Testing a Dynamic Model of Pharyngeal Articulation

Ahmed M. Elgendy
Department of Linguistics, Stockholm University, Sweden

ABSTRACT

A dynamical model of pharyngeal articulation was designed to account for the mechanism underlying the use of the pharynx in speech production and to examine the nature of coarticulation in the back cavity of the vocal tract. Some aspects of the model are tested and its ability to predict the properties of the natural system is discussed.

INTRODUCTION

In search for the invariance units blended in the acoustic signal of speech, models have been constructed in order to account for the observed behavior of various articulators and to attempt to predict the properties of the natural system controlling the process of speech production.

Speech has been proved to be a dynamic and context-conditioned process at all levels. Articulatory dynamics involve co-ordinated movements of the articulators expressed in space and time. The question whether timing control over a speech utterance is issued externally or internally, i.e., included in the motor program, still a debatable problem (cf. e.g., [1]).

A standard model of speech production must consider activities of all parts of the vocal tract continuum together with activities of respiratory system. The nasal, oral and laryngeal portions of the pharynx constitute more than one half of the vocal tract length. Therefore, it is important to obtain exhaustive account on the physiology of the pharynx during speech. Understanding the mechanism underlying distinctive pharyngeal speech sounds would increase the efficiency of current models build up mainly for languages lacking pharyngeal phonemes and would improve our insight into the process of speech motor control.

Articulatory Dynamics of Pharyngeal Segment

In Arabic, the pharynx is used to produce distinct speech sound units both as primary as well as secondary place of articulation. The pharyngeal consonants, i.e., lower pharyngeals /ʕ, ħ/ and upper pharyngeals /q, ʁ, χ/ have the pharynx as their primary place of articulation. The pharyngealized consonants, on the other hand, use the pharynx as a secondary place of articulation with a major constriction in the oral cavity.

The production of the true pharyngeal consonants in Egyptian Arabic is characterized by a complex mechanism involving the control of co-ordinated activities of the pharynx, the epiglottis and the larynx. Sphinctric contraction of pharyngeal wall at the point of constriction occurs simultaneously with upward movement of the larynx and hyoid bone. This is accompanied by a constriction in the glottis and active bending of the epiglottis towards the arytenoids. The timing of the epiglottis movement is synchronized with a downward pull of the velum [2].

The resulting coarticulatory effect causes the jaw to sustain certain mechanical constraints realized as antagonism to the tongue movement and temporal reorganization of the syllable containing pharyngeal segment. That is, the synergies involved in controlling the production of pharyngeal segments restrict the jaw and the tongue from anticipating the articulation of the upcoming segments until the motor command is completely executed.

As a consequence, vowels are found to accommodate mandible position assigned to the pharyngeal segment intervocalically but not initially or finally in a word. The degree of jaw lowering is greater for mono-syllabic than tri-syllabic words. Acoustic analysis showed that the excessive degree of jaw lowering associated with pharyngeal segment production, compared to oral segments, is reflected as a compensatory effect on vowel duration. The degree of readjustment depends on the position of pharyngeal segment in the word and the relative degree of jaw-height of the consonant embracing the vowel at the syllable margin. That is, low vowels are longer when preceded than when followed by a pharyngeal consonant [2].

Eventually, vowels depart to their inherent position right after the gestures for pharyngeal segment are completed in initial but not in intervocalic position. Seemingly the articulators seek a rhythmic pattern among successive syllables. Furthermore, the degree of contextual overlapping between segments in non-pharyngeal environment is much greater.

The Structure of Arabic Language

The Arabic word is basically composed of three consonantal elements embedded in a finite set of vocalic patterns by which derivational forms can be generated. A vocalic pattern |≈V_1≈V_2≈| (a "tenon") is inserted into a given consonantal pattern |C_1–C_2–C_3| (a "mortise") to generate a word. For example, the vocalic pattern |≈æ≈æ≈| can be tenoned to the mortise |k–t–b| so that the word /kætæb/ will mean "he wrote". Similarly, |≈i≈ææ≈| when tenoned to the mortise |ʕ–t–b|, will yield the noun /ʕitææb/ "blaming". The inflectional constructions, on the other hand, are paradigmatically obtained by adding prefixes, infixes or suffixes to the derived forms. For instance the tri-consonantal mortise |k–t–b| can be used to generate /kætæb/ "to write"; /kotib/ "it was written"; /kotob/ "books"; /kættib/ "to cause to write"; /kæætib/ "a writer"; /kæætæb/ "he corresponded" etc.

There are five types of syllabic-pattern in Egyptian Arabic: CV, CVV, CVC, CVVC and CVCC in which a consonant must begin the syllable. Syllabic boundaries are located at the left of each consonant starting from the rightmost side moving to the begining of the word, e.g., /ʔis.tik.tææb./. There is a strong correlation between syllabic structure, stress assingment and nucleus vowel duration in Egyptian Arabic. In addition, vowel length is phonemic in Arabic, e.g., /sææd/ "to dominate" vs. /sædd/ "to block".

The nature of Arabic word structure and the degree of compensatory lengthening or shortening exerted on vowels more than consonants strongly suggest that consonants are more "stable" than vowels in terms of their duration. Thus, consonants can be considered as "landmarks" linked together by vowels.

Vowels, on the other hand, being more flexible articulatory events than consonants, can tolerate greater amounts of compression or expansion. Accordingly, vowels are issued to preserve isochronical intervals between consonants in the syllable. They also can manage various coarticulatory effects resulting from the overlapping of successive segments due to inertial and mechanical constraints.

It is appropriate, then, to take the syllable as the basic unit of motor programming in Arabic (see [2] for an overview on the components gathered for dynamic modeling of pharyngeal articulation).

THE MODEL

Figure 1 shows schematic diagram for a dynamic model by which a word containing a pharyngeal segment can be generated. The development of the model divisions is based on current views describing strategies controlling motor programs of speech production (cf., e.g., [3]).

If the mortise |ʕ–l–m|, which bears the basic semantic unit "knowing", is selected from the lexicon and the tenon |≈ææ≈i≈| is inserted into it, the word /ʕæælim/ "a scientist" will be the concept input to the articulatory plan. The tenon is recalled from the morpho-semantic storage since the vocalic pattern partially provides the grammatical meaning of the word. Notice that |ʕ–m–l| and |l–m–ʕ| are permitted but not |m–ʕ–l|, |l–ʕ–m| or |m–l–ʕ|.

An articulatory plan must be available to decode the phonological rules before the motor plan can be executed. At this stage of high-level planning, the speech sound pattern is already designed as a series of consonants and vowels in a specific order comprises lexical and morpho-semantic items. Thus, the input for the articulatory plan takes the word as the minimal unit for phonological parameters applications.

Pharyngeal consonants prevent the anticipation of the ensuing segment to take place before their execution is completed. That, in turn, would have a perturbation effect on the temporal pattern. Hence, two articulatory strategies are presumably needed to control the timing organization.

The main strategy (the default) will fail to manage the existence of a pharyngeal segment in the utterance since it assumes that all consonants are apt to equally coarticulate with the coherent vowel(s). Accordingly, strategy (2) is proposed to

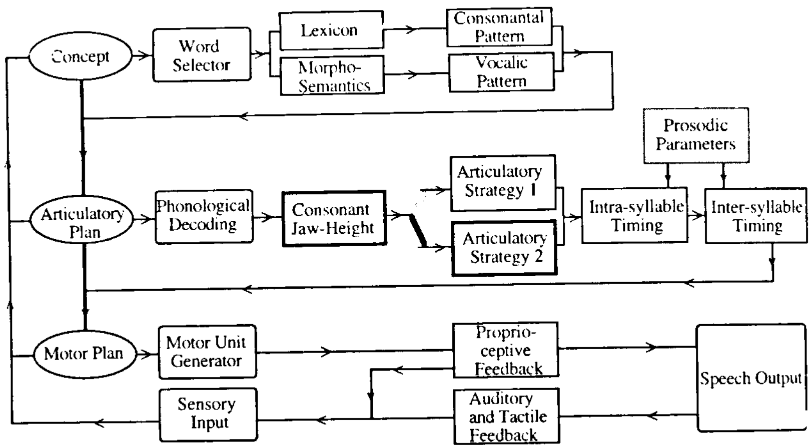

Figure 1. An articulatory model by which a pharyngeal segment can be generated.

tackle the inevitable temporal re-adjustments due to the constraints, imposed on the jaw and the tongue, characterizing the pharyngeal segment production.

The degree of re-adjustment in vowel duration, in case the final segment in a CVC is a pharyngeal segment, will depends on the degree of coarticulation between the first consonant and the following vowel. This is because the first consonant (non-pharyngeal) will allow the anticipation of the vowel. When the initial consonant is a pharyngeal, the vowel coarticulates only with the second consonant. Hence, strategy (2) will determine the degree of re-adjustment according to the degree of jaw-height of the non-pharyngeal consonant.

In both cases the value of jaw-height for both pharyngeal and non-pharyngeal consonants must be previously conceptualized. This is essential in order to maintain fixed intervals among syllables of the word. In order to keep a unitary syllable length, it is necessary that the duration of each syllable is calibrated as the combined duration of a consonant and a vowel. Hence, jaw-height must be determined for each consonant in the word since the total duration of the utterance will depend, to a great extent, on the trajectory of the jaw moving from one consonant constriction location to the next. Vowels, on the other hand, will occupy the intervals between consonants.

Thus, they serve to secure the timing regulations needed to overcome the coarticulatory effects.

The module responsible for "intra-syllable timing" determines the degree of adjustment for segment duration (mainly for the vowel). The timing among various syllables in a word will be controlled by the "inter-syllable timing" module. The temporal pattern which governs the inter-relationship between syllables will be the end product of this process.

Next, intonation contour and stress assignment rules as prosodic parameters will be applicable on the sound pattern. Recall that stress position is correlated with the duration of the syllable and its position in the word. The application of other factors effecting the rate and style of speech also pertains to this stage of phonological decoding. As soon as the articulatory plan is discharged, it is fed into an articulatory buffer before the motor plan can be commenced.

TESTING THE MODEL

Phonotactics

The validity of the model can be attested by examining its ability to predict the properties of the natural system. For this purpose the distribution patterns of pharyngeal consonants with respect to all other consonants were stated as manifested in the phonotactic rules governing spoken Egyptian Arabic word structure.

It was found that the severe mechanical constraints exerted on pharyngeal segment production has a prevailing effect over the construction of the entire language system. Consonants in a given sequence are selected according to their compatibility to preserve the temporal aspects of syllable structure. That is, the organization of consonants in a sequence depends, to a great extent, on their relative degree of jaw-height. The co-occurrence of different consonants in a word is based on the consonant's inherent degree of jaw-height. The word-length can be seen as the path the jaw takes from one consonant to the next along the word (cf., [2]).

The model can successfully predict the severe restrictions on the distribution of pharyngeal segments since it contains a module for estimating jaw-height for each element of the consonantal pattern. Thus, the degree of jaw-height for each consonant in the sequence is determined prior to the listing of timing instructions for the entire utterance.

Moreover, the model could provide an explanation for the tendency observed of pharyngeal consonants to favor initial or final rather than medial position in a word. The temporal specifications in the model are highly restricted. A pharyngeal segment in medial position will demands that the execution of the motor plan must be reset. This justifies the finding that two pharyngeal consonants do not co-occur in one and the same consonantal pattern (cf. [2]). Recall that the pharyngeal segment posses an extreme degree of jaw lowering. Furthermore, the vast majority of the vocalic pattern used in Arabic language was found to be based on the low back vowel /æ/. Low vowels are more susceptible to coarticulate with pharyngeal consonants.

Acquisition of Pharyngeals

The acquisition of pharyngeal segment production takes considerably long time compared to oral consonants and has a gradual emergence. The acquisition stage is correlated with the place of constriction in the pharynx and the relative degree of jaw-displacement associated with each consonant. That is, the greater the degree of jaw displacement of pharyngeal consonant, the longer is the acquisition time [4]. This may indicate that specifying consonant jaw-height is a primary issue in the acquisition process.

The model underscores the importance of consonant's jaw-height as a primary phonological parameter by which one of the two articulatory strategies will be chosen. The delay in the acquisition time may be due to the availability of two strategies operating the temporal organization within and between the syllables of the utterance in the mature system .

Concluding Remarks

The proposed model presupposes that well specified aspects of the lexical item are available for execution from high-level planning down to the motor commands level. The positions permitted for the pharyngeal segment to occupy in the sound patterns indicate that the effect resulting from the mechanical constraints is considered by the central planning.

On the other hand, the selection of temporally compatible segments to co-occur with pharyngeal consonants in a word indicates that "timing" is issued internally, i.e., at high level in the brain. It is suggested then, that long-term feedback is essential for most of the stages of motor control process. The model's implications lend support to the view which considers coarticulation as a preplaned articulatory process. It remains, however, to test the model's ability to account for auditory perception and the speech motor commands adaptation to peripheral contexts. The model should also be re-evaluated in the light of other models of coarticulation to examine its validity to account for a universal system.

REFERENCES

[1] **Fowler**, C.A. (1980), "Coarticulation and Theories of Extrinsic Timing". *Journal of Phonetics*, 8, pp. 113-133.

[2] **Elgendy**, A.M. (1994), "Components for Dynamic Modeling of Pharyngeal Articulation". ICSLP, Proceedings of the International Conference on Spoken Language Processing. Yokohama, Japan, September 18-22, 1994, 12-2, pp. 1-4.

[3] **Kent**, R.D. and Minifie, F.D. (1977), "Coarticulation in Recent Speech Production Models". *Journal of Phonetics*. 5, pp. 115-133.

[4] **Elgendy**, A.M. (in preparation), "Delayed Acquisition of Pharyngeal Segment Production as a Result of Mechanical Constraints".