no central contacts intervocalically. Speaker MM, on the other hand, shows no variation in CC indeces for the first segment in the cluster, but the central contacts for the second segment are significantly higher than in singleton /θ/, indicating that the constriction degree of C2 becomes more similar to that of the alveolar stop. Thus, the adaptation of the constriction degree for both segments in the cluster go in opposite direcions for the two speakers. CA values of /θ/ preceded by /n, d/ increase significantly vis-à-vis intervocalic position indicating a more constricted articulation of /θ/ in clusters. Although the articulatory configurations of both segments in the cluster are modified, becoming more similar to each other, there seems to be no blending of the constriction location or constriction degree of original /n, d/ and /θ/ since the CA and CC values of /θ/ in the cluster are significantly lower than those for the first segment. Alternatively, the lower CA values of /θ/ may reflect the lesser degree of contact for fricatives with blending of the constriction location for the cluster at the dental region.

Rate effects are speaker dependent. Speaker MM shows no significant effect of rate. For speaker AR differences in rate affect the contact indeces of C1 and C2 in the cluster showing more adaptation of C1 to C2 in fast than in slow speech. Word category effects show that C1 (and C2) in lexical words tend to have a higher constriction degree than in function words both in isolation and in clusters.

Catalan /nt/ clusters

Compared to intervocalic /n/, /n/ followed by dental /t/ shows a significantly higher CA and CC values for both subjects. The second consonant in the cluster shows a smaller and less consistent increase in both indeces. Such increase in anteriority and centrality in both segments in the cluster indicates an extension in the area of closure, that is, a larger spatial displacement of the tongue tip/blade for the two similar lingual gestures. Thus, coarticulation between the two same-tier gestures has the effect of reinforcing each individual gesture (similar reinforcement effects in clusters, resulting in a generally more advanced and higher maximum position, were found by [4, 5]). As a result of reinforcement, which affects C1 in a

larger degree than C2, the significant difference in CA indeces found at the two points of maximum constriction for /n/ and /t/ intervocalically disappears at the corresponding points in the cluster, which is articulated as a single gesture in the region of dental /t/.Thus, /nt/ trajectories in Catalan show shifting of the constriction location of /n/ to that of dental /t/.

No significant effects of rate and word category were found except for speaker MJ who exhibits significantly more similar central contacts for the two consonants in the cluster in functional words in fast speech.

Catalan /n(t)ʃ/ /t(t)ʃ/ clusters and English /ntʃ/, /dtʃ/ clusters

In Catalan /ʃ/and /tʃ/ are in free variation in word initial position. Speaker AF consistently produced [tʃ] intervocalically and in clusters whereas speaker MJ consistently produced [ʃ]. Thus, the CC index for [ʃ] for speaker MJ is much lower, indicating an open constriction, than that for AF's [tʃ]. Compared to intervocalic /n, t/, Catalan /n, t/ followed by an alveolopalatal obstruent show a significantly more retracted constriction location (lower CA), indicating that the articulatory configuration of C1 becomes more similar to that of C2. No significant differences in constriction location between cluster and intervocalic C2 were found, indicating that all the coarticulatory effects go in the anticipatory direction (these results are in line with those found for Italian [4]). The CC values of C1 and C2 in oral and /n/ clusters increase vis-à-vis intervocalic position, indicating that the constriction degree is higher for both segments in the clusters than in single consonants, thus confirming the effect of reinforcement suggested above.

For speaker MJ, the difference in CA between the two segments in isolation disappears in the oral and /n/ clusters. This suggests blending of the individual gestures constriction location. The cluster appears as a single gesture realized in a region in the vicinity of single /ʃ/. Whereas /n/ clusters show blending of constriction degree (no difference in CC between the two elements in the cluster) with /ʃ/ adapting to the constriction degree of single /n/, /t/ clusters exhibit a significant discontinuity in the central contacts for the two elements, indicating the transition from a stop to a fricative degree of constriction. For speaker AF there is no blending of the individual gestures constriction location or degree (significantly different CA and CC in the two segments in the cluster).

For the English speakers both consonants in the cluster show a higher CA index than in intervocalic position, indicating that the extent of contact is higher in the cluster. In all cases, the constriction location of C2 in the cluster becomes more similar to that of single C1 (carry over assimilation). The alveolar segment in the cluster shows a much wider contact in the central region (significantly higher CC) due to the effect of the following alveolopalatal (anticipatory assimilation). No changes in central contacts for the /tʃ/ are observed in the cluster. Thus, there is a mutual influence of both segments in the cluster: the first segment adapts to the wider central constriction of the second element whereas the /tʃ/ adapts to the constriction location of the alveolar.

For speaker MM the CA values of /d/+/tʃ/ differ significantly whereas those of /n/+/tʃ/ do not differ (showing values in the region of single /n/). The CC values for both elements in the cluster, however, differ significantly, suggesting two different gestures. This confirms the visual impression that the cluster is produced with two distinct but strongly coarticulated /n,d/ + /tʃ/ gestures. Speaker AR shows no blending of the individual gestures.

Overall, fast rate tends to exhibit lower CA and CC indices than slow rate both in clusters and in isolation, indicating a smaller spatial elevation of the articulator in fast speech. However, rate differences seem to be speaker dependent (English speaker MM shows no significant effect of rate on contact indices). All consonants (C1 and C2), and specifically oral alveolars, tend to show a wider extent of contact in lexical than in functional words both in clusters and in isolation. Comparison of the contact index values for both consonants in the cluster at slow and fast rate and in lexical and functional words shows no significant interaction effects. Thus, rate and word category do not affect the degree of blending in clusters.

It is necessary to be cautious when drawing conclusions about differences in blending processes between languages in view of the small number of speakers and the differences between speakers within each language group. However, it can be observed that in alveolar + alveolopalatal clusters English shows a mutual influence of both segments in the cluster and no cases

of articulatory shift whereas Catalan speakers tend to adapt the articulatory configuration of C1 to that of C2, the latter dominating the articulatory configuration of the cluster. This observation is in line with the stronger tendency in Catalan to weaken coda consonants, and to show a greater tendency to articulatory overlap than English and other languages [6].

CONCLUSIONS

The results indicate that featural phonology cannot fully account for the data obtained. Evidence of the alveolar gesture suggests that the alveolar segment is present in the input to speech production although its realization may be modified in connected speech. Thus, assimilatory processes cannot be modelled in terms of substitution of features. Articulatory Phonology predicts than when two same-tier gestures overlap in time, blending of the constriction locations of C1 and C2 result in one single gesture with an intermediate target [1]. Furthermore, blending of constriction location is claimed to occur only when the specifications for constriction degree in both segments in the cluster are the same [5]. In this view blending is interpreted as a reorganization of the articulatory movements at the gestural level, and not at the tract variable level. These predictions are not fully borne-out by the data obtained, which show no cases of intermediate targets for the blended gestures. In cases of articulatory shift the constriction location of C1 in the cluster is modified in the direction of that of C2. When two separate trajectories are found, there is mutual adaptation of the constriction location of C1 and C2. As regards the relationship between constriction location and constriction degree in blending, the predictions of AP are only partially borne out. Blending of constriction location occurs in Catalan /n/+/t/ clusters, which share constriction degree, but also possibly in English /n,d/+/θ/ clusters involving different specifications for constriction degree. In alveolar + alveolopalatal clusters, blending occurs for Catalan /n, t/ +/ʃ/, involving two different constriction degrees whereas it does not occur in clusters involving /tʃ/.

The rate, word category and speaker dependency in blending processes, along with the predominance of C2 constriction location, suggest that the neural commands for the alveolar segment are overlapped and modified by the conflicting commands for the upcoming segments, rather than reorganized and modified at a higher gestural level. The language-specific nature of blending processes and the higher occurrence of articulatory shift in Catalan than in English does not allow an explanation in terms of the organization of articulatory gestures in fast speech alone -- as claimed by gestural phonology -- but requires an explanation in terms of language dependent structural factors. The language-specific ranking of universal constraints, as proposed by Optimality Theory, can explain the observed cross-language differences.

The fact that nasal alveolars are more likely targets of place assimilation than oral alveolars can be explained in terms of trade-offs between articulatory effort and enhanced perceptibility: oral stops have stronger place cues than nasals and are more likely preserved in articulation than acoustically less salient segments.

ACKNOWLEDGEMENTS

This research was supported by a DGICYT grant PB93-859 to the Universitat Autònoma de Barcelona.

REFERENCES

[1] Browman, C.P. and L.M. Goldstein (1990), Tiers in Articulatory Phonology. In J. Kingston and M. E. Beckman (eds.), *Papers in Laboratory Phonology I.* Cambridge University Press: Cambridge, pp. 341-376.

[2] Nolan, F. 1992. The descriptive role of segments: Evidence from assimilation. In G.J. Docherty and D.R. Ladd (eds.), *Papers in Laboratory Phonology II.* C.U.P.: Cambridge, 261-280.

[3] Fontdevila, J., M.D. Pallarès and D. Recasens. 1994. The contact index method of EPG data reduction. *Journal of Phonetics*, 22, 141-154.

[4] Farnetani, E. and M.G. Busà. 1994. Italian clusters in continuous speech. *Proceedings of the 3rd International Conference on Spoken Language Processing*, Vol. 1, pp. 359-362.

[5] Romero, J. in press. Articulatory blending of lingual gestures. *Proceedings of the 1993 ACCOR Meeting*. Barcelona.

[6] Gibbon, F., W. Hardcastle and K. Nikolaidis. 1993. Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: a cross-linguistic investigation. *Language and Speech*, 36 (2, 3), 261-277.