Coarticulatory organization for lip rounding in Turkish and English

Suzanne E. Boyce
Speech Communication Group, Research Laboratory of Electronics, Room 36-545, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 2 October 1989; accepted for publication 16 July 1990)

A number of studies, involving English, Swedish, French, and Spanish, have shown that, for sequences of rounded vowels separated by nonlabial consonants, both EMG activity and lip protrusion diminish during the intervocalic consonant interval, producing a "trough" pattern. A two-part study was conducted to (a) compare patterns of protrusion movement (upper and lower lip) and EMG activity (orbicularis oris) for speakers of English and Turkish, a language where phonological rules constrain vowels within a word to agree in rounding and (b) determine which of two current models of coarticulation, the "look-ahead" and "coproduction" models, best explained the data. Results showed Turkish speakers producing "plateau" patterns of movement rather than troughs, and unimodal rather than bimodal patterns of EMG activity. In the second part of the study, one prediction of the coproduction model, that articulatory gestures have stable profiles across contexts, was tested by adding and subtracting movement data signals to synthesize naturally occurring patterns. Results suggest English and Turkish may have different modes of coarticulatory organization.

PACS numbers: 43.70. – h, 43.70.Bk, 43.70.Kv

INTRODUCTION

Theories of coarticulation in speech have taken as an axiom the notion that, by coarticulating segments, a speaker is aiding the efficiency of his or her production (Liberman and Studdert-Kennedy, 1977). Although discussion of the forces affecting coarticulation has tended to concentrate on articulatory and/or perceptual pressures operating within particular sequences of segments (Martin and Bunnell, 1982; Ohala, 1981; Beckman and Shoji, 1984; Recasens, 1985), there is a growing body of cross-linguistic work exploring the influence of language-particular phonological structure on coarticulation (Öhman, 1966; Lubker and Gay, 1982; Magen, 1984; Perkell, 1986; Manuel, 1990; Keating, 1988, among others). These studies have generally been concerned with the interaction of coarticulation and segment inventory, or coarticulation and the properties of some particular segment; the question of how coarticulation interacts with phonological rules has been relatively neglected (but cf. Cohn, 1989). Phonological rules, for instance, determine the typical structure of words in a language; we might speculate that languages with different constraints on the possible sequencing of segments pose different challenges to the articulatory planner, and thus that speakers of these languages would vary in the way they implement coarticulation. To take an example, speakers of Turkish, a vowel harmony language with strict rules for the possible sequencing of rounded and unrounded vowels, might feel more pressure to employ rounding coarticulation than English speakers, whose language freely combines rounded and unrounded vowels.

Rounding coarticulation for sequences of rounded and unrounded vowels in English has been extensively studied. A number of studies have shown that, for strings of two rounded vowels separated by nonlabial consonants, e.g., /utu/ or /ustu/, both EMG and lip protrusion movement traces show double peaks coincident with the two rounded vowels plus an intervening dip or "trough" in the signal (McAllister, 1978; Gay, 1978; Engstrand, 1981; Perkell, 1986). (A schematized version of this pattern, representing EMG from the orbicularis oris muscle for the utterance /utu/, is illustrated in Fig. 1.) This result has been the focus of a good deal of controversy in recent years, primarily because different theories of coarticulation tend to treat it in different ways. For instance, much previous work on the control mechanisms underlying anticipatory coarticulation has centered on the predictions of one class of models, the "look-ahead" or "feature-spreading" models (Henke, 1966; Daniloff and Moll, 1968; Benguerel and Cowan, 1974). Generally, these models view coarticulation as the migration of features from surrounding phones. In the most explicit form of this type of model, Henke's (1966) computer implementation of articulatory synthesis, the articulatory planning component scans upcoming segments and implements features as soon as preceding articulatorily compatible segments make it possible to do so.[1] In the case of /utu/ or /ustu/, the nonlabial consonants separating the vowels are made with the tongue and presumably do not conflict with simultaneous lip

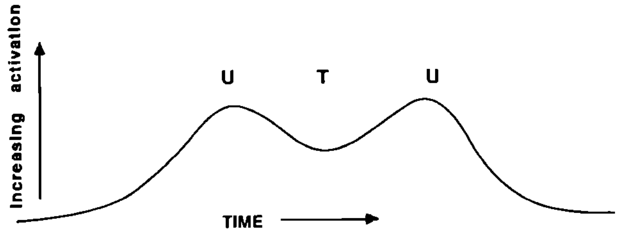

FIG. 1. Schematized version of "trough" pattern, representing EMG from the orbicularis oris muscle for the utterance /utu/.

rounding. Thus the fact that troughs occur is a problem for the look-ahead model, because the model would normally predict that the rounding feature for the second vowel would spread onto the preceding consonant, producing a continuous plateau of rounding from vowel to vowel.

Explanations of the trough results have varied widely. Following a more general suggestion of Kozhevnikov and Chistovich (1965), Gay (1978) proposed that the trough represents the resetting of a "syllable-sized articulatory unit," and that coarticulation is allowed to take place only within that unit. Some evidence against this explanation was provided by Harris and Bell-Berti (1984), who found no sign of a trough in sequences such as /uhu/ and /u?u/. Addressing himself to the sequences typically used in these experiments, Engstrand (1981) took issue with the assumption that alveolar consonants are compatible with full lip rounding. He argued instead that rounding as found in the vowel /u/ may interfere with optimal acoustic/aerodynamic conditions for these consonants, and that the trough may result from lip movement toward a less-rounded configuration. That such acoustic/aerodynamic constraints may not hold for all subjects was shown by Gelfer *et al.* (1989), who reported data from a subject with lip protrusion and orbicularis oris muscle activity for /t/. Perkell (1986) hypothesized that a dipthongal pattern of movement for the /u/ vowels (i.e., from a less to a more extreme lip position), in addition to acoustic and other constraints might reduce the extent of rounding in the vicinity of the intervocalic segment(s). In his own work, however, he found little evidence for diphthongal behavior in those subjects who showed troughs.

Each of these proposals, it should be noted, can be seen as a modification to the look-ahead class of models, in which features of one segment spread to another segment if context conditions allow it. Alternatively, a class of models known as "coproduction," "frame," or "time-locking" models (Fowler, 1980; Bell-Berti and Harris, 1981) assumes that coarticulation results from temporal overlap between independent articulatory gestures belonging to neighboring segments. In these models, lip movement for the utterance /utu/ involves two overlapping rounding gestures (assuming /t/ is not independently rounded). Thus the presence of a trough is controlled by the degree of overlap between gesture peaks. If the peaks overlap one another, no trough will be discernible, but if the peaks are temporally separated from one another, the model predicts the occurrence of a trough. Another provision of these models is that gestures associated with a particular segment should show a stable profile across different segmental contexts. [It is acknowledged that characteristic gesture profiles may be affected by stress and possibly other prosodic contexts (Tuller *et al.*, 1982).] Thus the temporal extent of coarticulation is predicted by the temporal extent of the gesture. Attempts to test the latter provision of this model by measuring the lag times between the acoustic onset of rounded vowels and related articulatory activity have had varied and sometimes conflicting results, with studies by Bell-Berti and Harris (1979, 1982) and Engstrand (1981) supporting the coproduction prediction of stable lag times, and studies by Lubker (1981), Lubker and Gay (1982), Sussman and Westbury (1981), and Perkell (1986) indicating more variable behavior supportive of the look-ahead view. Thus, although it is unclear how the look-ahead class of models can account for the trough pattern, both types of models remain viable options for explaining coarticulation.

Regardless of which interpretation of the trough pattern is correct, however, the pattern itself has been found in each of the languages so far surveyed, appearing in English (Bell-Berti and Harris, 1974; Gay, 1978; Perkell, 1986), Swedish (McAllister, 1978; Engstrand, 1981), Spanish and French (Perkell, 1986). It is notable that these languages, while differing in such variables as syllable structure, the tendency to diphthongize vowels, and the presence of a phonological contrast between rounded and unrounded vowels, are alike in their tolerance for mixed sequences of rounded and unrounded vowels. It seemed plausible, at least, that the finding of troughs in lip-rounding activity for these languages might be related to this tolerance, and that a language like Turkish, in which words with mixed rounded and unrounded vowels are the exception, might show lip-rounding patterns other than the trough pattern. In particular, it seemed that Turkish might provide particularly favorable conditions for anticipatory coarticulation of the kind predicted by the look-ahead class of models. In brief, the hypothesis was that Turkish speakers would exhibit plateau patterns of activity for lip rounding. The experiment described below was part of a larger study designed to test this hypothesis (Boyce, 1988). A second aim of the experiment was to test the degree to which the coproduction models' explicit prediction of stable, independent gestures could be used to predict both English and Turkish movement patterns.

I. EXPERIMENT

A. Method

1. Stimuli

Four speakers of American English and four speakers of Standard Turkish produced similarly structured nonsense words designed to show the presence or absence of troughs in lip rounding. Corpus words for this purpose consisted of the series /kuktluk/, /kuktuk/, /kukuk/, /kutuk/, /kuluk/. Because arguments concerning the trough pattern often hinge on questions concerning the production of the intervocalic consonants in an unrounded environment (Benguerel and Cowan, 1974; Gelfer *et al.*, 1989), the words /kiktlik/, /kiktik/, /kikik/, /kitik/, /kilik/ were included as controls. Additionally, words with rounded vowels followed by unrounded vowels /kuktlik/, /kuktik/, /kukik/, /kutik/, and /kulik/, and words with unrounded vowels followed by rounded vowels /kiktluk/, /kiktuk/, /kikuk/, /kituk/, and /kiluk/ were included to provide data on single protrusion movements. In the remainder of the paper, words with vowel sequences U-U, I-I, etc., will be referred to as U-U, I-I, U-I, and I-U words. The words with intervocalic KTL, which had the longest vowel-to-vowel intervals, were included to provide the clearest test case for the presence of a trough pattern. Words with shorter intervocalic consonant intervals were included to provide control information on the lip activity patterns for different consonants. The carrier phrase for Turkish speakers was "Bir daha__deyiniz" (pronounced as phonetically spelled and meaning 'Say__once again'). The English carrier phrase was "It's a__again."

Although Turkish has final stress, the degree of difference between stressed and unstressed syllables is much less than in English (Boyce, 1978). Therefore, English speakers were encouraged to use equal stress on both syllables of the disyllabic nonsense words, and if equal stress felt unnatural, to place stress on the final rather than the initial syllable. Turkish subjects were given no instructions about stress. All subjects were instructed to speak at a comfortable rate, in a conversational manner.

Additional facts about Turkish that impinge on the arguments made in this paper have been summarized in the Appendix. For the present, it is sufficient to note that rounding in Turkish operates according to a vowel harmony rule which, in essence, causes sequences of high vowels to acquire the rounding specification of the preceding leftmost vowel. (With minor exceptions, consonants do not participate in this process.) The effect is to produce long strings of rounded or unrounded vowels whose rounding is predictable given the first vowel in the sequence. While vowel harmony is a productive rule for the vast bulk of the lexicon, there are numerous exceptions, mainly from Arabic and Persian borrowings. Real word counterparts exist for each of the vowel sequences in the experimental corpus, although U-U and I-I words conform to vowel harmony, while I-U and U-I words do not.

2. Subjects

English-speaking subjects included one male (AE) and three females (MB, AF, and NM), each of whom spoke a variety of General American with no marked regional or dialectal accent. The Turkish speakers included one female (IB) and three males (AT, EG, and CK). All spoke similar varieties of Standard Turkish.

For Turkish subject EG, utterances were randomized and the randomized list was repeated 15 times. Utterances in later subject runs were blocked, so that utterances were repeated in groups of five tokens (three for MB), utterances with the same vowel combinations were grouped together, and the same order of consonant combinations was repeated for each vowel combination. The order of vowel and consonant combinations was different for each subject, except that one Turkish speaker (IB) and one English speaker (AF) had the same order of presentation.

3. Instrumentation

Movement data from the nose, upper and lower lip, and jaw were obtained by means of an opto-electrical tracking system, similar to the commonly used Selcom SELSPOT system. The system consists of infrared light emitting diodes (LEDs) attached to the structure of interest. LED position is sensed by a photo-diode within a camera positioned to capture the range of LED movements in its focal plane. The output of this diode is translated by associated electronics into pairs of X and Y coordinate potentials for each LED, each with a maximum frequency response of 500 Hz. Calibration is achieved by moving a diode through a known distance in the focal plane.

LEDs were attached to the subject's nose, upper lip, lower lip, and jaw with double-sided tape. The nose LED was placed on the bridge of the nose, slightly to the left side, at a point determined to show the least speech-related wrinkling, waggling, etc. LEDs were placed just below the vermilion border of the upper lip and just above the vermilion border of the lower lip, in a plane with the nose LED, at a point judged to show the axis of anterior–posterior movement for each articulator. The movement of the subject's skin between the lower lip and chin was observed during production of rounded vowels, and the jaw LED positioned to best reflect anterior–posterior movements of the mandible rather than skin and muscle. Generally, this was at the point of the chin or under it, in a plane with the higher LEDs.

The LED-tracking camera was positioned at 90 deg to the left of the subject's sagittal midline, at a camera-to-subject distance (21 in.) that provided a 10×10-in. field of view. When centered approximately on the upper/lower lip junction, during maintenance of a position appropriate for bilabial closure, this field is large enough to capture the full range of anterior–posterior LED movement, as well as allowing for some degree of head movement.

A video camera was positioned 90 deg to subject midline on the subject's right, and focused as narrowly as possible, while continuing to keep all 4 LEDs within the field of view. Four subjects were videotaped throughout the experiment: English subjects AF and NM, and Turkish subjects AT, and IB. An additional videotape of English subject MB producing the words /kitklik/, kuktluk/, /kiktluk/, and /kuktlik/ was obtained in a separate session.

A simultaneous audio recording of the subject's speech during the experiment was made on a Sennheiser "shotgun" microphone.

The EMG recordings were made with adhesive surface silver–silver chloride electrodes. These were placed just below and above the vermilion border of upper and lower lips, laterally to the midline. According to Blair and Smith (1986), an electrode at this location is likely to pick up relatively more activity from orbicularis oris, and less of nearby muscles, than at other locations along the lip edge. Pickup from the desired muscle, orbicularis oris inferior (OOI) and orbicularis oris superior (OOS), was checked by having the subject produce repeated /u/ or /i/ vowels several times in succession; if a strong signal was evidenced for /u/ and little or no signal for /i/, the EMG electrode was assumed to be well placed.

4. Procedure

The EMG and movement signals, together with audio and clock signals, were recorded onto a 14-channel FM tape recorder (EMI series 7000). The EMG signals were rectified, integrated over a 5-ms window, and sampled at 200 Hz. Movement signals were also sampled at 200 Hz. The audio channel was filtered at 5000 Hz and sampled at 10 000 Hz. By means of the simultaneous clock signals, data from all channels were synchronized to within 2.5 ms.

The signal from the nose LED was numerically subtracted from respective lip and jaw signals to control for changes in baseline due to head movement. Differences in baseline between early and late portions of the experiment remained for some speakers, presumably due to vertical rotational movement of the head, in which the lips and nose, or jaw and nose, moved by different amounts in space. The speakers most affected were English speaker AE, for whom the total horizontal lower lip baseline change was approximately 3 mm, and Turkish speakers IB, EG, and CK, for whom the total changes were approximately 3.5, 6.5, and 6 mm, respectively. In each case, baseline change reflected movement in the posterior direction. Baseline change for other speakers was within 1 mm of movement. Rotational movement of this type, in which the chin sank gradually toward the base of the neck, was confirmed in the videotape for subject IB (other videotaped subjects showed little baseline change). These baseline changes did not appear to affect the data in any substantial way.[2]

Because of recording or calibration problems, the upper lip movement signal and both EMG signals for Turkish subject AT, the EMG OOS signal for Turkish subject EG, and the EMG OOI signal for Turkish subject CK were eliminated from the study. Except for AT, therefore, the full complement of movement signals, and at least one EMG signal, was available for each subject. Recording or calibration problems also caused some of the 15 tokens planned for words in the experimental corpus to be discarded. The upper lip signal level for English subject AE deteriorated after the first block of utterances. Thus only the first five tokens for this signal are reported.

Two acoustic reference points, or lineups, were identified for each token. The first, the V_1 offset, was defined as the point where the formant structure disappeared from the waveform at the onset of closure for /k/ or /t/ or the point of sudden amplitude change marking the change between the vowel and the voiced approximant /l/. The second, the V_2 onset, was defined as the release of the consonant occlusion for /k/ and /t/, or the point of amplitude change for /l/. Consonant interval duration measurements consisted of the time between these two points. The audio waveform, movement, and EMG signals for each repetition of an utterance in the experimental corpus were extracted into a separate computer file. Each file contained a 2000-ms slice of speech with constant dimensions before and after the V_1 offset point.

The main body of movement data reported here comes from the anterior–posterior upper and lower lip signals. These signals are referred to in the text as upper lip X (ULX) and lower lip X (LLX). Both signals reflect lip protrusion, which is generally acknowledged to be the most reliable single index of lip rounding. However, because rounding may also involve vertical motion of the lips, to narrow the lip aperture, and because vertical movement and protrusion of the lower lip may be affected by movements of the jaw, anterior–posterior jaw (JX), inferior–superior jaw (JY), and lip signals (ULY, LLY) were also consulted.

As a rule, token-to-token variability was minimal in both movement and EMG signals. [Full sets of token movement traces for all subjects' /kuktluk/, /kiktlik/, /kuktlik/ and /kiktluk/ words are displayed in Boyce (1988).] Accordingly, much of the presentation in this paper is based on movement and EMG traces produced by ensemble averaging. Those cases where token-to-token variability was greater than implied by the averaged signal are mentioned in the text. Signals were ensemble averaged using the acoustic V_1 offset as a lineup point.

B. Results

Movement and EMG signals were examined separately for Turkish and English subjects, with a view to determining characteristic movement and muscle activity patterns for U-U words. The major issues were whether (a) the trough pattern reported in the literature for English speakers would be replicated and (b) Turkish and English speakers would show the same pattern. The Turkish cases will be discussed first, followed by the relatively more complex English data.

1. Turkish speakers

The right side of Fig. 2 shows the averaged ULX, LLX, and EMG traces for /kuktluk/ and /kiktlik/ as produced by the four Turkish speakers AT, IB, EG, and CK. Overall, the /kuktluk/ movement traces for these subjects tended to resemble a plateau, with the protrusion traces being flat or slightly falling over the course of the word. Exceptions to this pattern are the occurrence of a peak, or "bump" during the consonant interval in the LLX traces for subjects EG and CK, and the slight trough located at the beginning of V_2 in the ULX signal for subject EG.

To verify the impressionistic analysis, a statistical procedure was designed to test the degree to which lower lip movement patterns for the Turkish speakers' U-U words were flat or convex rather than troughlike.[3] For each subject's averaged U-U word trace, the number of sample points from the lineup point to the left and right edges of the plateau pattern was identified (by visual inspection), and the same number of sample points to the left and right of the lineup point was extracted from each token trace. Baseline differences between tokens were normalized by (a) calculating the difference between the mean of token sample points and the grand mean of sample points for all tokens of a particular word and (b) adding or subtracting the difference for each sample point in a token.

These data were subjected, as a function of displacement versus time, to a stepwise regression analysis testing for fit to linear and quadratic regression equations. The intent was to determine to what extent the data between these points showed quadratic (i.e., parabolic) trend, and, if so, whether the orientation of the parabola

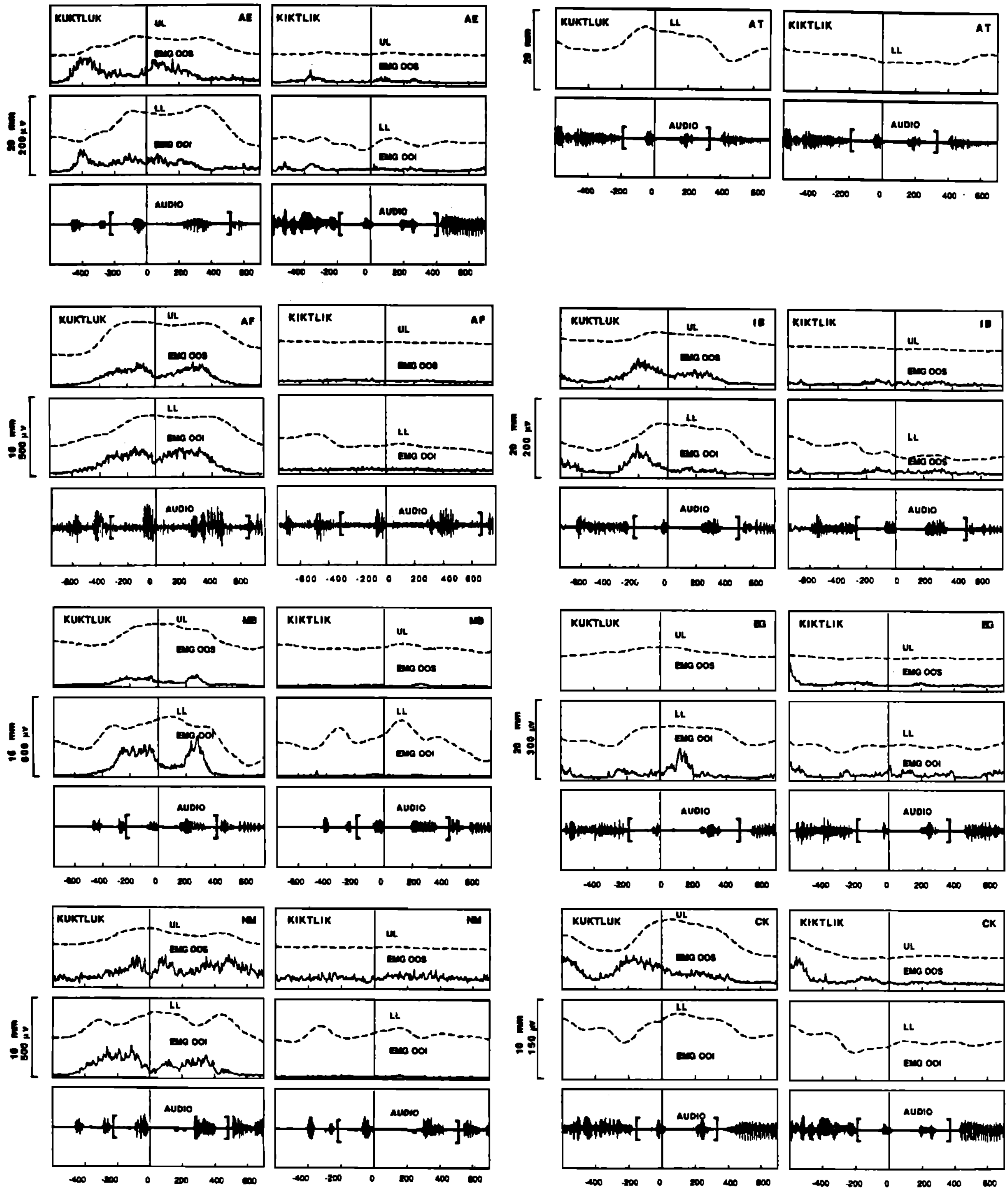

FIG. 2. Averaged upper and lower lip movement (dashed line) and EMG (solid line) traces, plus single token acoustic waveform traces, for /kuktluk/ and /kiktlik/ as produced by Turkish speakers AT, IB, EG, and CK, and English speakers AE, AF, MB, and NM. The number of tokens for averaging ranged from 10–15, except for AE's upper lip trace, which was averaged from five tokens. Upward deflection represents anterior movement. The vertical line indicates the lineup point for ensemble averaging, which was at the acoustic offset of V_1. Square brackets in the lower panel indicate approximate acoustic boundaries for these words. The horizontal scale is time in ms.

was concave, i.e., troughlike, or convex. (Only data from /kuktluk/, which was representative of that for all u-u words, is reported here.) Of the four Turkish speakers, all but AT consistently showed negative quadratic regression coefficients, indicating convex rather than concave shape. AT's positive quadratic coefficients, however, were accompanied by small quadratic R^2 terms and large linear R^2 terms (for /kuktluk/: 0.0001 and 0.79, respectively), indicating that the concave quadratic trend accounted for a tiny proportion of the variance and that AT's pattern was overwhelmingly linear. For subjects EG and CK, on the other hand, the quadratic term accounted for a considerable proportion of the variance (subject EG: /kuktluk/ quadratic $R^2 = 0.28$, linear $R^2 = 0.21$; subject CK: /kuktluk/ quadratic $R^2 = 40$, linear $R^2 = 0.02$), indicating a substantial degree of convex shape probably attributable to the "bump" visible on their movement traces. Like AT, subject IB showed a large linear component combined with a very small degree of quadratic trend (/kuktluk/ linear $R^2 = 0.61$, quadratic $R^2 = 0.004$).[4]

Of the three Turkish speakers with EMG data (OOS and OOI for IB, OOI for EG and OOS for CK), there was no conspicuous diminution of EMG activity during the consonant interval. The general pattern was unimodal. For IB and CK, there was an early peak on V_1 followed by a long, sustained offset and some indication of increased activity during V_2. For subject EG, the EMG peak was located close to V_2 in /kuktluk/ and to V_1 in all other U-Ụ words. (Movement patterns, however, were similarly plateaulike over EG's different U-U words.[5])

As noted in the Introduction, it is well known that some English speakers may protrude and/or narrow their lips for nonlabial consonants, most notably /t/ (Gelfer *et al.*, 1989) and /l/ (Leidner, 1973; Brown, 1981). Looking at the /kiktlik/ traces, it appears that Turkish subjects AT, IB, and EG do not produce significant independent protrusion during the intervocalic consonants. Although small fluctuations in LLX signals may indicate some degree of active lower lip protrusion, these may also be due to lip relaxation from a retracted position during the flanking /i/ vowels. The strongest degree of movement during the /kiktlik/ consonant interval is seen for subject CK. It is hard to tell if this reflects active movement rather than passive relaxation, however, as CK retracts lip and jaw heavily for the sustained /ɑ/ of "dɑhɑ" (pronounced [dɑɑ]) in the carrier phrase (Boyce, 1988). EMG traces for all subjects during I-I words were flat.

2. English speakers

English subjects in this study could be divided into two groups based on the appearance of their horizontal lower lip signals for U-U and I-I words. (Upper lip signals were less clearly differentiated.) One pattern can be seen in the upper two sets of traces on the left side of Fig. 2, showing the averaged ULX, LLX, and EMG traces for /kuktluk/ and /kiktlik/ as produced by English subjects AE and AF. For both of these speakers, /kuktluk/ movement traces showed double-peaked "trough" patterns. A similar double-peaked pattern can be seen for subject AF's EMG OOI and OOS. For subject AE, the EMG OOS trace is clearly double peaked. The EMG OOI trace also shows two peaks, but with an additional peak between.[6]

The extent to which lower lip U-U word movement patterns for subjects AF and AE were troughlike as opposed to flat was examined statistically as described above for the Turkish speakers' data, in a stepwise regression analysis testing for the degree of linear and quadratic trend between the two movement peaks. For both subjects, the regression showed a positive quadratic coefficient, indicating a concave, or troughlike, shape, and quadratic R^2 terms represented a considerable proportion of the variance (for /kuktluk/, subject AE: quadratic $R^2 = 0.55$, linear $R^2 = 0.17$; subject AF: quadratic $R^2 = 0.18$, linear $R^2 = 0.14$).[7]

Looking at the averaged ULX, RLLX, and EMG traces for the word /kiktlik/, it appears that for subject AF there is little or no movement for either lip during the intervocalic consonant interval, and little or no activity in the EMG signal. Some fluctuation in the movement signal for LLX is present for subject AE. Comparison with the relatively neutral position of the lips during the /ɛ/ vowel from "again" at the end of the carrier phrase (between 400 and 600 ms after the V_1 offset point) suggests that this may be due to lip retraction during the flanking vowels, with relaxation of the lips during the consonant interval. Alternatively, some small active forward movement of the lower lip and/or jaw may be involved.

The lower sets of traces on the left-hand side of Fig. 2 show averaged ULX, LLX, and EMG traces for English subjects MB and NM. Looking at the /kuktluk/ words, we see that, for both MB and NM, the lower lip trace shows three peaks of movement. The first peak is located during the "It's" of the carrier phrase (at approximately 350 ms before the V_1 offset point) and probably indicates protrusion associated with /s/. The central, and largest, peak is located during the intervocalic consonant interval (after the vertical line). The third peak is located during the second vowel. At the same time, the EMG patterns for MB's and NM's EMG OOI traces show trough patterns like those of English subjects AE and AF (NM's EMG OOS trace, like AE's OOI trace, shows an additional peak after V_1 offset). The upper lip pattern for subject NM also resembles those for subjects AF and AE. Subject MB's upper lip movement pattern contains two peaks, which correspond roughly in time to the central and final peaks of the lower lip trace. Again, regression analyses as described above for Turkish subjects and English subjects AE and AF were performed on data extracted from between the left- and rightmost peaks. These showed negative regression coefficients for quadratic trend and high quadratic R^2 values, indicating convex shape (for /kuktluk/, MB: quadratic $R^2 = 0.60$, linear $R^2 = 0.01$; NM: quadratic $R^2 = 0.14$, linear $R^2 = 0.001$).

There is less apparent consistency between upper lip, lower lip, and EMG traces for these subjects than for subjects AE and AF. Looking at their /kiktlik/ traces, however, we see that both subjects MB and NM show protrusion in the lower lip signal during the consonant interval. There is also some protrusion in MB's upper lip /kiktlik/ trace. The latter is likely to reflect active for-

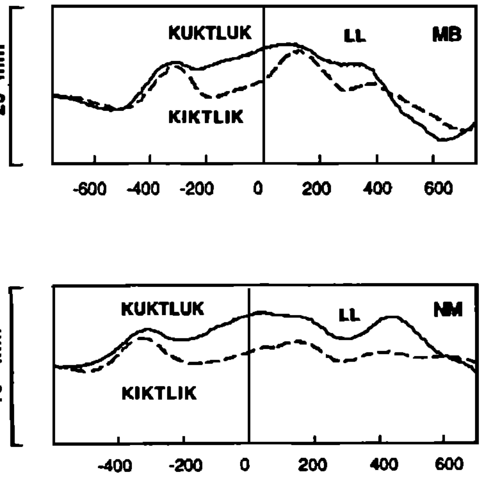

FIG. 3. Superimposed /kuktluk/ and /kiktlik/ averaged lower lip protrusion traces as produced by English subjects MB and NM. The vertical line is V_1 offset. Vertical and horizontal scales are as in Fig. 2.

ward movement since there is no sign of upper lip retraction on the flanking /i/ vowels.

In Fig. 3, we see the lower lip signal for /kiktlik/ overlaid with that for /kuktluk/ for English subjects MB and NM. (Baseline differences between averaged traces have been adjusted when necessary, so as to visually align carrier phrase portions of each trace.) From this, it is clear that the timing of the consonant interval protrusion peak in /kiktlik/ is very similar to the timing of the central peak in these subjects' /kuktluk/ traces. This is most striking for MB, whose protrusion peak in /kiktlik/ was also similar in amplitude to that of /kuktluk/. For NM, the central peak in /kuktluk/ was slightly bimodal, and the protrusion peak in /kiktlik/ is more nearly matched in timing to the second inflection. [For both subjects, a similar congruence of peaks can be seen when lower lip /kiktluk/ traces (shown in Fig. 5) are overlaid with /kiktlik/ traces.]

These observations suggest that the central peak of the lower lip trace for /kuktluk/ may be largely due to protrusion for one or more of the intervocalic consonants. The protrusion in subject MB's upper lip trace may also reflect movement for consonants.[8] The implication of these observations is that protrusion movement during the consonants in /kuktluk/ is independent of its rounded vowel context. It is interesting, in this context, that both EMG OOS and OOI signals for /kuktluk/ are less strong during the consonant interval than during the rounded vowels. It seems likely that some or all of the consonant interval protrusion for these signals is due to jaw activity, perhaps as a consequence of jaw raising for the consonantal occlusions.[9]

C. Discussion

The main question behind the experiment was whether English and Turkish speakers would show the same articulatory patterns when producing similar words with rounded vowels separated by rounding-neutral consonants. The data presented here suggest that they do not. Rather than showing the consistent troughlike movement and EMG patterns exhibited by English speakers AE and AF, and reported in the literature for speakers of English, Swedish, Spanish, and French, Turkish speakers show a consistent plateaulike pattern of movement and a unimodal pattern of EMG activity (with the possible exception of the ULX signal for subject EG). Similarly, the Turkish subjects' patterns of movement and EMG contrast with the multipeaked movement pattern and troughlike EMG patterns of English speakers MB and NM. Additionally, the latter two groups differ in the degree of consonant-related protrusion seen in I-I utterances.

The look-ahead model, as modified by Engstrand (1981), might account for these data in the following way: (1) For English speakers such as AE and AF, full lip protrusion (i.e., to the degree found in rounded vowels) is prohibited during one or more of the intervocalic consonants used in this study (see footnote 5); (2) for English speakers such as MB and NM, full lip protrusion for one or more consonants is required; (3) for Turkish speakers lip protrusion is compatible with but not required during these consonants, so that the degree of lip protrusion seen is dictated by feature spreading from the segmental context. Thus, for the English speakers, consonants must have some phonetic feature specification associated with protrusion, although this may be either plus or minus, while, for the Turkish speakers, consonants are allowed to have neutral specification for this feature. It should be noted that, for this version of the look-ahead theory, because the context-independence of gestures is not a theme, there is no straightforward prediction of relationship between the U-U and I-I word data. It is possible to say, for instance, that, for speakers such as AE and AF, lessened protrusion on consonants is a reaction to a strongly protruded environment, and the behavior of the same consonants in an unrounded environment is irrelevant.

For the coproduction model, in which articulatory output trajectories are the result of combining sequences of relatively stable, independently organized gestures, data from other contexts such as the I-I words becomes more important. In this model, the fact that the central peak in the U-U word movement traces for English subjects MB and NM has a counterpart in the I-I word traces is particularly relevant, as it suggests that the consonant-related peak in the U-U word traces may be independent of the gestures for the flanking vowels. Similarly, the relative lack of movement in the I-I word traces for speakers AE and AF suggests that the trough patterns in their U-U word traces result from combining overlapping vowel gestures with a small or nonexistent consonant gesture. For the Turkish data, on the other hand, the lack of movement associated with the consonant(s) in the I-I word traces, together with the lack of a trough pattern in the U-U word traces, means that a different explanation is called for. According to the coproduction model, there are several possibilities. First, gestures for rounded vowels in Turkish (in contrast to those for English) may simply combine so as to produce a plateau pattern. This could happen, for instance, if Turkish gestures were larger or if the gesture-to-gesture interval were shorter, such that their overlap results in little or no trough. Alternatively, Turkish may have a different algorithm for combining gestures. Finally, the peculiar phonological properties of vowel harmony may result in successive rounded segments being associated with the same protrusion gesture.

The differences seen here between Turkish and English are also interesting in terms of the other theories mentioned above. For instance, if the trough in English is assumed to be a marker of syllable boundary, then the plateau pattern in Turkish may be taken to indicate that Turkish does not mark syllable boundaries in this way. Further, Turkish vowels such as /u/ are not diphthongized, so that the lack of a trough in Turkish is compatible with a diphthongal account of the trough in English. Note, however, that for these theories the lack of a trough for English subjects MB and NM is somewhat problematic. It is necessary to assume either that the explanation does not apply to all English speakers or that the specification of protrusion for the intervening consonants obscures, in some fashion, the marking of syllable boundaries or the pattern of diphthongization.

II. FURTHER TESTING

Given the data reported here, it is not possible to test either the look-ahead, the syllable marker, or the diphthongization theories further. However, the (phonetic) context-free provision of the coproduction theory makes it amenable to testing based on articulatory behavior in different phonetic contexts. In essence, the logic is as follows: If articulator trajectories over several segments reflect the combination of gestures for each of the segments, then it should be possible to deduce the basic shape of each gesture from its behavior in different contexts. It should also be possible to synthesize articulatory contours by combining their elements.

Accordingly, this section of the paper describes a series of tests based on the context-free provision of the coproduction model. In the first test, the consonant-related protrusion gestures seen for English subjects MB and NM in I-I words are subtracted from corresponding protrusion traces for U-U words. Success is a function of correspondence, for the same speaker, between subtracted traces and other U-U word traces, such as EMG traces, with no suggestion of consonant interval protrusion. In other words, because the coproduction interpretation of inconsistencies between upper lip, lower lip, and EMG signals for these speakers involves the presence of an independent consonant-related protrusion gesture, removing the additional gesture should resolve the inconsistencies. In the second test, it is assumed that, if the vowel- and consonant-related gestures seen in the corpus are independently organized, then it should be possible to construct a viable U-U word from elements in I-U and U-I words. Thus the original protrusion traces from I-U and U-I words are added together and the result compared to original U-U word traces. Success here is a function of degree of correspondence between original U-U word signals and the synthesized versions. The use of subtraction and addition for gesture combination is based on data reported by Saltzman *et al.* (1987), Saltzman and Munhall (1989), and Munhall and Löfqvist (1990).

Both tests require similar intersegment timing of consonant and vowel gestures in the I-I, U-U, and mixed-vowel words. Although explicit measures of gestural timing were not made, mean intervocalic consonant intervals (from acoustic offset of V_1 to acoustic onset of V_2) among one-, two-, and three-consonant words varied by less than 35 ms for any English or Turkish subject. This was taken as evidence that speech rate and gesture phasing were similar enough for corresponding gestures to be equated.

A. Subtraction test

Those I-I word signals showing protrusion in the consonant interval consisted of upper and lower lip movement traces for subject MB and lower lip traces for subject NM. For the first test (henceforth called the Subtraction Test), these traces were subtracted, point by point, from corresponding U-U word movement traces. The theory behind this procedure was that the underlying movement during the consonant interval, i.e., the portion of movement associated with the vowel gestures, would be the same for both I-I and U-U words.[10]

Figure 4 shows the results of subtracting averaged /kiktlik/ from averaged /kuktluk/ movement traces, superimposed on original /kuktluk/ movement traces, for these two English subjects. For comparison pur-

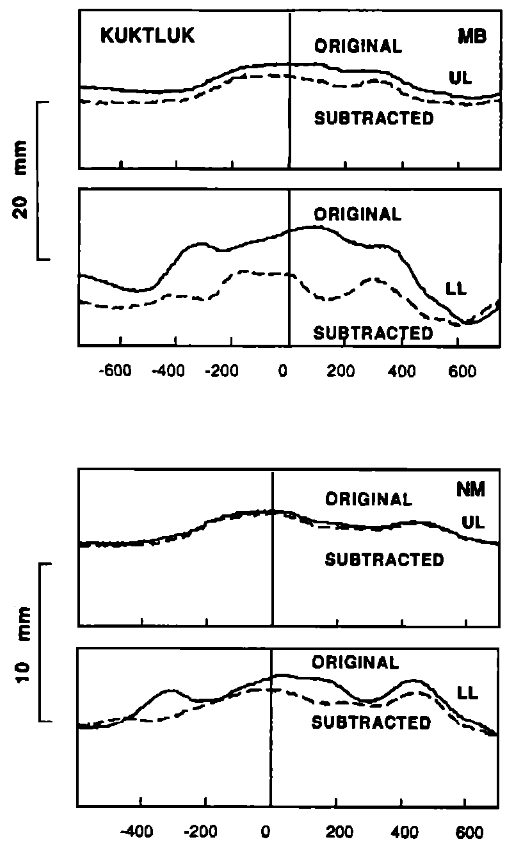

FIG. 4. Original averaged /kuktluk/ protrusion traces (solid lines) with superimposed trace achieved by subtracting original averaged /kiktlik/ trace from original averaged /kuktluk/ trace (dashed line), for English subjects MB and NM. Upper panel shows original and subtracted traces for the upper lip; lower panel shows the same for the lower lip. The vertical line is the V_1 offset. Vertical and horizontal scales are as in Fig. 2.

poses, Fig. 4 also shows the results of subtracting NM's averaged upper lip /kiktlik/ movement trace, which showed no sign of protrusion during the consonant interval, from her averaged upper lip /kuktluk/ trace.

All four subtracted traces in Fig. 4 show a trough pattern. This is to be expected for NM's upper lip trace, since her original /kuktluk/ traces showed a trough, and her /kiktlik/ trace was essentially flat. It is striking, however, that the trough patterns for both subjects' lower lip traces, and for MB's upper lip trace, correspond more neatly to these subjects' EMG trough patterns (seen in Fig. 2) than did the original U-U traces. Further, NM's subtracted lower lip trace is nearly identical to both her subtracted and original upper lip trace.

To confirm that this result of a trough pattern applied to token data for these subjects, a regression analysis of the type described above for original traces was performed, testing for quadratic and linear trend between the two peaks. Data for this analysis were produced by randomly pairing I-I word tokens and U-U word tokens with the same number of intervocalic consonants, and subtracting the I-I member of the pair from the U-U member. (Each token was paired only once, and unpaired tokens were not used.) For both subjects, coefficients for quadratic trend were positive, indicating a concave, or troughlike shape.[11] Terms for R^2 indicated that the quadratic trend accounted for a substantial portion of the variance (for /kuktluk/, subject MB: quadratic $R^2 = 0.21$, linear $R^2 = 0.23$; subject NM: quadratic $R^2 = 0.21$, linear $R^2 = 0.06$).

This result supports the hypothesis that subjects NM and MB have separate vowel and consonant-related behavior for protrusion. Further, it suggests that their vowel-related protrusion behavior—presumably con-

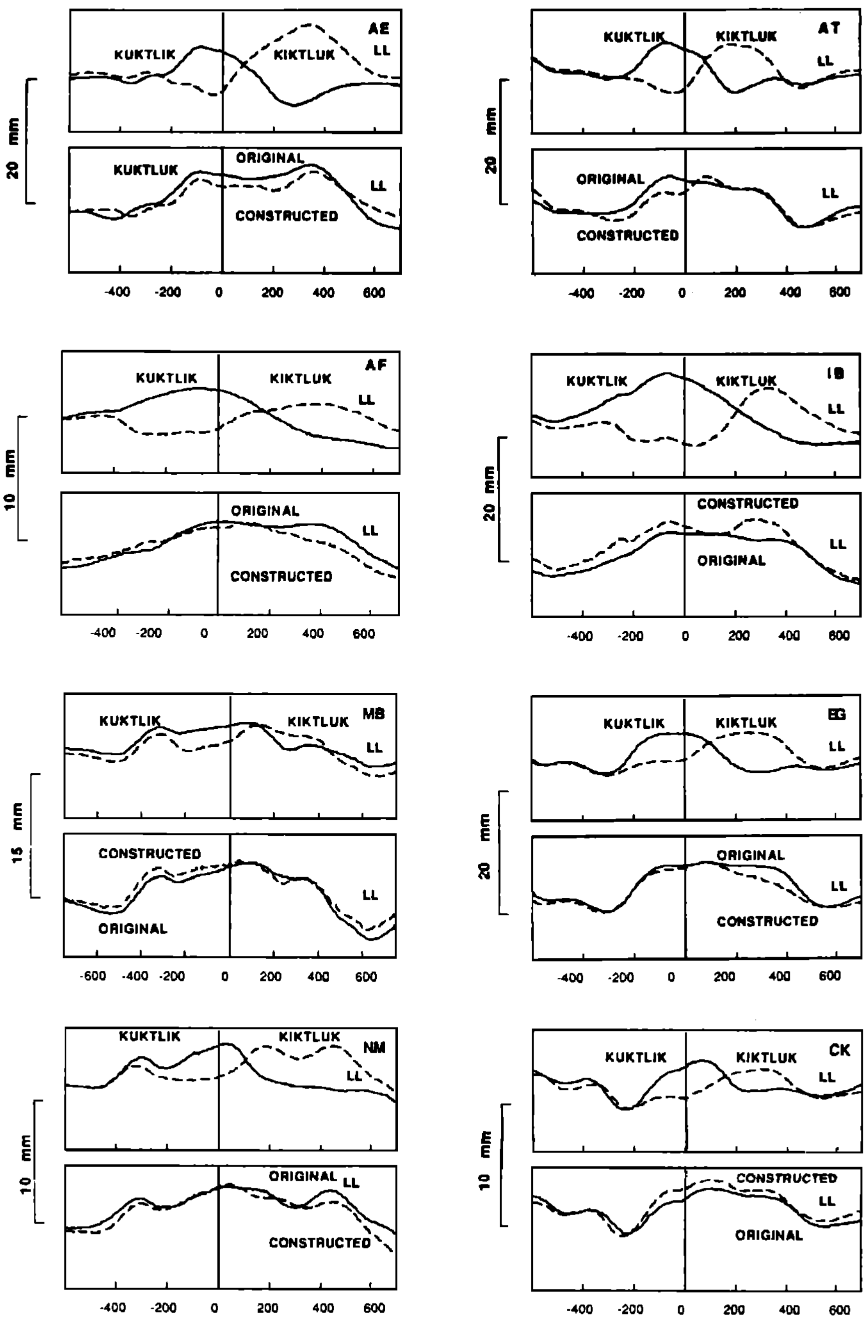

FIG. 5. Overlaid averaged lower lip protrusion traces for /kuktlik/ and /kiktluk/, plus original averaged protrusion trace for /kuktluk/ (solid line) with superimposed trace constructed by adding averaged traces for /kuktlik/ and /kiktluk/ and subtracting averaged trace for /kiktlik/ (dashed line), for Turkish speakers AT, IB, EG, and CK, and English speakers AE, AT, MB, and NM. The vertical line is the V_1 offset. Vertical and horizontal scales are as in Fig. 2.

nected to articulatory instantiation of the vowel feature of rounding—resembles that of other English speakers in being troughlike. The presence of lip protrusion during consonant articulation suggests, not a different articulatory organization, but an additional gesture overlapping with vowel-related gestures. At a more general level, this result can be taken as support for the coproduction model notion that gestures are independent entities and for the notion that gesture combination is approximately additive.

B. Addition test

For the second test (henceforth known as the Addition Test), averaged upper and lower lip movement I-U and U-I word traces were added together for each English and Turkish subject. Because the result of adding I-U and U-I word traces is theoretically equal to the result of adding U-U and I-I word traces, the averaged I-I word traces were then subtracted from each added trace[12] to produce a "constructed" U-U trace. Figure 5 shows the results of this procedure for lower lip data from /kuktluk/, /kiktlik/, /kuktlik/, and /kiktluk/ (upper lip data are substantially the same). The top panels show averaged /kuktlik/ and /kiktluk/. The traces resulting from adding these and subtracting /kiktlik/ (henceforth known as constructed traces) are shown in the bottom panels together with superimposed original U-U traces.

As the figure shows, the constructed traces paralleled the original traces quite closely for three out of four English subjects, and for two out of the four Turkish subjects. For English subjects MB and AE, in particular, the traces show a high degree of resemblance. For Turkish subject CK the principal difference is the slightly lower amplitude of the constructed trace.[13] For English subject NM, differences are also minimal. For Turkish subject EG, differences are intensification of a slight "bump" existing in the original trace plus a slightly lowered amplitude of movement during the final /u/ vowel. For the remaining subjects, however, differences are qualitatively more serious. English subject AF's constructed trace shows a single broad peak rather than a trough as in the original trace. In contrast, Turkish subject IB's constructed trace shows a trough rather than a plateau as in the original trace. Turkish subject AT's constructed trace, while paralleling the original trace during the final vowel, has a generally different shape from the plateau pattern of the original trace.

C. Discussion

The results of the Subtraction Test constitute relatively strong evidence for the generalization that English speakers produce troughs for words such as /kuktluk/, and for the notion that gestures are independent entities whose trajectories combine when overlapped in time. The subtraction test results also suggest that additivity is at least a reasonable approximation of the way that gestures combine for these articulators and these segments.

The results of the Addition Test are less clear. While the predicted and actual trajectories were close for some subjects, for other subjects they were qualitatively different. While the results were slightly better for English subjects than for Turkish subjects, the distinction between three subjects out of four (for English) versus two subjects out of four, or even one out of four (for Turkish), is hardly great enough to warrant concluding the two languages are different. It is also not clear how to interpret a lack of correspondence between constructed and original traces; for instance, the assumption of similar conditions of speech rate, stress and gesture phasing between averaged I-U, U-I, U-U, and I-I words may not be accurate. The fact that in Turkish I-U and U-I words are nonharmonic is also relevant. It is possible, for instance, that /u/ and /i/ vowels in Turkish words are always produced with independently organized gestures, but that these gestures are different in harmonic and nonharmonic words. A fuller discussion of these issues can be found in Boyce (1988).

III. GENERAL DISCUSSION

Overall, the results of this study suggest there is something very different in the way English and Turkish speakers organize articulation, at least in the way they use lip protrusion for rounded segments. The simplest index of this difference is the plateau pattern of protrusion evinced by the Turkish speakers, which contrasts with the English patterns found here, and with the trough patterns reported in the literature to date. The results for English subjects, and in particular the results of the subtraction test for subjects MB and NM, confirm that the underlying articulatory pattern for U-U words in English follows a trough pattern.

With regard to the competing coproduction and look-ahead types of models, interpretation of these results is both straightforward and complex. The straightforward interpretation is as follows. Since the coproduction model predicts a trough pattern in U-U words, and English shows a trough pattern, then English speakers employ a coproduction articulatory strategy. Since the look-ahead model predicts a plateau pattern in U-U words, and Turkish shows a plateau pattern, then Turkish speakers employ a look-ahead strategy. Thus English and Turkish have different articulatory strategies.

This interpretation gains strength from the fact that, for each model, explaining the patterns of both English and Turkish requires an additional mechanism. To explain the English trough pattern the look-ahead model must posit additional effects such as syllable-boundary marking, diphthongization, or consonant-specific unrounding in a rounded context. Similarly, to explain the Turkish plateau pattern the coproduction model must posit an unknown effect that causes I-U and U-I vowel gestures to differ from those in U-U words, or an unknown principle of gesture combination, or a loosening of the notion that gestures may be associated with only one segment. While any of these posited effects may ultimately prove to be valid, their status at this stage of investigation appears to be weak.

The complexity of this interpretation lies in the conclusion that different languages may employ different articulatory strategies. In some sense, this is to be expected, since the combination of phonology, lexicon, and syntax in different languages may impose entirely different challenges to articulatory efficiency. In fact, the hypothesis behind this comparison of Turkish and English was the notion that, in contrast to English, Turkish provides ideal conditions for articulatory look-ahead. At the same time, human beings presumably come to the task of language acquisition with the same tools and talents. The finding that current models of coarticulation are insufficient to account for language diversity indicates how difficult it may be to penetrate to the universal level of speech production. Further research, and in particular more cross-linguistic research, is needed in order to close this gap.

ACKNOWLEDGMENTS

The research in this paper was supported by NIH Grants NS-13617 and BRS RR-05596 to Haskins laboratories and by NIH Grant NS0-7040-15 to MIT. Advice on Turkish linguistics was provided by Jaklin Kornfilt and Engin Sezer. Suggestions, comments and criticism supplied by Katherine Harris, Louis Goldstein, Michael Studdert-Kennedy, Ignatius Mattingly, Fredericka Bell-Berti, Sharon Manuel, Rena Krakow, Marie Huffman, Michel Jackson, Joe Perkell, John Westbury, and the JASA review process are gratefully acknowledged.

APPENDIX

Turkish has eight vowels /iɨɑeoɵuy/ and thus (like Swedish but unlike English) has vowels which contrast only in rounding. The consonants /t/, /k/ and /l/ are nonlabial and phonemically unrounded in both languages (Lewis, 1967; Ladefoged, 1975). English and Turkish have somewhat different patterns of allophonic variation for /k/ and /l/. In Turkish /k/ and /l/ tend to be front or back according to the front/backness of the vowel of the same syllable (Clements and Sezer, 1982). In contrast, for most English dialects syllable-initial /l/ is front and syllable-final /l/ is back, i.e. velarized (Ladefoged, 1975; Keating, 1985), while /k/ varies primarily in syllable-initial position, becoming front before front vowels and back before back vowels. (Although it is sometimes referred to as "consonant harmony," the Turkish rule for /l/ and /k/ is distinct from that for front/back harmony in vowels.) Neither /l/ nor /k/ participates in roundness harmony. The sequence /ktl/ is rare in both languages, but exists, cf. English *tactless* and Turkish /paktlar/ 'pacts.' Neither language allows the initial cluster /tl/; therefore, /kuktluk/ would have the syllable structure /kukt–luk/ in both languages.

[1] The models tend to differ in the level at which compatibility is assessed. In Henke's program, the limitation was explicitly defined on an articulatory basis. Other investigators (Keating, 1988; Cohn, 1989) have postulated that coarticulation spreads by reference to feature specification at the phonological level.

[2] Rotational head movement could theoretically affect the results in two ways. First, the LEDs might have been moved into a more peripheral area of the focus field where tracking is less accurate. Subjects were constantly monitored against this possibility during the course of the experiment, and videotaped experiments were checked *post hoc*. Second, rotation of the head changes the relationship between the vertical and horizontal axes of the LED tracking system (the X and Y coordinates) and the subject's sagittal midline. Thus less or more of the subject's anterior–posterior movement relative to the midline may be detected. For some of the subjects, movement amplitude declined over the course of the experiment (although movement pattern remained stable). Note, however, that this amplitude change was mirrored in the corresponding EMG signal, suggesting an explanation in terms of some production-related factor, such as subject fatigue, rather than a tracking failure caused by baseline change.

[3] This was not done because of token variability [in terms of shape, token variability was minimal (cf. Boyce, 1988)] but because apparently flat patterns, such as those for AT, might be found to involve a slight trough.

[4] All linear and quadratic terms were significant to at least the 0.05 level. It should be noted, however, that, because of limitations built into available statistical packages, it was not possible to enter token membership as a factor in the regression analysis. Thus the terms for degrees of freedom, representing number of sample points, were extremely large (AT: $df=986$;IB: $df=1436$;EG: $df=1214$;CK: $df=1301$).

[5] Interestingly, for subjects IB and CK, the EMG pattern for U-U words resembled that for U-I words. The pattern for I-U words showed a strong peak associated with V_2. For subject EG, on the other hand, the pattern for /kuktluk/ resembled that for /kiktluk/, while the patterns for shorter words /kukuk/, /kuluk/, etc., resembled those for /kukik/, /kulik/, etc., EMG traces for these words are reported in Boyce (1988).

[6] To some extent this difference between OOS and OOI signals is a consequence of averaging, as token traces for the two signals showed differing proportions of double- and triple-peaked patterns.

[7] As noted above for the Turkish data, all linear and quadratic terms for the English data were significant to at least the 0.05 level. Again, degrees of freedom terms were large (AE: $df=1436$;AF: $df=1515$;MB: $df=1766$;NM: $df=2022$). Data from shorter U-U words generally showed a decreasing (positive) quadratic trend with decreasing numbers of intervocalic consonants (Boyce, 1988).

[8] Perusal of the traces for /kitik/, /kilik/ and /kikik/ suggest that NM shows some lower lip protrusion for each intervocalic consonant. For MB the lower lip trace shows marked protrusion for /t/ and some protrusion for /l/, while the upper lip shows protrusion only for /l/. Each of these protrusion peaks matches the consonant-interval peak of the corresponding U-U word.

[9] Comparisons of JX, JY, LLX, and LLY signals, as well as observations of the videotapes, suggest that, in addition to lip movement, these speakers use forward movement of the jaw (both rotational and translational) to produce protrusion during rounded vowels. Thus, while jaw closing movements certainly account for some of the protrusion visible during the intervocalic consonant interval, it is not clear how much of this protrusion is due to jaw closing and how much to rounding-associated jaw protrusion. It should also be noted that subject MB (and to some extent subject NM) curled her lower lip downward during rounded vowels. Because of the position of the LED, a good deal of this lip movement is reflected in the LLY trace. Interestingly, MB's LLY traces for U-U words show a trough pattern. These topics are discussed further in Boyce (1988).

[10] One intractable problem with this procedure is that /i/ and /u/ vowels also should have characteristic patterns. Thus, if retraction for /i/ vowels is present, it will be subtracted from protrusion for /u/ vowels at the same time the consonant protrusion is subtracted. In these data, the relative magnitude of protrusion dwarfed that of retraction. Thus it was assumed that the effects of subtracting the one outweighed the effects of the other.

[11] Exceptions were MB's subtracted data from /kukuk/ and /kuluk/. In general, subtracted token data followed the trend of original data in showing decreasing trough size with decreasing numbers of intervocalic consonants. This topic is treated further in Boyce (1988). Again, the degrees of freedom terms were very large (MB: $df=1497$;NM: $df=1479$).

[12] As in the first test, some residue of possible /i/ vowel-associated movement remains in these constructed traces.

[13] Note that the theory of independent gestures does not require that all gestures have identical amplitude or be produced with identical force. Such a requirement would leave no room for the effect of fatigue or for prosodic variables such as stress and syllable position. It is a common observation, for instance, that EMG or movement signals for the same word may show less amplitude at later stages during the same experiment. In this study, the fact that I-U, U-U, U-I, and I-I utterances were blocked separately may have caused some differences in overall amplitude among them.

Beckman, M., and Shoji, A. (1984). "Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese," Phonetica **41**, 61–71.

Bell-Berti, F., and Harris, K. S. (1974). "More on the motor organization of speech gestures," Haskins Lab. Stat. Rep. Speech Res. **SR-37/38**, 73–77.

Bell-Berti, F., and Harris, K. S. (1979). "Anticipatory coarticulation: some implications from a study of lip rounding," J. Acoust. Soc. Am. **65**, 1268–1270.

Bell-Berti, F., and Harris, K. S. (1981). "A temporal model of speech production," Phonetica **38**, 9–20.

Bell-Berti, F., and Harris, K. S. (1982). "Temporal patterns of coarticulation: Lip rounding," J. Acoust. Soc. Am. **71**, 449–454.

Benguerel, A.-P., and Cowan, H. (1974). "Coarticulation of upper lip protrusion in French," Phonetica **30**, 41–55.

Blair, C., and Smith, A. (1986). "EMG recording in human lip muscles: can single muscles be isolated?, " J. Speech Hear. Res. **29**, 256–266.

Boyce, S. (1988). "The influence of phonological structure on articulatory organization in Turkish and in English: Vowel Harmony and Coarticulation," unpublished Ph.D. dissertation, Yale University, New Haven, CT.

Boyce, S. (1978). "Accent or Accident?: The Acoustical Correlates of Word-level Prominence in Turkish," unpublished A. B. thesis, Harvard University (Linguistics Department), Cambridge, MA.

Brown, G. (1981). "Consonant rounding in British English: The status of phonetic descriptions as historical data," in *Towards a History of Phonetics*, edited by R. E. Asher and E. J. A. Henderson (Edinburgh U.P., Edinburgh).

Clements, G. N., and Sezer, E. (1982). "Vowel and consonant disharmony in Turkish," in *The Structure of Phonological Representations*, edited by H. van der Hulst and N. Smith (Foris, Dordrecht, The Netherlands).

Cohn, A. C. (1989). "Phonetic evidence for configuration constraints," Proc. New England Linguist. Soc. (NELS) **19**, 63–77.

Daniloff, R. G., and Moll, K. L. (1968). "Coarticulation of lip-rounding," J. Speech Hear. Res. **11**, 707–721.

Engstrand, O. (1981). "Acoustic constraints of invariant input representation? An experimental study of selected articulatory movements and targets," Rep. from Uppsala University, Department of Linguistics **7**, 67–94.

Fowler, C. (1980). "Coarticulation and theories of extrinsic timing control," J. Phon. **8**, 113–133.

Gay, T. J. (1978). "Articulatory units: Segments or syllables?, " in *Syllables and Segments*, edited by A. Bell and J. B. Hooper (North-Holland, Amsterdam).

Gelfer, C. E., Bell-Berti, F., and Harris, K. S. (1989). "Determining the extent of coarticulation: Effects of experimental design," J. Acoust. Soc. Am. **86**, 2443–2445.

Harris, K. S., and Bell-Berti, F. (1984). "On consonants and syllable boundaries," in *Language and Cognition*, edited by L. Raphael, C. Raphael, and M. Valdovinos (Plenum, New York).

Henke, W. L. (1966). "Dynamic articulatory model of speech production using computer simulation," unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Keating, P. (1985). "CV phonology, experimental phonetics and coarticulation," UCLA Work. Pap. Phon. **62**, 1–13.

Keating, P. (1988). "Underspecification in phonetics," Phonology **5**, 275–292.

Kozhevnikov, V. A., and Chistovich, L. A. (1965). "Speech: Articulation and perception," Washington, DC: Joint Publications Research Service No. 30543.

Ladefoged, P. (1975). *A Course in Phonetics* (Harcourt Brace Jovanovich, San Diego).

Leidner, D. R. (1973). "An electromyographic and acoustic study of American English Liquids, unpublished Ph.D. dissertation, University of Connecticut.

Lewis, G. L. (1967). *Turkish Grammar* (Clarendon, Oxford).

Liberman, A. M., and Studdert-Kennedy, M. (1977). "Phonetic perception," in *Handbook of Sensory Physiology, Vol. VIII, Perception*, edited by R. Held, H. Leibowitz, and H.-L. Teuber (Springer-Verlag, Heidelberg).

Lubker, J. (1981). "Temporal aspects of speech production: anticipatory labial coarticulation," Phonetica **38**, 51–65.

Lubker, J., and Gay, T. (1982). "Anticipatory labial coarticulation: Experimental, biological, and linguistic variables," J. Acoust. Soc Am. **71**, 437–447.

Magen, H. (1984). "Vowel-to vowel coarticulation in English and Japanese," J. Acoust. Soc. Am. Suppl. 1 **75**, S41.

Manuel, S. Y. (1990). "The role of contrast in limiting vowel-to-vowel coarticulation in different languages," J. Acoust. Soc. Am. **88**, 1286–1298.

Martin, J. G., and Bunnell, H. T. (1982). "Perception of anticipatory coarticulation effects in vowel-stop consonant–vowel sequences," J. Exp. Psychol.: Human Percept. Perform. **8**, 473–488.

McAllister, J. (1978). "Temporal asymmetry in labial coarticulation," Papers Institute Linguist. **35**, 1–29 (University of Stockholm, Stockholm).

Munhall, K. G., and Löfqvist, A. (1990). "Gestural aggregation in speech," J. Phon. (in press).

Ohala, J. J. (1981). "The listener as a source of sound change," in *Papers from the Parasession on Language and Behavior*, edited by C. Masek, R. Hendrick, and M. F. Miller (Chicago Linguistic Society, Chicago, IL).

Öhman, S. E. G. (1966). "Coarticulation in VCV utterances: spectrographic measurements," J. Acoust. Soc. Am. 39, 151–168.

Perkell, J. S. (1986). "Coarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movements," Speech Commun. **5**, 47–68.

Recasens, D. (1985). "Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences," Lang. Speech **28**, 97–114.

Saltzman, E. L., and Munhall, K. G. (1989). "A dynamical approach to gestural patterning in speech production," Ecolog. Psychol. **1**, 333–382.

Saltzman, E., Rubin, P. E., Goldstein, L., and Browman, C. P. (1987). "Task-dynamic modeling of inter-articulator coordination," J. Acoust. Soc. Am. Suppl. 1 **82**, S15.

Sussman, H. M., and Westbury, J. R. (1981). "The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation," J. Speech Hear. Res. **46**, 16–24.

Tuller, B., Kelso, J. A. S., and Harris, K. S. (1982). "Interarticulator phasing as an index of temporal regularity in speech," J. Exp. Psychol.: Human Percept. Perform. **8**, 460–472.