Laryngeal compensation and linguistic stress in children with velopharyngeal dysfunction

Ann Kathryn Lundberg

University of Northern Iowa

Follow this and additional works at: https://scholarworks.uni.edu/pst
Part of the Speech Pathology and Audiology Commons

Let us know how access to this document benefits you

Recommended Citation
https://scholarworks.uni.edu/pst/107

This Open Access Presidential Scholars Thesis is brought to you for free and open access by the University Honors Program at UNI ScholarWorks. It has been accepted for inclusion in Presidential Scholars Theses (1990 – 2006) by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.
LARYNGEAL COMPENSATION AND LINGUISTIC STRESS IN CHILDREN WITH VELOPHARYNGEAL DYSFUNCTION

Ann Kathryn Lundberg
University of Northern Iowa
May 1991
INTRODUCTION AND REVIEW OF THE LITERATURE

Normal Speech Physiology

The audible aspect of communication known as speech is the product of three physiological systems: respiration, phonation, and articulation. Air flow from the lungs is set into vibration at the level of the larynx. The vibrated air is then shaped by the pharyngeal, oral, and nasal cavities, and articulators to produce phonemes we understand as speech.

Respiration

Respiration is the basis for speech and provides the energy needed for vibration of the vocal folds. During quiet respiration, little muscular effort is used. Gases are exchanged between the atmosphere and the lungs based on a principle known as Boyle's Law. Boyle's Law states that air moves from regions of high pressure to regions of low pressure until the system has equal pressures. As we inhale, muscles of inspiration (diaphragm and external intercostals) expand the thoracic cavity, stretching lung tissue, thus increasing lung volume. This increase in lung volume results in decreasing
(alveolar) pressure. Air will then enter the lungs until alveolar pressure equals atmospheric pressure. When we exhale, the lungs recoil from the release of the stored potential energy in the stretched system. The compression force of this recoil causes an increase in alveolar pressure due to reduced lung volume. That is, the air molecules are now compressed within the lungs, creating greater alveolar pressure within the lungs, which will exceed atmospheric pressure, resulting in air rushing from the respiratory system.

For speech, the basic phases of respiration are modified, in that inhalation is more rapid, while expiration is extended. Further, the speaker needs to maintain a relatively constant lung volume and alveolar pressure, as alveolar pressure ultimately determines the intensity (loudness) of voice. At the beginning of expiration, the inspiratory muscles provide a checking action, since the recoil force of the lungs provides more alveolar pressure than necessary. Checking action involves the use of inspiratory muscles through the initial phase of expiration to keep air from flowing out too rapidly (Daniloff, et al., 1980). As the air continues to flow out the respiratory system, the muscles of expiration (primarily internal intercostals and abdominal muscles) serve to keep lung pressure
Phonation

The larynx acts as a valve between the lungs and the vocal tract. Laryngeal musculature regulates the glottis (opening between the vocal folds), which allows air to exit the lungs via the trachea. When the glottis is closed, air from the lungs ceases to flow. This results in a buildup of air pressure in the trachea inferior to the glottis (subglottal pressure). The interaction between laryngeal musculature and subglottal pressure causes the vocal folds to vibrate. Elements of this interaction are related to changes in vocal intensity and pitch. That is, to increase intensity and/or pitch, there is a need for increased subglottic pressure.

The vocal folds are composed of several layers of tissue which are positioned horizontally within the cartilaginous framework of the larynx. The folds come together anteriorly (in the front of the laryngeal space) to form a V-shape. They can be abducted (opened) or adducted (closed) through the contraction of various laryngeal muscles. However, it is not the action of individual muscle movements which causes vocal fold vibration for speech. The vocal folds vibrate hundreds of
times per second. This activity could not be mediated neuromuscularly, as it would be impossible for muscles to contract and relax at this rate.

Vocal fold vibration can be partially explained by a principle known as Bernoulli's Effect. Bernoulli's Effect states that when a gas travels through a constricted opening, it accelerates, creating a pressure drop at the point of constriction. The laryngeal muscles are needed to initiate the phonatory cycle. Their role is to adduct the vocal folds to build subglottal pressure. The pressure increases until it is so great that it blows the folds apart. Air rushes through the constricted opening and the pressure just below the level of the glottis drops, in accordance with Bernoulli's Effect. This drop in pressure creates negative pressure (a vacuum) which pulls the folds back together and subglottal pressure begins to build again (Daniloff, et al., 1980). This process is repeated hundreds of times per second and the valving of the exhaled air generates the sound we modify or articulate into distinctive phonemes (meaningful speech sounds).

The vocal mechanism can be described as a source-filter system (Borden & Harris, 1984). The vocal folds act as the sound source and the upper vocal tract serves to filter the
sound. The rate at which the vocal folds come together determines the lowest frequency of the sound produced, called the fundamental frequency. As sound emanates upward in the vocal tract, certain other frequencies (multiples of the fundamental frequency) are intensified depending upon the configuration of the articulators. This increase of energy at certain frequencies is known as resonance.

Articulation

The articulators of the vocal tract may be classified as either soft or bony. The lips, tongue, and facial muscles are soft articulators, while the teeth and mandible are bony articulators. The palate is bony anteriorly (hard palate) and soft posteriorly (soft palate or velum).

The velum serves to separate the oropharynx from the nasopharynx. Upon contraction of the levator muscle, the velum raises in a vector movement (upward and posteriorly) to contact the pharyngeal wall, while sphincteric muscular contraction in the superior portion of the pharynx completes closure of the velopharyngeal (vp) port. When the velum is lowered, the oropharynx and nasopharynx are said to be coupled (linked). When the vp port is closed, air rising from the lungs cannot
escape through the nose, thus pressure builds up in the oral cavity. This intraoral pressure is needed to produce many of the consonants, especially plosives and fricatives. Plosives, /p/, /t/, /k/, and /g/, result when an articulator occludes a portion of the oral cavity, permitting pressure to build up, which is then released as a burst of acoustical energy (Minifie, et al., 1973).

Other consonants, called nasals, require the vp port to be open. These consonants include /m/, /n/, and /ŋ/ (as in sing). When the vp port is open, the vocal tract increases in length. Because of the added length, the lower frequencies are amplified (Borden & Harris, 1984). Also, nasalized sounds are less intense. This is due to the fact that the nasal cavity is composed of a relatively large surface area which absorbs the sound (Daniloff, et al., 1980). The reduction in intensity of nasalized sounds is called damping.

**Suprasegmental Aspects of Speech**

Words and sentences are created by combining speech sounds into meaningful units, but in the actual production of speech, phonemes are not independent. Producing the individual phonemes in temporal order would not have the same result as producing a
whole utterance. This is due to the fact that suprasegmental features are present during speech. Suprasegmentals (or prosodic features) are those variations which are overlaid upon words, phrases and sentences (Borden & Harris, 1984). There are four main suprasegmental features: stress, intonation, duration and juncture.

Stress is used as a source of contrast in English, often creating different meanings, such as in the words 'digest and di`gest. Stress occurs on a syllable when there is a change in articulatory effort, intensity, pitch and duration. Stressed syllables (which may be classified by primary or secondary stress) are greater in intensity, higher in pitch, and longer in duration.

Even if stress is held constant, the intonation (melodic pattern) of an utterance adds a great deal of meaning. It often communicates the attitudes and feelings of the speaker. Statements and questions that cannot be answered with yes or no generally have falling intonation patterns. Falling intonation coincides with the fact that subglottic pressure gradually decreases as the end of the utterance approaches, due to decreased lung volume. This natural pattern can be overcome, however, for linguistic reasons, such as when a yes-no question
is asked.

The duration of a phoneme, vowels in particular, is dependent upon the surrounding phonemes. For example, the vowel in the word had, produced before a voiced consonant, is longer in duration than in the word hat, where it precedes a consonant made without movement of the vocal folds. This represents a general pattern in all languages, but the fact that the difference in vowel length is larger in English suggests it is a partially learned phenomenon (Borden & Harris, 1984).

Juncture refers to the differences that occur from a combination of duration changes and other sound changes. An example of juncture is the contrast between the words "I scream" and the words "ice cream." Such contrasts are important in the perception of speech.

Abnormal Speech Physiology

Abnormalities can occur in any of the three physiological systems involved in speech production. Since the systems are interdependent, a difficulty in one system may profoundly affect the function of another system. For example, respiration that is insufficient for speech purposes may not create the subglottic pressure necessary to achieve normal intensity or
One type of abnormality affecting the articulatory system is velopharyngeal incompetence (VPI). Velopharyngeal incompetence may be defined as the inability to use the velum and/or pharyngeal muscles to create the closure between the oropharynx and nasopharynx necessary for the undisturbed performance of the functions of swallowing, respiration, middle ear ventilation, or speech (Hirschberg, 1986). For speech, velopharyngeal incompetence results in inappropriate closure between the oropharynx and nasopharynx. As a result, the intensity of the speech signal is reduced, and the speaker exhibits too much nasal resonance, which is termed hypernasality. In addition, since the velum is no longer an effective valve, it is difficult for the speaker with velopharyngeal incompetence to build up intraoral pressure needed for aspects of oral articulation. This causes difficulty in producing certain consonants, especially plosives and fricatives.

Velopharyngeal incompetence can occur in many populations. For instance, it may be a functionally based, learned behavior. Deaf speakers may be functionally hypernasal because it is difficult to sense velar movement; that is, they cannot effectively monitor the difference between nasalized and non-
nasalized speech. VPI can also be the result of an organic cause such as developmental shortening of the palate, a disproportionately large pharynx, paresis of the velum, structural changes as the result of surgery, or a cleft palate. A cleft can occur in the hard and soft palate or soft palate alone and may include lip clefting as well. Clefts may be unilateral (on one side) or bilateral, overt (visible) or covert (submucous). In the case of a submucous cleft, the velum may appear to be structurally normal, but there are deficits in the velar musculature. Velopharyngeal incompetence may occur even after surgical palatal repair because deficits in the musculature of the velum persist (Bzoch, 1989).

**VPI and Suprasegmental Aspects of Speech**

The possible relationship between velar valving and laryngeal valving has been previously explored (Curtis, 1968; McWilliams, et al., 1969; Hamlet, 1973; Leder and Lerman, 1985; D'Atonio, et al., 1988). Zajac and Linville (1989) reported that cycle-to-cycle variations in fundamental frequency (jitter) were positively correlated with perceived nasality ratings by three experienced speech-language pathologists. Cycle-to-cycle variations in amplitude (shimmer) were correlated positively
with ratings of perceived hoarseness. The relationship between jitter and perceived nasality in particular suggests the larynx may be used as a compensatory valve by children with velopharyngeal incompetence in an attempt to regulate subglottal pressure for phonation. Increased subglottal pressure may also be a means of compensating for the loss of intensity due to damping when coupling occurs. However, the relationship between vp incompetence and suprasegmental aspects of speech remained unclear. While cycle-to-cycle variations suggest laryngeal compensation in children with VPI, it was not known if compensation by the laryngeal valve had an effect on linguistic stress patterns for this same population.

The purpose of this study was to examine the effects of vp insufficiency on suprasegmental aspects of speech. That is, the present investigation was designed to determine whether hypernasal children exhibit differences in linguistic stress as compared to their normal speaking peers.

METHODS

Subjects

Participants in this study were six children, ranging in age
from five to eight, judged to be hypernasal by two certified speech-language pathologists. These subjects were each matched by sex and age with three "normal" speaking peers. The members of this control group were all students enrolled at the Malcom Price Laboratory School, Cedar Falls, Iowa. The normal speakers had no known history of speech, hearing, or neurological anomalies.

Speech Sample

A speech sample of each subject was recorded on a Marantz (PMD 340) audiocassette recorder. A Superscope (EC-7) condenser microphone was placed 15 cm from the lips. Subjects were asked to prolong the production of the vowels /a/, /i/, and /u/ as long as possible. WH-Question pictures (Teaching Resources Corporation, 1982) were used to elicit spontaneous speech. Each subject was asked to describe the actions that appeared on the stimulus pictures. If necessary, specific questions were asked by the examiner from a list of questions printed on the back of each stimulus picture.

Stress type analysis

The speech samples were orthographically transcribed and a
corpus of linguistic utterances commonly exhibited by all subjects was identified. Primary and secondary stress was assigned to each word in the common sample, according to the procedure described by Campbell and Shriberg (1982). For the purpose of this study, the primary stressed syllable was defined as the strongest syllable in the utterance. All other stress types, including secondary, tertiary, and minimal stress types were considered non-primary stressed syllables.

Sample analysis

The speech samples were analyzed by the use of the Visipitch (Kay Elemetrics, Model 6097). Segments of speech were truncated and line fed from the Marantz (PMD 340) recorder directly into the Visipitch to obtain fundamental frequency measurements. The Visipitch was programmed to analyze a 135-535 cycle/second bandwidth over a two second time interval.

Duration and intensity measurements were obtained by line-feeding samples into a Macintosh Two computer, using GW Instrument's Mac Audios data acquisition hardware and MacSpeech Lab II software. Durational measurements were computed in milliseconds and average intensity measurements in volts were calculated by the software program.
Results

For each utterance, fundamental frequency, duration, and intensity measurements were calculated for the entire phrase and each syllable within the phrase. This data was then used to calculate measurements of each syllable as a proportion of the total utterance. In this way, data could be compared across subjects. These proportional measurements were subjected to three independent analyses of variance.

Analyses of variance revealed significance in stress type (p<.05) for all three prosodic features. In other words, the primary stressed syllables differed significantly in their intensity, frequency, and duration when compared to the non-primary stressed syllables.

Interaction between speaker type and stress type was also significant (p<.05). Post hoc t-tests were performed to determine the sources of significant differences in this data. These t-tests revealed that the hypernasal speakers and normal speakers produced significantly more intensive primary than non-primary stressed syllables. This result was consistent with expectations, since greater subglottic pressure is needed to produce stressed syllables, which also would result in an increase in intensity. For the data concerning frequency, post
hoc t-tests showed that normal speakers used a significant increase in frequency when producing primary stressed syllables, as compared with non-primary stressed syllables. This result would also be consistent with expectations, since an increase in subgottic pressure causes the vocal folds to vibrate more often, increasing pitch. Post hoc t-tests for the durational measurements revealed that the hypernasal speakers used significantly longer primary than non-primary stressed syllables.

Discussion

The results of this study indicated that the normal-speaking children increased their intensity and pitch when producing primary stressed syllables. They did not rely on an increase in duration in the production of stress. The stressed syllables they produced were slightly shorter in duration than the non-primary stressed syllables. This result was inconsistent with previous data that indicated a durational increase with the production of primary stress. A possible explanation for this inconsistency may be related to the fact that all subjects were under eight years of age. Research has indicated that suprasegmentals continue to develop throughout adolescence.
(Kent, 1981). Therefore, the results may indicate an incomplete development of suprasegmentals in our subjects.

When examining the results obtained from the hypernasal subjects, it is evident that they used vocal intensity for primary stressed syllables in a manner similar to their normal speaking counterparts, increasing their intensity on these syllables. The hypernasal speakers also exhibited an increased fundamental frequency level when producing primary as opposed to non-primary stressed syllables, although not to the extent of the normal-speaking control group. This decreased difference in frequency may be a result of their physiologically deviant vocal tract hampering them from implementing frequency as a variable in the production of suprasegmentals. It may also be explained by the fact that these speakers have become habitually less dependent on frequency in the production of primary stressed syllables. Due to their deviant system, the hypernasal children may have learned an alternate pattern of prosodic speech production.

Analysis of the durational data of the hypernasal subjects revealed that these subjects produced significantly longer primary stressed syllables. This result may indicate a type of laryngeal compensation on the part of these speakers. Since
they did not rely on an increase in frequency when producing stress, they may have used an increase in duration to compensate. Therefore, a deviance in the upper vocal tract may be causing differences in the physiology of the lower vocal tract for this population.

Further research is needed in this area to substantiate the present study. More information should be obtained in the area of stress production and suprasegmentals in both normal children and children with velopharyngeal incompetence. With the advent of further research, coupled with this investigation, speech-language pathologists may be better able to serve this population and aid them in acquiring more normal speech patterns.
References


Table 1. Mean Relative Intensity, Frequency, and Durational Measures (expressed as a ratio of phrase values) for Hypernasal and Normal Speaking Children.

<table>
<thead>
<tr>
<th>SUPRASEGMENTALS</th>
<th>SYLLABLE TYPE</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HYPERNASAL SPEAKERS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELATIVE INTENSITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>1.230</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td>DURATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>0.389</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.304</td>
<td></td>
</tr>
<tr>
<td>FUNDAMENTAL FREQUENCY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>1.034</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.961</td>
<td></td>
</tr>
<tr>
<td>NORMAL SPEAKERS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELATIVE INTENSITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>1.323</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.823</td>
<td></td>
</tr>
<tr>
<td>DURATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>0.345</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.354</td>
<td></td>
</tr>
<tr>
<td>FUNDAMENTAL FREQUENCY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIMARY STRESS</td>
<td>1.125</td>
<td></td>
</tr>
<tr>
<td>SECONDARY STRESS</td>
<td>0.961</td>
<td></td>
</tr>
</tbody>
</table>
INTENSITY EXPRESSED AS A RATIO OF SYLLABLE TO PHRASE INTENSITY.

- NORMAL
- HYPERNASAL

LINGUISTIC STRESS
FREQUENCY EXPRESSED AS A RATIO OF SYLLABLE TO PHRASE FUNDAMENTAL FREQUENCY.

- NORMAL
- HYPERNASAL
DURATION EXPRESSED AS A RATIO OF SYLLABLE TO PHRASE DURATION.

- NORMAL
- HYPERNASAL

LINGUISTIC STRESS

DURATION (proportion of phrase length)

Primary
Secondary