—대한치과의사협회지 : Vol. 15, No. 4, April, 1977—

SPEECH PRODUCTION MECHANISM

Byoung W. Kim

Speech Science Laboratory University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

발음생성기전(發音生成機轉)

위스콘신—밀워키대학 언어병리학교실

조교수 김 병 욱

≫국문초록≪

우리인체의 언어기관은 중추신경에서 부터 근육에 이르기까지 광범위한 부분이다. 언어(발성)기관을 근육의 단계에서 볼때는 다음의 세가지 model(모형)으로써 언어생성(발성) 과정을 설명할 수 있다. 첫째 모형은 해부학적인 모형으로써 언어생성에 어떠한 근육들이 작용하는가를 볼 수 있고 두째번 모형은 생리학적인 모형으로 언어생성과정의 기능적인 면을 볼 수 있고 셋째번 모형은 물리학적 모형으로 말소리의 물리학적인 성질을 이해하는데에 도움을 준다. 이 세가지 모형을 이용해서 언어생성 과정을 간결하게 설명했다.

Traditionally, the speech producing apparatus is often divided into respiratory (subglottal; pulmonary), phonatory(glottal; laryngeal), and articulatory-resonatory (supraglottal; oral-nasal-pharyngeal) subsystems. The glottal and supraglottal subsystems are collectively called the vocal tract or upper airway. Speech production is also arbitrarily divided into respiration, phonation, and articulation-resonation. However, the speech production system is a unitary system of which various functional components coordinate their movements spatially and temporally. Speech production system is made of several functional components(FC) which either generate or valve the speech air stream. In this context, functional components include only muscles and movable structures. Fixed structures such as the upper teeth, the roof of the mouth(alveolus and hard palate), and the nasal cavity are

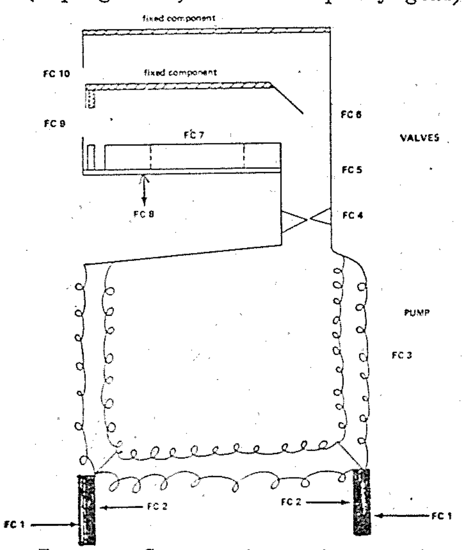
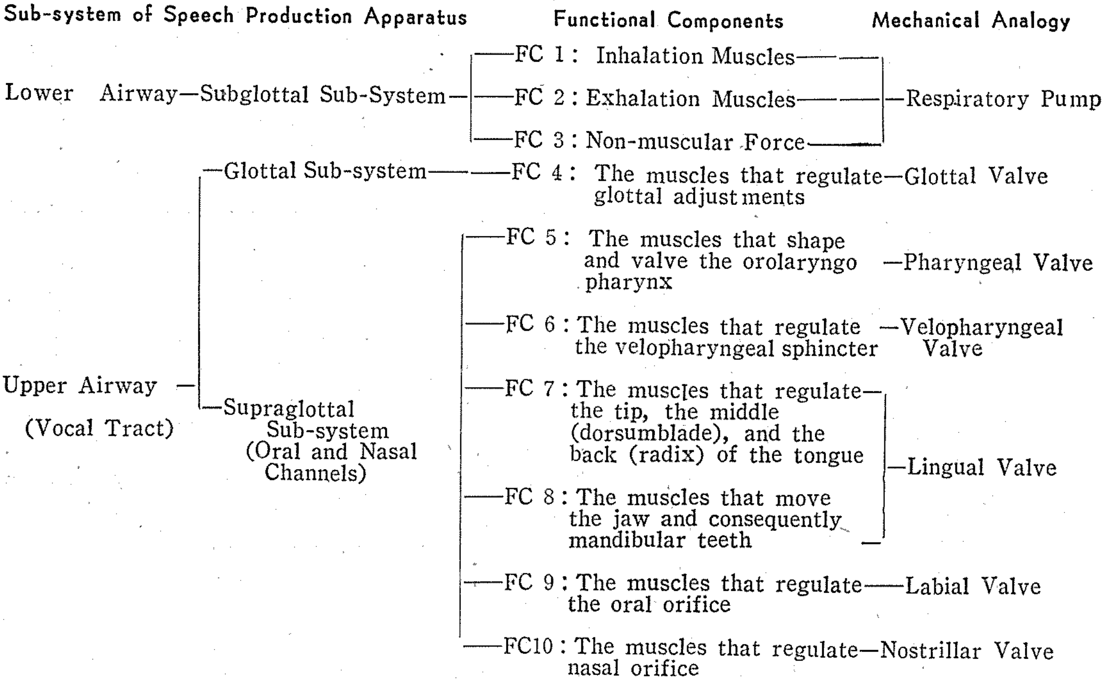


Figure 1. Structural-Functional Model of Speech Production Mechanisms

the mechanical components which are supportive to the functional components(See Figure 1). Although skeletal frameworks(laryngeal skeletal framework, thoracic skeletal framework, and facial skeletal framework) provide bio-mechanical protections and supports for speech and other biological functions, they are not described in this model. A functional component in this model is defined as a muscle group or movable structure working to generate or valve the speech air stream. A simplified description of the functional components is as follows. The multiple structural inter-connections among the functional components 4, 5, 6, 7, 8, 9 are significant: Respiratory



pump generates airflow, which is basic physiological energy(driving force) for speech, and several upper airway valves such as glottal, pharyngeal, velopharyngeal lingual and labial valves(nostrillar valve may be active only in pathological speech due to velopharyngeal incompetency) valve airflow to produce acoustic features. In essence, speech production is temporospatial coordination of respiratory pump and several selected upper airway valves to produce articulatory distinctive features required for a given phoneme to be categorically perceived as the intended phoneme.

The aerodynamic parameters of speech production mechanisms is shown in Figure 2. The basic role of the functional components and 2 is to provide and keep the air pressure inside the lungs(subglottal air pressure: Ps) relatively constant at the desired level for speech(with some variations for suprasegmental features) in the face of changing static relaxation pressure(non-muscular force) of the respiratory system provided by the functional component 3. The subglottal air pressure responsible for producing glottal airflow is the physiological index or the basic driving force for speech production. The effect of the coordinated movements of

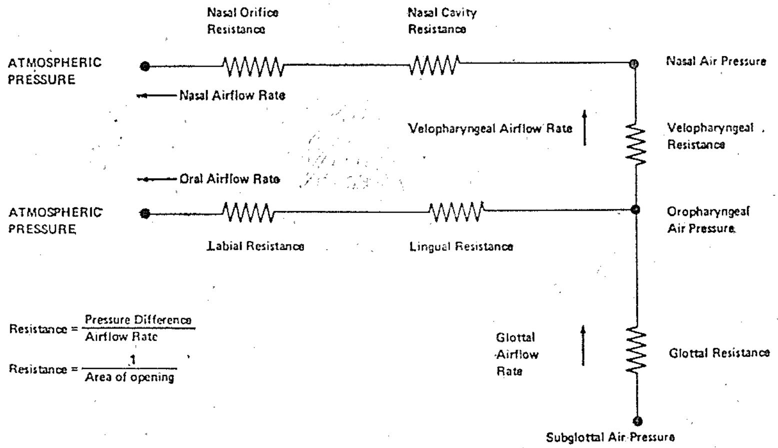


Figure 2. Aerodynamic Model of Speech Production Mechanisms

these 10 functional components is to create a series of rapidly changing air pressures and flows at various points along the vocal tract. In other words, when the respiratory pump and several upper airway valves coordinate their movements, aerodynamic pattern is determined in terms of air pressures at several points within the vocal tract, pressure differentials across the valves, aerodynamic resistances at the valves, and volume velocity of airflows at the valves. The rate at which air flows(volume velocity of airflow or airflow rate) through any of these valves is determined by the aerodynamic resistance provided by the valves and the magnitude of the pressure drop across the valve. Thus glottal airflow rate is determined by the glottal resistance and the pressure drop across the glottis. Aipflow rate through the nasal channel, in turn, is determined by the combined series resistance of the velopharyngeal port, the nasal cavity, and the nostrils. Finally, airflow through the oral channel is a function of oral channel resistance which s a net total of the combined series resistance of the tongue, the lips, and the pressure drop across these structures. In essence, the aerodynamic pattern which is a result of tempor-ospatial coordinations of functional components of the speech production system determines the nature of acoustic sources and filters.

The acoustic process involved in the production of speech sounds can be viewed as the excitation of the vocal tract configuration by one or more sources of acoustic energy according to acoustic source-filter theory of speech production proposed by Fant(1960). The basic driving force for speech production is provided by the respiratory pump in the form of airflow. The aerodynamic energy present in the form of flowing air converted into acoustic energy at the glottal valve or at any one of the several valves along the vocal tract which can be constricted to different degrees by muscular force. Valving is the mechanism for generating acoustic sources(a voice source at the glottal valve, and turbulent and transient noise sources at the glottal and/or supraglottal valves) and determining the nature of their modulation(excitation). All the upper airway valves except for the velopharyngeal valve serve dual functions; they generate acoustic sources(source function) and the alter the shape of the vocal tract, which, in turn, determines the resonating characteristics(filter function). Nasal channel is relatively fixed filter and oral channel is variable and dynamic filter. The velopharyngeal valve determines the degree of acoustic coupling between the two filters. The acoustic model(Figure 3) is based on the acoustic source-filter theory which is, in essence, that the acoustic characteristics of speech sound at the mouth opening are dependent upon the acoustical properties of the sound source(or sources) and the time-varying change

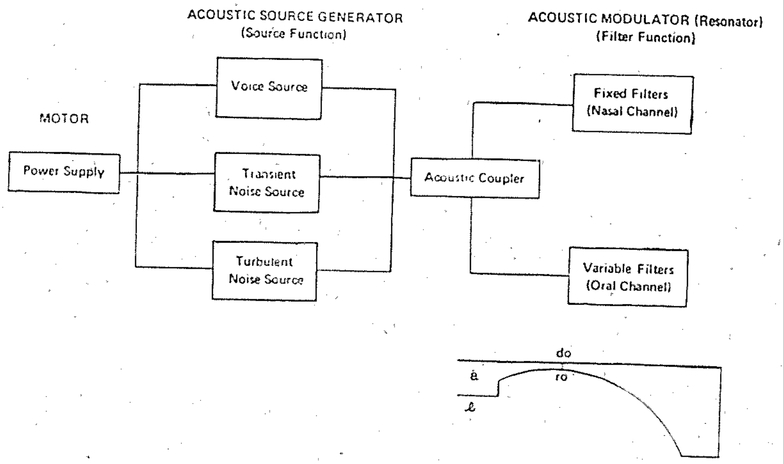


Figure 3. Acoustic Model of Speech Production Mechanisms

in the three dimensional geometry of the vocal tract. The segmental features of Kim's feature system were directly derived from the acoustic model of speech production mechanisms. Three source function features were derived from the source generators. They are voice source, transient noise source and turbulent noise source. Eight filter function features(resonation features) were derived from the first-order approximation of the time-varying change of the three dimensional geometry of the vocal tract. They are oral-nasal coupling, retroflexion, lateralization labial configuration, speed of vocalic transition, number of lingual target, point and degree of major upper airway constriction. The source features are not fully independent of filter features since movements of upper airway valves have dual function: they generate sources while they determine vocal tract shape. In essence, differing speech sounds are associated with differing acoustic sources and differing vocal tract shapes. The acoustic characteristics of speech sound at the mouth opening are dependent upon the acoustical properties of the sound source (sources) and the time-varying changes in the three-dimensional geometry of the vocal tract.

REFERENCES

Fant, G. Acoustic theory of speech production, The Hague: Mouton, 1960.

Kim, B. W. Clinical segmental feature system of English syllabic nuclei, semivowels, and consonants, Acta Symbolica, Vol. 1, No. 1, 1975.

Kim, B. W. Clinical Segmental Feature System of Korean Syllabic Nuclei, Semivowels, and Consonants. Submitted to Journal of Dental College, Seoul National University, Vol. 1, 1976