LARYNGEAL AND ORAL GESTURES IN ENGLISH /P, T, K/

André M. Cooper

The University of Michigan, Ann Arbor, Michigan and Haskins Laboratories, New Haven, CT, USA

ABSTRACT

The present results support recent findings showing that VOT is shorter for /p/ than for the two lingual stops /t, k/ and that VOT for lingual stops are generally equivalent. Further, the results offer no support for a compensatory relationship between closure duration and VOT and show that the laryngeal devoicing gesture differs for stops produced at different places of articulation, thus ruling out several articulation-based explanations for place-related differences in VOT. Finally, the results suggest that the timing of glottal adduction relative to oral release most nearly accounts for observed differences in VOT.

1. Introduction

A number of researchers have found that VOT increases as the place of articulation of a stop progresses from the front to the back of the vocal tract [4, 6, 8]. One possible explanation for this finding is based on the assumption that the devoicing gesture (i.e., the opening and closing of the glottis for devoicing) is invariant while supralaryngeal gestures get progressively shorter the further back a stop is articulated [7, 8]. Other proposed explanations refer to automatic aerodynamic or mechanical consequences or to perceptual requirements associated with stops produced at different places of articulation [3, 5].

Results from a number of recent studies of both American [1] and British [2] English, however, have cast doubt on the conventional view of place-related differences in VOT and their explanations. These findings indicate that VOT for labial stops is shorter than for lingual stops, while VOT for /t, k/ tend not to differ from one another. Indeed, most earlier studies reporting place-related differences in VOT show smaller VOT differences between /t/ and /k/ than between /p/ and /t, k/, a difference that may not have been statistically significant [2]. Thus, explanations of VOT differences which crucially refer to a stop's place of articulation cannot account for the data from these recent findings.

The purpose of this study is threefold: (1) to examine place-related differences in closure duration and VOT in different word positions and under different stress conditions to test whether there is a compensatory relationship between the two; (2) to determine whether there is a single invariant devoicing gesture for all stops across different places of articulation; (3) to explore the role of oral-laryngeal timing with respect to VOT.

2. Methods

Two male speakers of English, ES and KM, spoke the nonsense words /pipip, titit, kikik/ with primary stress either on the initial or the final syllable in the carrier phrase "say __ again". Both acoustic and transillumination signals were collected simultaneously. Since the two speakers sometimes exhibited different articulatory patterns, separate statistical analyses were performed for each.

3. Results and Discussion

3.1. Acoustics

3.1.1. Closure duration

Separate ANOVAs for each word position indicate that stops produced at different places of articulation differ in closure duration in both initial and medial positions for both speakers (fig. 1). For ES, individual protected t-tests indicate that closure duration is significantly longer for /p/ than for /t, k/ but that there is no significant difference between closure durations for /t/ versus /k/. Similarly, for KM, closure duration is longest for /p/, and although mean closure duration is consistantly longer for /t/ than for /k/, the effect is only significant in medial stressed position.

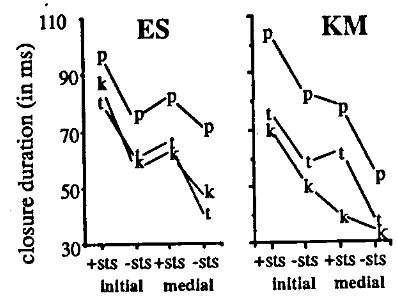

Fig. 1. Closure duration results for ES are presented on the left and results for KM on the right. The letter corresponding to the stop category is plotted in the graph. +sts = stressed and -sts = unstressed.

3.1.2. VOT

In general, the well-documented place-related VOT pattern for English voiceless stops is exhibited for each stress category in both word-initial and word-medial positions for both speakers (fig. 2). Separate one-way ANOVAs confirm that stops produced at different places of articulation significantly differ in VOT for both speakers. For ES, VOT is significantly longer for /t/ than for /p/, and significantly longer for /k/ than for /t/. The only exception is that VOTs for medial unstressed /t, k/ are not significantly different from one another, although they manifest the same rank order as the other groups.

Results for KM differ somewhat from those for ES. Like ES, VOT for /p/ is significantly shorter than that for /t, k/ for each level of stress within each word position. Unlike ES, however, there is no significant difference in VOT for /t, k/ even though there is a tendency for mean VOT for /k/ to be slightly longer than for /t/.

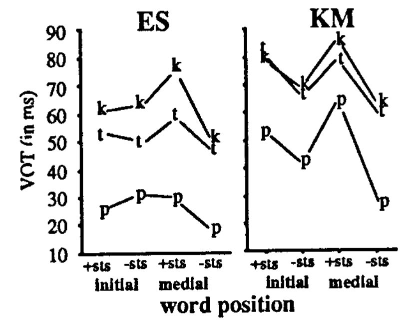

Fig 2. VOT results presented as in fig. 1.

The acoustic data offer no support for VOT as a function of closure duration across an invariant devoicing gesture. Since closure duration for /t, k/ is equivalent for both speakers, one might expect VOT to be equivalent for both speakers under the invariant devoicing gesture proposal. However, ES shows VOT differences between /t, k/, while KM does not. It is, nevertheless, still possible that there is an invariant devoicing gesture for all stops. The devoicing gesture may simply be shifted in time with respect to oral closure for /t, k/. We examine these possibilities in the following sections.

3.2. Transillumination

3.2.1 Devoicing Gesture Duration (DGD)

ANOVAs show that there is also a significant effect of place of articulation on DGD for both speakers (fig. 3). However, a clear pattern of results does not emerge unless DGDs for lingual stops are considered as a group separate from labial stops. For ES, DGDs for /t/ are significantly longer than those for /k/ regardless of stress or word position. For KM, DGDs for /t, k/ only differ significantly from one another in medial stressed syllables, although mean DGD for /t/ is longer than for /k/ for each condition.

DGDs for labial stops in general appear to differ from DGDs for lingual stops. In word-initial position, differences between stressed and unstressed DGDs for labial stops are small and nonsignificant, but are comparatively large and significant for the lingual stops. Within stress categories mean DGD for initial stressed /p/ is somewhat similar to that for medial stressed /p/, while DGDs tend to be longer in initial than in medial position regardless of stress for lingual stops. The aforementioned patterns are especially evident for KM.

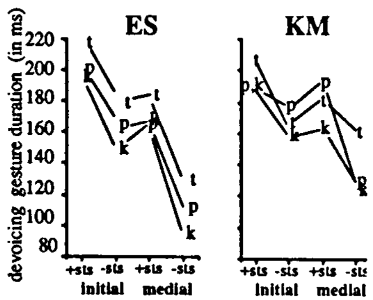

Fig. 3. DGD results presented as in fig. 1.

3.2.2. Peak Glottal Magnitude (PGM)

PGM (i.e., the greatest distance between the vocal folds during the devoicing gesture) results are similar to DGD results for both speakers in that there is a significant effect of place of articulation on PGM for both speakers and in that labial stops behave somewhat differently than lingual stops (fig. 4). For ES, PGM is always significantly greater for /t/ than for /k/. For KM, PGM is only significantly greater for /t/ than for /k/ in medial unstressed syllables; otherwise, they are equivalent in magnitude.

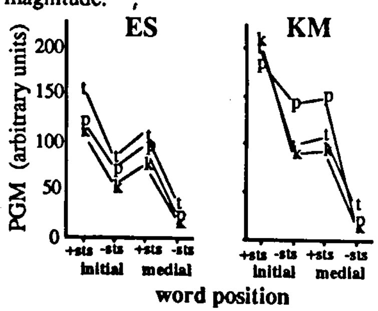

Fig. 4. PGM results presented as in fig. 1.

The present results show that there is no single invariant devoicing gesture for stops across place of articulation. Indeed, it seems that the devoicing gesture may be influenced by both the supralaryngeal constriction location as well as the primary supralaryngeal articulator. Specifically, it appears that devoicing gestures are generally sensitive to whether the supralaryngeal constriction is produced with the lips or the tongue since the data suggest that stress and word position have different effects on labial versus lingual stops.

It seems possible that VOT differences among the lingual stops arise from differences in the duration and magnitude of the devoicing gesture since ES shows consistant differences for /t/ versus /k/ for both DGD, PGM and VOT, and since KM shows no difference in DGD, PGM or VOT for /t/ versus /k/. However, such a relationship seems especially doubtful since one might expect larger devoicing gestures to give rise to longer VOTs, whereas just the opposite result obtains for ES. In any case, these findings suggest that the timing of oral and laryngeal gestures must play a crucial role in VOT since variations in neither oral nor laryngeal gestures alone can account for the observed VOT patterns.

3.3. Interarticulator Timing

The coordination of laryngeal and supralaryngeal gestures has been intimately linked with VOT [4]. Here we consider the coordination between two pairs of articulatory events associated with the beginning and the end of voiceless stop-related gestures (namely, the interval from oral closure to the onset of glottal opening and the interval from oral release to the onset of glottal adduction) in order to determine whether the relationship between either of these events covaries with VOT.

3.3.1. Closure to Onset of Glottal Opening (C-OGO)

There is a significant effect of place of articulation on C-OGO for both speakers (fig. 5). For ES, OGO always occurs significantly later for /k/ than for /p, t/, but only occurs significantly earlier for /t/ than for /p/ in initial stressed and medial unstressed syllables. For KM, C-OGOs for the lingual stops are not significantly different from one another in any word position or for any stress category. There is no clear pattern for labial stops.

The C-OGO results closely mirror the patterns found for DGD and PGM suggesting that the larger the devoicing gesture, the earlier it begins relative to closure. When considering the labial stops in conjunction with the lingual stops, it becomes even more clear that the onset of the devoicing gesture does not simply shift in time relative to oral gestures to achieve a specific VOT. Rather, C-OGO is related to the size of the devoicing gesture. In fact, even the mean C-OGO data closely follow the same rank order as for DGD and PGM.

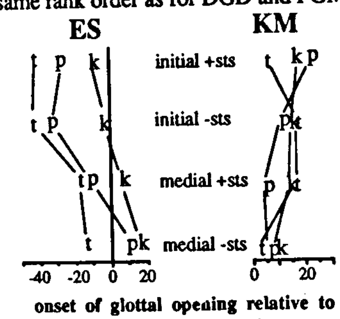

Fig. 5. Results for the interval from oral closure (at 0ms) to the OGO (the points plotted) presented as in fig. 1.

3.3.2. Release to Onset of Glottal Adduction (R-OGA)

There is a significant effect of place of articulation on R-OGA for both speakers (fig. 6). For ES, mean OGA relative to release occurs earliest for /p/, latest for /k/, and intermediate for /t/; this effect is significant except in medial unstressed position where R-OGAs for /t, k/ are not significantly different from one another.

Like ES, the OGA for KM always occurs significantly earlier for /p/ than for /t, k/ in both word positions and for both stress categories. Unlike ES, however, OGAs for /t/ only occur significantly earlier than for /k/ in medial stressed position; otherwise, R-OGAs for /t, k/ do not differ significantly.

R-OGA results are practically identical to the corresponding VOT results. Specifically, the earlier the OGA, the shorter the VOT for all stop categories (cf. fig. 2). Thus it appears that R-OGA is responsible for differences in VOT, and not variations in closure duration plus an invariant devoicing gesture, or differences in the size of the devoicing gesture.

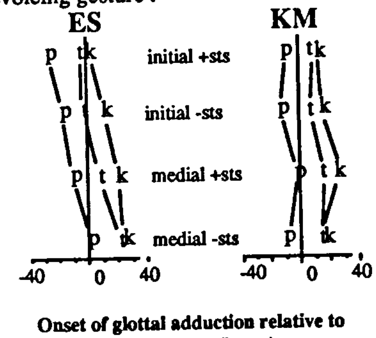

Fig. 6. Results for the interval from oral release (at 0ms) to the OGA (the points plotted) presented as in fig. 1.

Finally, it is important to note that the OGA is an active gesture rather than a passive aeromechanical consequence of oral release. Since both oral release and the OGA are controlled by muscular forces, it follows that oral release and the OGA are actively timed relative to one another. Thus it is possible that VOT differences are simply a byproduct of, rather than the motivation for, the timing of OGA relative to oral release.

REFERENCES

[1] Crystal and House A., (1988) Segmental durations in connected speech signals: Current results. *JASA*, 83(4): 1553-1573.

[2] Docherty, G. J. (1989) *An experimental phonetic study of the timing of voicing in English obstruents.* University of Edinburgh unpublished Ph.D. Dissertation.

[3] Fischer-Jorgenson, E. (1980) Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *Annual Report Institute of Phonetics University of Copenhagen*, 14:207-261.

[4] Lisker, L. & Abramson, A. (1967) Some effects of context on voice onset time in English stops. *Language and Speech*, 10:1-28.

[5] Stevens, K., N., & Klatt, D., H. (1974) The role of formant transitions in the voice-voiceless distinction for stops. *JASA*, 55: 653-659.

[6] Weismer, G. (1979) Sensitivity of voice onset measures to certain segmental features in speech production. *Journal of Phonetics*, 7: 194-204.

[7] Weismer, G. (1980) Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations. *Journal of Phonetics*, 8: 427-438.

[8] Zue, V. (1976) *Acoustic characteristics of stop consonants.* Indiana University Linguistics Club.

Work supported by NIH Grant DC 00121 to Haskins Laboratories and the University of Michigan.