An Acoustic Study of Coarticulation Patterns in Turkish and English

by

Hale H. Ozsoy

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2000
[June 2000]

Author ....
Department of Electrical Engineering and Computer Science
May 2000

Certified by ....
Kenneth N. Stevens
Clarence J. LeBel Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

An Acoustic Study of Coarticulation Patterns in Turkish and English

by

Hale H. Ozsoy

Submitted to the Department of Electrical Engineering and
Computer Science in May 2000, in partial fulfillment
of the requirements for the degree of

Master of Engineering in Electrical Engineering and
Computer Science

Abstract

Studies in Turkish and English had shown that there is a remarkable difference in the way native speakers of these two languages employ coarticulation in production of rounded vowels separated by nonlabial consonants. The motivation for such studies was to determine whether a correlation between coarticulation and the phonetic structure of the language exists. Understanding specifics of coarticulation is of great interest since it sheds light onto motor organization of speech production. Most of the previous studies were based electromyographic (EMG) signals generated by upper and lower lip movements. Not only is such methodology very cumbersome and impedes research, but also the equipment attached to the speakers' mouth to obtain the measurements is likely to induce the speaker to alter the way he/she normally produces utterances. This present study suggests that acoustic analysis is sufficient to characterize differences in coarticulation patterns of languages. Spectral analysis of recorded signals corresponding to nonlabial consonants in different vowel environments showed that speakers of Turkish and English employ different coarticulatory organizations.

Thesis Supervisor: Kenneth N. Stevens
Title: Clarence J.LeBel Professor of Electrical Engineering

Acknowledgements

My deep gratitude goes to my thesis advisor, Ken Stevens. Without his guidance, encouragement, insightful comments and patience, this thesis would not have been possible. I feel very lucky to have had the chance of working with Ken. I have always thought he is an incredible combination of knowledge kindness.

Special thanks go to talented and supportive researchers in the Speech Communication Group who always took the time to answer my questions and to Arlene who always made me feel at home with her warm smile. I would also like to thank to every single one of my subjects who took the time to help me in spite of their busy schedules.

Heartfelt thanks go to Cagri for supporting and keeping me sane through this exhausting process. Without his help, this thesis would have taken me twice longer.

Last but not the least, I would like to thank my parents and brother who always made me feel that they will be on my side no matter what I do. Their love and unending support has always amazed me.

Table of Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Speech Production and Articulation

Speech production begins with a flow of air passed by the respiratory system. The respiratory system, consisting of lungs, rib cage, diaphragm, and abdominal wall muscles, provides energy for sound production in the form of a relatively constant subglottal air pressure through the control of lung volume. The vibrating vocal folds convert this air stream to a series of air puffs which vibrate at or near their natural frequencies once excited. Above the vocal folds, the vocal tract consists of pharyngeal, nasal and oral cavities. In the source-filter acoustic theory of speech production, the vocal-tract is considered as an acoustic tube with varying cross-sectional area that is excited by a source at either the glottal end or at some point along its length. The shape of the vocal tract determines how the source is to be filtered. Different filtering leads to distinct resonance patterns associated with different vowels or consonants.

The process of modifying the vocal tract shape is called articulation, and is achieved through dynamic control of the articulators: tongue, lips, velum and soft palate. Articulators not only change configurations of the vocal tract but also can generate speech sounds that consist of noise due to turbulence, and do not depend on vocal fold vibration. If vocal fold vibration accompanies this process, the noise generated at constrictions is superimposed on the vocal tone and the output is referred to as a voiced sound. If there is no vocal fold activity, then the output is referred to as an unvoiced sound. In production of a sequence of vowels and consonants, articulators change the configuration of the vocal tract to achieve the resonances associated with each distinct sound. Thus, from an articulatory point of view, production of an utterance composed of segments with different feature specifications is a continuous event without obvious segment markers.

Perceptually, speech is considered to consist of discrete units also referred as phonemes or segments. Each unit has different articulatory requirements. How these discrete units are fitted into the continuous acoustic stream is explained by coarticulation. Since the speech organs are unable to change states instantaneously, the overlap of articulatory gestures corresponding to neighboring segments is inevitable. The question is whether this overlap is a natural consequence of simple sequencing of segments with different feature specifications or is the output of a motor planning mechanism that scans an utterance beforehand and moves toward an articulatory goal associated with an upcoming segment in advance. Coarticulation is the mechanism behind this overlapping of the acoustic and articulatory consequences of neighboring segments.

1.2 Models of Coarticulation

Theories of coarticulation can be grouped into two main categories, the first one being centered on the widely cited "look-ahead" model by Henke(1966). In this model, each phoneme in an articulatory string is conceived as composed of a bundle of articulatory features. When phonemes for which a given feature e.g., rounding, is unspecified, precede phonemes for which that feature is specified, i.e, +rounded, the human speech apparatus moves toward the goal of lip rounding in advance of the phoneme marked +rounded. The model suggests that the central motor control of speech is constantly seeking goals that change in time, so it scans utterances in advance and initiates the required gesture toward an articulatory target as soon as there are no conflicting prior requirements on the relevant articulators.

The second category that is in direct contrast with the first one is built around the "time locked coarticulation" model by Bell-Berti and Harris(1981). This model is also referred to as coproduction. It suggests that articulation is context-independent and the articulatory gesture associated with a phonetic segment begins at a specific time period prior to the onset of that segment independent of context. What this means in speech motor control context is that there is no scanning mechanism but a simpler sequencing of articulatory configurations. Coarticulation then is a result of temporal overlap between independent articulatory gestures associated with neighboring segments. Thus, the overlap between configurations is not a sign of movement toward a goal in advance but an unavoidable consequence of limitations on articulators in modifying their configuration instantaneously.

Many studies have been conducted to resolve the conflict in these opposing models. Some of them provided evidence for a look-ahead model whereas the others for a time-locked model. The difficulty of making measurements on movements of articulators such as tongue, velum and soft palate has led scientists to build articulation models on a specific form: labial coarticulation.

1.3 Anticipatory Labial Coarticulation

Anticipatory labial coarticulation refers to the early onset of rounding for rounded vowels. It has received a lot of attention for two reasons: i) lips are more accessible to observation than the lingular, pharyngeal, or laryngeal systems by virtue of their location and ii) lip movements reflect extremes of anticipation since many phonemes are unmarked for labiality and allow the speaker greater freedom to anticipate labial features.(Sussman and Westbury, 1981; Lubker, 1981).

To understand the time course of labial activity, electromyographic signals (EMG) generated by labial musculature have been recorded as subjects uttered words in which two rounded vowels are separated by a nonlabial consonant or a cluster of such consonants. Not surprisingly, the two models have contradicting predictions for articulation of these segments of interest.

The look-ahead model of coarticulation suggests that as a speaker produces a sequence of phonetic segments, he spreads the features of the following segments onto the preceding parts as long as these parts are unspecified for a given feature. What this translates into in the environment of interest, i.e., where two rounded vowels are separated by an alveolar consonant, is that the rounding feature spreads unto the consonant. Once the speaker rounds his lips for the production of the first vowel and moves into the consonant, the scanning mechanism in the motor control system identifies the +rounding feature for the upcoming vowel. Thus, it preserves lip protrusion during the consonant as long as the production of the consonant does not conflict with rounding. Lip protrusion is terminated at the end of the second vowel.

The time-locked model of coarticulation brings forth a contrasting mechanism of articulation for the same environment where rounded vowels are separated by a nonlabial consonant. The model treats articulation to be context-independent and suggests the initiation of gestures to be locked to a specific period before the onset of the associated segment. Thus, it predicts lip protrusion to be terminated at the end of the first rounded vowel and reinitiated at a specific period before the onset of the second rounded vowel, even though the production of the intervening consonant does not have requirements conflicting with lip rounding.

The electromyographic signals have been observed to display a double-peak phenomenon in rounding activity such that if two rounded vowels are separated by even one alveolar consonant, labial EMG activity is minimized during the consonant and is restarted for the second rounded vowel. The consistency in results of follow-up studies led scientists to reevaluate the models of coarticulation in light of this double peak phenomenon.

A variety of explanations have been proposed. Obviously, the results constituted a major problem for the proponents of the look-ahead model. The articulatory planning mechanism introduced by the model required scanning of the upcoming rounded vowel and thus preservation of rounding during the production of the intervening consonant -a prediction which is in direct contrast with the observations. To deal with this discrepancy in the model, a new definition of articulatory unit came forward. Base on a proposal by Kozhevnikov and Chistovich (1965), Gay (1978) stated that articulatory units are syllable-sized, and coarticulation occurs only within a single syllable. Since the two rounded vowels belong to different syllables, there is no overlap of rounding features associated with the two. Evidence against syllable-sized articulation was provided by a study of Harris and Bell-Berti (1984).

The double peak pattern seemed to confirm predictions of the time-locked model. However, further studies carried out to identify the context-independent temporal window proposed by the model had varied and conflicting results. Gestures associated with a particular segment did not show a stable profile across different segmental contexts as suggested by the model. It is acknowledged that gesture profiles are impacted by stress and other prosodic contexts. (Tuller et al., 1982). Studies by Bell-Berti and Harris (1979, 1982); Engstrand(1981) supported that there is a time window of a stable duration attached to the onset of the rounded vowel at which rounding is initiated. On the other hand, studies by Lubker(1981), Lubker and Gay (1982) Sussman and Westbury(1981) and Perkell (1986) reported a lot of variation in how long before the onset of the rounded vowel, lip rounding is initiated. Thus, they provided support for the hypothesis that coarticulation is heavily context-dependent.

Some studies took a different approach by questioning the basic assumption that alveolar consonants are rounding neutral and therefore questioned the validity of the previous interpretations. Engstrand(1981) argued that rounding may interfere with the optimal acoustic/aerodynamic conditions for alveolar consonants, and the termination of rounding during the production of the consonant was a consequence of this fact. A follow-up study by Gelfer et al. (1989) showed that such acoustic/aerodynamic conditions may not hold for all subjects.

Since all models had weaknesses in explaining the double peak behavior, they continued to coexist. To identify the mechanism of coarticulation further and to determine if there is a consistency in behavior of native speakers of different languages, scientists conducted cross-linguistic studies. A proposal was made that phonological structure of a language could have an influence on coarticulation behavior.

1.4 Language and Coarticulation

Phonology determines the typical structure of words by introducing constraints on possible sequencing of segments. Languages differ in their inventories of distinctive sounds and in their constraints on sequencing of these sounds. The constraints pose different challenges to the central articulatory planning unit. As a child acquires language, he becomes familiar with linguistic information about the structure of the language through interaction with fluent speakers. Different combinations of phonology, lexicon and syntax occurring in different languages impose entirely different challenges to articulatory efficiency. From this point of view, it is natural for speakers of different languages to use different strategies for coarticulation. On the other hand, all humans are born with identical articulatory apparatus and their speech systems are subject to same physiological limitations. Thus, there should be at least some sort of similarity in the speech of speakers of different languages as they produce articulatory strings with identical requirements.

As far as coarticulation is concerned, such an invariance in mechanism was observed in some crosslinguistic studies and disputed by the others. An important hypothesis deduced from this variability in results was that languages with similar phonological constraints are similar in their coarticulatory mechanisms and those with different phonological specifications employ coarticulation in an entirely different manner. Thus it may not be possible for a single model of coarticulation to grasp the behavior observed in different languages.A formal attempt to identify differences in coarticulation structures of languages came from Boyce in 1990.

1.5 Differences in Coarticulation patterns of Turkish and English

Boyce began by acknowledging that the double peak pattern seen in English was also observed in French, Spanish and Swedish as subjects uttered a sequence of rounded vowels separated by nonlabial consonants. She noted that these languages are alike in their phonological tolerance for combining rounded vowels with unrounded vowels in a single word. Thus, it is natural for speakers of these languages to employ the same pattern of labial coarticulation as English speakers. The next step then was to check whether a constraint in combining rounded and unrounded vowels causes a different coarticulation mechanism. Boyce utilized Turkish, a vowel harmony language with strict rules for the possible sequencing of rounded and unrounded vowels.

In her study, she compared the labial activity of native English speakers to that of Turkish speakers as they produced two rounded vowels separated by alveolar consonants. Her beginning hypothesis was that the difference in phonology would lead Turkish speakers feel more pressured to utilize anticipatory labial coarticulation and to spread lip rounding onto an intervening alveolar consonant. As predicted, EMG signals from Turkish subjects had a “plateau” pattern of rounding whereas those from English speakers had a "trough" pattern. Turkish speakers clearly preserved lip protrusion whereas English subjects terminated it during the production of the alveolar consonant.

The results of Boyce's study led to an important conclusion: There is variation in coarticulation of features in languages with different phonological constraints for these features. English was better represented by the time locked model of labial coarticulation whereas Turkish data was more consistent with the look ahead model. Boyce pointed out that the insufficiency of a single model of coarticulation to account for language diversity indicates the difficulty of penetrating into the universal level of speech production. It is necessary to pursue further cross-linguistic studies through which interaction between phonology and coarticulation can be better understood.

1.6 Motivation for this thesis

Understanding coarticulation has been of great interest in speech research since it sheds light upon the motor organization of anticipatory coarticulation. Whether the motor control system scans the string of segments and adjusts the onset of rounding activity to the amount of time available or simply sequences independent segments has a big impact on dealing with context dependency in speech synthesis and recognition systems. Once the rules of coarticulation are identified, mechanisms can be sought to incorporate them into recognizers and synthesizers.

Determining the correlation between phonological structure and coarticulation is a major advancement in understanding human speech production and mechanism of language acquisition. Once this correlation is clear, phonological constraints of a language can be utilized to enhance speech recognition and synthesis in that language. This can only be achieved through extensive cross-linguistic studies where languages are compared to each other.

The previous studies discussed above have already revealed important aspects of coarticulation. However, for the most part, they utilized EMG signals from lip movements that are detected by placing adhesive surface silver-silver chloride electrodes to lips of subjects. The electrodes pick up any activity from orbicularis oris inferior and orbicularis oris superior which are the muscles that control lip rounding. Not only is this methodology complex, but also the apparatus attached to the subjects' lips is very likely to induce them to alter the way they normally produce utterances. An alternative to this tedious method is to study the associated acoustic output.

Acoustic analysis is undoubtedly requires less effort in terms of data acquisition, is noninvasive, and hence does a better job of grasping natural speech. Many scientists did not base labial coarticulation studies on acoustical analysis since they were uncertain about the possibility of screening the time course of rounding effectively by paying attention to only the acoustic output. This concern arose from the lack of reliable parameters to detect spreading of rounding unto the intervening consonant. However, there is a well-developed theoretical model of the relation between vocal tract configuration during consonant production and the resultant acoustic patterns. The effects of rounding on consonant spectra are known. These can be utilized to devise parameters that capture the extent to which rounding is spread into a consonant adjacent to a rounded vowel.

This thesis argues that it is possible to confirm Boyce's results through acoustic analysis. The initial step will be to review the acoustic modeling of alveolar consonants and effects of rounding on the modeling. Then, by comparing the spectrum of the same consonant in rounded versus unrounded vowel environments in light of these effects, we will be able to estimate the time course of rounding.

1.7 Thesis Outline

Acoustic consequences of labial activity of six Turkish and six English subjects are analyzed as they produce consonants in rounded versus unrounded environments. Four different consonants are utilized: unvoiced alveolar fricative /s/, its voiced cognate /z/, unvoiced alveolar stop /t/ and palatoalveolar fricative /sh/. Chapter 2 reviews acoustic modeling of these consonants and effects of rounding on their spectra. Chapter 3 describes the experiment, parameter design and methods of data analysis. Chapters 4 and 5 discuss the results for utterances with /s/ and /z/ respectively. Chapter 6 discusses the additional study done on /t/ and /sh/ utterances. The following three chapters present the data for each of the four consonants. Chapter 7 takes a consolidated view and combines results from the different parts to reach general conclusions about i) use of acoustic analysis in coarticulation studies ii) the relation between phonology and coarticulation, and gives suggestions for further research.

Chapter 2

THEORETICAL BACKGROUND

2.1 Alveolar fricatives /s/ and /z/

All fricatives are produced by forming a narrow constriction in the front part of the vocal tract. For alveolar fricatives /s/ and /z/, the constriction is made by holding the tongue blade against the alveolar ridge. The air pressure developed at the constriction makes the air flow turbulent, resulting in a frication noise. The alveolar fricative /s/ is unvoiced since no vocal fold vibration occurs during its production, whereas /z/ is voiced. The exact place of the constriction has a strong influence on the spectral shape.

The production of alveolar fricatives can be modeled as an acoustic tube with two constrictions; one at the glottis, with area A_g and one at the alveolar ridge, with area A_c, as shown in Figure 2.1. The area of the glottal opening, A_g, is adjusted to be greater than the cross-sectional area, A_c, of the constriction. The subglottal pressure P_{sub} causes air to flow through the system with volume velocity, U.

The system can be decomposed into a quarter wave resonator representing the cavity in front of the constriction and a half-wave resonator representing the portion between the constriction and the glottis.

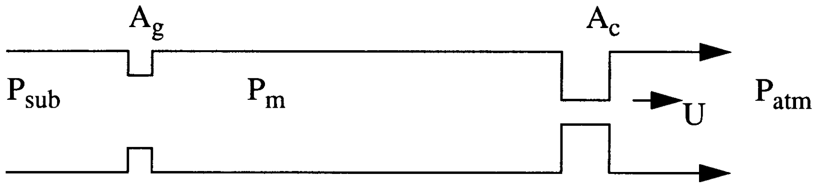

Figure 2.1: Simple model of the production of a fricative consonant

The low-frequency circuit analog of this concatenated tube is given in Figure 2.2 with pressure analogous to voltage and volume velocity to current. (Stevens, 1993).

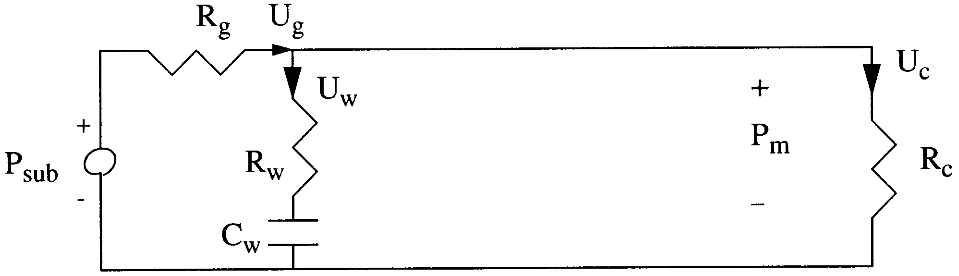

Figure 2.2: Low frequency equivalent circuit of the model in figure 2.1. (From Stevens, 1993). P_{sub} represents subglottal pressure. R_g and R_c are resistances due to glottal and supraglottal constrictions respectively. C_w is the acoustic compliance of the vocal tract walls. R_w is the resistance that represents the behavior of the walls at low frequencies. U_g, U_w, U_c are the airflow at the glottis, the walls and the constriction respectively. P_m is the intraoral pressure.

The vocal-tract pressures and flows can be estimated from Figure 2.2. The pressure drop across the constriction, ΔP_c, is given by the equation:

$$\Delta Pc = \frac{Rc}{(Rc + Rg)} Psub = \frac{1}{1 + \frac{A_c^2}{A_g^2}} \tag{2.1}$$

where $R_c \sim 1/A_c^2$ and $R_g \sim 1/A_g^2$

Since the subglottal pressure is assumed to be constant, the pressure drop at the constriction is determined by the relative areas of the constriction and glottis. The pressure drop at glottis can be found by using the relation $P_{sub}=\Delta P_c + \Delta P_g$. To solve for U, the volume velocity in cm^3/sec, the following formula is used where ρ is the density of air (ρ=.00114 g/cm3), A is the area of the constriction in cm^2, ΔP the pressure drop across the constriction in $dynes/cm^2$ and k is a constant that depends on cross-sectional shape of the constriction and is approximately equal to 1:

$$U = \sqrt{\frac{2\Delta P}{k\rho}}A \tag{2.2}$$

The amplitude of the frication noise source is proportional to $P_m^{1.5}A_c^{0.5}$(Stevens, 1971; Shadle, 1985) where P_m is the pressure in the mouth.

The production of /z/, the voiced cognate of /s/, has two deviations from this scheme.The first one is a lower P_m. There needs to be a balance between maintaining enough pressure drop across the glottis to sustain vocal fold vibration. The second effect of the vocal fold vibration is the increased laryngeal resistance, i.e., reduction of the air-flow. Figure 2.3 displays examples of the spectrograms for an /s/ and /z/ pronounced by an American male. Both have energy concentrated at frequencies above 4.5kHz. Vocal fold vibration accompanies production of /z/ as shown in the figure.

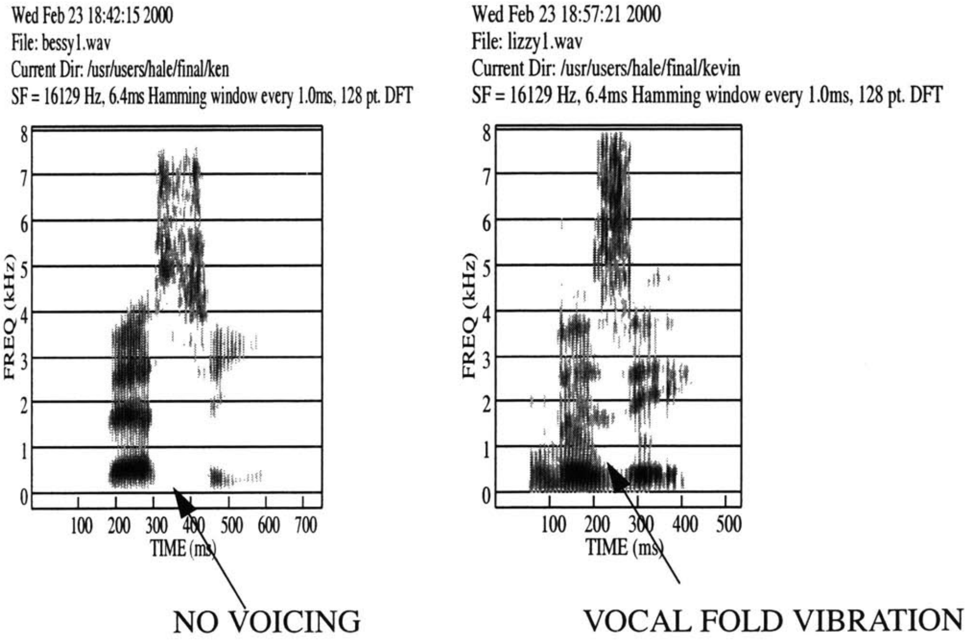

Figure 2.3: Spectograms of /bessy/ and /lizzy/ respectively.

2.2 Spectrum of an alveolar fricative

The spectrum of the sound radiated from the lips can be considered as the product of the spectrum of the noise source which is taken to be a sound pressure source, the transfer function from the source to volume velocity at the lips and the radiation characteristic. The filtering function is determined by the natural frequencies of this configuration. It can be analyzed in terms of its poles and zeros. Poles are the natural resonances of the system whereas zeros are frequencies for which the input to the system produces zero output.

When area of the constriction, A_c, is very narrow compared to the area of the concatenated tubes shown in Figure 2.4, the coupling between the cavity in front of the constriction and the cavity behind the constriction is negligible. Thus, to a first approximation, the constriction can be replaced by a rigid wall as shown in Figure 2.5.

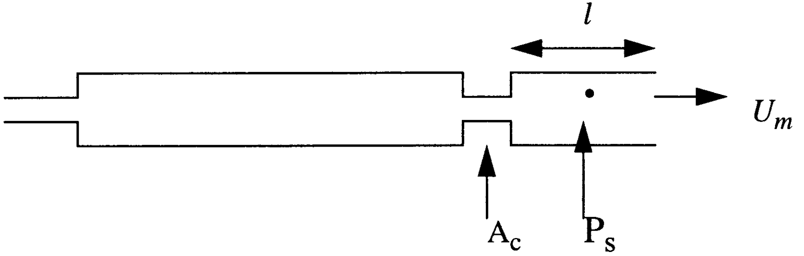

Figure 2.4: Model of vocal tract for an alveolar fricative in which a pressure source P_s is located at a fixed position in front of the constriction.

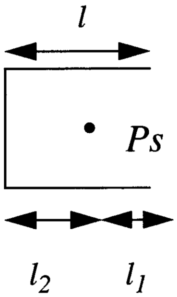

Figure 2.5: The model in Figure 2.4 when the constriction is considered as a rigid wall.

With this decoupling approximation, the only resonances are those due to the cavity in front of the constriction assuming no acoustic loss. The transfer function of the vocal tract becomes

$$T = \frac{jA\sin(kl2)}{\rho c\cos kl} \tag{2.3}$$

where l_2 and l are as shown in Figure 2.5 and $k=w/c$. The poles are given by $coskl=0$ which yields $f_{pn}=c/4l;\ 3c/4l;\ 5c/4l;....(2n-1)c/4l$ (Stevens, 1998). As is apparent from the formula, the length of the portion in front of the constriction l determines the location of the poles.

If the length of the front portion is taken to be 2cm, the average value for an adult male speaker producing /s/ or /z/, the lowest resonance of the configuration as given by the equation $c/4l$ is around 4425Hz. This resonance is excited by the turbulence noise in the front cavity and appears as a peak in the spectrum of the fricative. Minor peaks corresponding to resonances of the cavity behind the constriction may be evident at lower frequencies in the spectrum. Figure 2.6 shows the calculated spectra of the alveolar fricatives /s/ and /z/. As seen from the figure, the prominent peak of frication occurs around 4.5 kHz in a neutral vowel environment. This corresponds to the F4 or F5 range of the following vowel. Experimentally, it is observed that the amplitude of frication noise is higher than that of the following or preceding vowel in this frequency range.

The effect of rounding is to lengthen the front portion of the vocal tract. Since the length of an acoustic tube is inversely proportional to its natural frequencies, the resonances associated with the front portion will be shifted downwards.

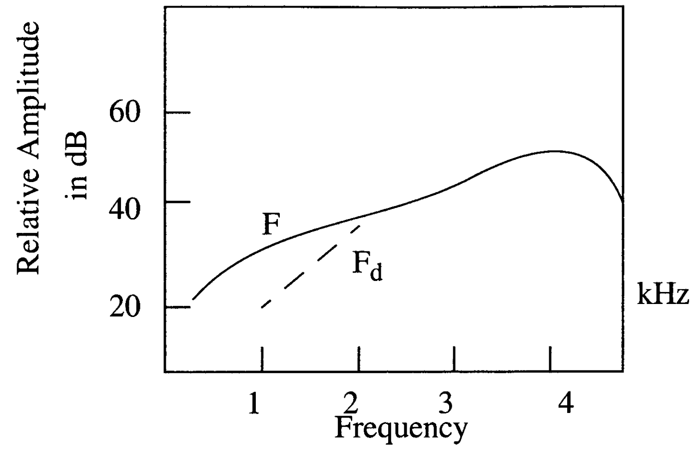

Figure 2.6: Spectrum of an alveolar fricative, including the effects of monopole and dipole sources. F_d is the spectrum of dipole frication noise in middle of fricative. (From Stevens, 1998).

Thus, if a speaker preserves rounding in production of an /s/ in between two rounded vowels, the prominent peak in frication will occur at a lower frequency than that of an /s/ in a rounding neutral environment. Using the equation c/4l, we note that the principal pole of the spectrum will occur at 2950Hz if lip protrusion causes *l* to increase from 2cm to 3cm. Monitoring the frequency of the prominent pole in unrounded versus rounded environments will reveal if rounding is preserved during the production of the alveolar fricative.

2.3 Alveolar Stop /t/

The production of a stop consonant can be decomposed into three stages: i) a closure is made at a particular point along the vocal tract, ii) pressure is built up behind this closure, iii) the closure is released. The consequences of these events are the inhibition of vocal fold vibration during the closure interval, an abrupt decrease in amplitude at the implosion and an abrupt increase at the release in some frequency regions, and spectral changes at the implosion and the release depending on the location of the closure (Stevens,1998).

In the case of alveolar stops, the closure is formed with the tongue tip. The tongue body is placed in a forward position to enable contact between tongue tip and alveolar ridge. The constriction in an idealized vocal tract shape is shown by Figure 2.7.

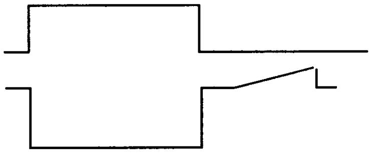

Figure 2.7: Idealized vocal tract shape for an alveolar stop consonant (Stevens 1998). The location of the constriction, i.e., the place of articulation, highly impacts the spectrum of the stop and the formant transitions from the adjacent vowels. These issues will be discussed later.

The fronting of the tongue results in tapering of the area in the anterior part and widening in the posterior area. As the constriction is released, there is an increase in the area of the constriction formed by the tongue tip. If the following vowel is a front vowel, there will only be the downward movement of the tongue blade. If the upcoming vowel is a back vowel, then an additional backward movement will occur. The sequence of events following the release of the closure is illustrated in Figure 2.8. There is an initial transient which is followed by an interval of turbulence noise at the constriction. Aspiration noise may be produced prior to reinitiation of vocal fold vibration.

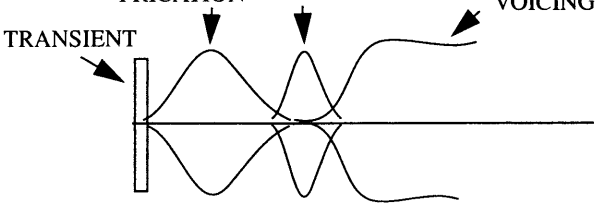

Figure 2.8: Sequence of events during the production of a stop (Stevens,1998).

The circuit equivalent of the configuration as the closure is made is identical to the one shown in Figure 2.2. Using the formulas given for the alveolar fricative, the transfer function from the volume velocity at the constriction to the volume velocity at the mouth can be calculated. As in the spectrum of the fricative, there will be a prominent peak at 4.5kHz assuming that the length of the front portion is 2cm. Again if lip protrusion is preserved, that is the front portion is longer than 2 cm, then the prominence will occur at lower frequencies. So by comparing the spectra of alveolar stop in rounded vowel environment versus unrounded vowel environment we will be able to detect spreading of rounding unto the consonant. Figure 2.9 shows spectogram of a /t/ in unrounded vowel environment in English.In Turkish, t and d are said to be dental, so the burst can be different.

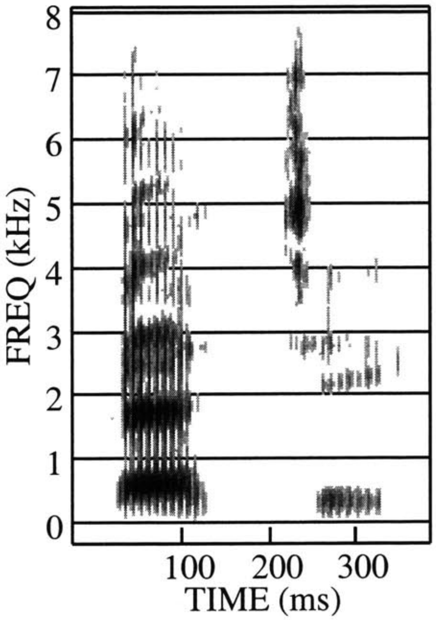

Figure 2.9: Spectrogram of /betty/ produced by an American male.

2.4 Palatoalveolar Fricative /sh/

The constriction in the case of /sh/ is made a few millimeters posterior to the alveolar ridge. Behind the constriction, a narrow channel is formed between the tongue blade and the hard palate. The portion of vocal tract posterior to the channel has a cross-sectional area that is relatively larger. The airflow becomes turbulent due to the air pressure behind the constriction, which results in a frication noise. Figure 2.10 displays an approximate model of the anterior part of the vocal tract configuration for the production of a palatoalveolar fricative.

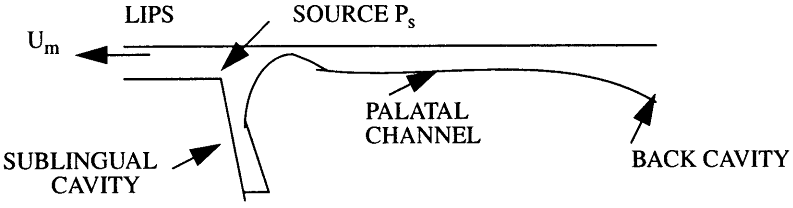

Figure 2.10: Vocal tract configuration for /sh/. (From Halle and Stevens,1991).

Since the area of the back cavity is significantly larger, its effects on the resonances of the front part will be small. Then the front portion can be modeled as an acoustic tube open at both ends whose natural frequencies will be given by c/2l where c is the velocity of sound and l is the total length of the branches. To a first approximation, the overall length of the branches is about 8.5 cm (Stevens, 1998). The resonances then will be spaced by approximately 2kHz which results in two natural frequencies in the region below 4kHz.

Another simplifying assumption is the division of the system into two parts by the constriction. Then the resonances of the anterior and posterior parts can be calculated separately. The palatal channel, having an average length of 4cm. for a male, would have a resonance at 2200 Hz if it were closed at the constricted end. Since it is not fully closed at the constriction, a natural frequency of 2.6kHz can be estimated for this portion.The sublingual cavity is roughly 2.6cm long and will have a resonance around 3500 Hz if the lip cavity is ignored. The zeros of the system estimated by a similar argument appear at roughly 1000 Hz and 3000Hz (Stevens, 1998).

A pole-zero pair below 2kHz is inserted to account for acoustic coupling to the cavity behind the constriction. The effect of rounding on the transfer function is to shift the poles downward by increasing the length of the front portion. Figure 2.11 shows an example of the spectogram for /sh/.

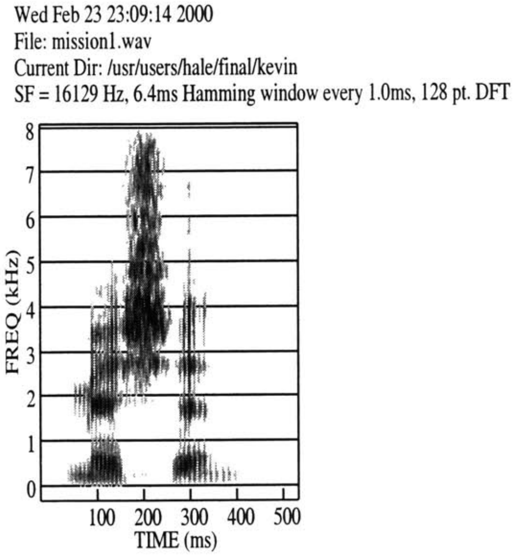
Wed Feb 23 23:09:14 2000
File: mission1.wav
Current Dir: /usr/users/hale/final/kevin
SF = 16129 Hz, 6.4ms Hamming window every 1.0ms, 128 pt. DFT

Figure 2.11: Spectogram of a /sh/ in /mission/

Chapter 3

EXPERIMENT

3.1 Subjects

Six speakers of American English and six speakers of Standard Turkish participated in this study. Subjects were recorded as they produced five randomized versions of word lists of twenty four utterances. Out of the six subjects for each language, three were female.

3.2 Wordlists

In Turkish, rounding operates according to a harmony rule that causes sequences of high vowels within a word to acquire the rounding specification of the preceding leftmost vowel. The result is long strings of rounded or unrounded vowels whose rounding is predictable given the first vowel in the sequence. (Boyce, 1990).There are a lot of exceptions that come from Arabic and Persian borrowings. In contrast to Turkish, English combines rounded and unrounded vowels freely.

Utterances of the form V_1CV_2 were utilized where C was /s/, /z/, /t/ or /sh/ and V_1 and V_2 were vowels. For the analysis of each consonant, six utterances were selected in which the consonant was preceded and followed by two vowels with identical rounding specification. When formal words that fit the purpose were not found, nonsense words were made up abiding by the rules of the language.

Utterances with /s/ in rounded versus unrounded vowel environments

Three of the words in lists for both languages consisted of /s/ in between rounded vowels and three words with /s/ in between unrounded vowels. All of the Turkish words were formal words. Some words in English were made up. Lists were randomized and five tokens were recorded for each subject. The utterances with /s/ for Turkish and English are displayed in Tables 3.1 and 3.2 respectively.

/s/ in unrounded environment	/s/ in rounded environment
hasan	kusur
karasin	korusun
keser	tosun

Table 3.1: Wordlist for Turkish containing /s/

/s/ in unrounded vowel environment	/s/ in rounded vowel environment
bessy	husu
lissa	losso
tissi	russo

Table 3.2: Wordlist in English containing /s/

Utterances with /z/ in rounded versus unrounded environments

The production of /z/ is almost identical to that of its voiceless cognate /s/ except for the accompanying vocal fold activity. There were six utterances with /z/ in the wordlists of both languages. In three of these utterances, /z/ was surrounded by two unrounded vowels and in the remaining three, by two rounded vowels. Again, some words were nonsense words made up to fit the analysis abiding by the phonological constraints of the languages. Tables 3.3 and 3.4 display the utterances with /z/ for Turkish and English respectively.

/z/ in unrounded environment	/z/ in rounded environment
bezek	bozuk
ezin	ozon
kazi	kuzu

Table 3.3: Wordlist in Turkish containing /z/

/z/ in unrounded vowel environment	/z/ in rounded vowel environment
buzzy	bozzo
lizzy	kuzzu
tazzi	tozzu

Table 3.4: Wordlist in English containing /z/

Utterances with /t/ in unrounded versus rounded environments

Six utterances contained /t/. As usual, in three, the consonant was surrounded by unrounded vowels and in the remaining three, by rounded vowels. Turkish speakers do not flap /t/ in any setting and English speakers were warned to avoid flapping their /t/'s as much as possible. Table 3.5 and 3.6 show the words for both languages.

/t/ in unrounded vowel environment	/t/ in rounded vowel environment
atar	kutu
beter	motor
kati	otur

Table 3.5: Wordlist in Turkish containing /t/

/t/ in unrounded vowel environment	/t/ in rounded vowel environment
beater	huttu
betty	kottu
rita	motto

Table 3.6: Wordlist in English containing /t/

Utterances with /sh/ in unrounded versus rounded vowel environment

In the Turkish alphabet, there is a distinct symbol for /sh/ which is an /s/ with a tail attached to the bottom. For convenience, utterances were written with /sh/ instead of that special symbol. Of the six utterances, /sh/ was surrounded by rounded vowels in three and by unrounded vowels in the other three. Tables 3.7 and 3.8 display the /sh/ utterances.

/sh/ in unrounded environment	/sh/ in rounded vowel environment
bashak	burushuk
besher	koshu
bitishik	kushum

Table 3.7: Wordlist in Turkish containing /sh/

/sh/ in unrounded environment	/sh/ in rounded environment
audition	hushu
mission	kushu
station	moshu

Table 3.8: Wordlist in English containing /sh/

3.3 Instrumentation and Method

The recording of subjects was done with a ceiling-hung microphone placed approximately 20 cm. away from the speaker's mouth in a sound attenuated room. All utterances were low pass filtered at 7500 Hz and digitized at 16kHz. Analyses conducted on the waveforms varied depending on the spectral properties of the consonant that intervened the rounded vowels.

3.3.1 Analysis on Average Spectra Of Fricatives /s/ and /z/

The average spectra for fricative consonants /s/ and /z/ were calculated using a series of 6.4 ms windows starting 10 ms after the first vowel/consonant boundary and ending 10 ms before the consonant/second vowel boundary. The motivation for using a smaller window was to smooth the signals to simplify the identification of the peaks. The spectra were divided into high and low frequency regions. 4kHz was taken as the boundary for separation of the high/low frequency regions in the male data, since these consonants had a major prominence above 4kHz in a neutral vowel setting. If lip rounding was preserved during the production of the consonant, this prominence shifts down to roughly 3kHz. For the female data, a boundary of 5Khz was used to separate the high and low frequency regions.

The average spectrum for an English alveolar fricative was compared to that of a Turkish one.The peak picking feature of the Klatt spectrum analysis program in the Speech Communication Group at MIT was activated to identify the two peaks with highest amplitude above the boundary and below the boundary. Figure 3.1 illustrates the identification of high and low frequency peaks in the calculated average DFT spectrum of /s/. Peaks in the spectrum of /z/ are identified in a similar fashion. The upper curve is a smoothed version of the average spectrum and is not utilized in this study. Peaks marked as Peak1 and Peak2 in the figure belong to low frequency region and peaks marked as Peak3 and Peak4 belong to the high frequency region.

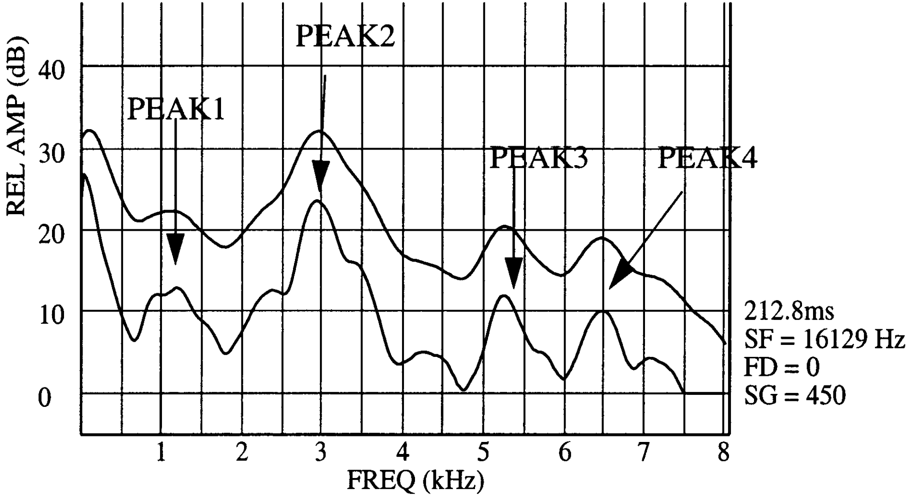

Figure 3.1: Identification of peaks in average DFT spectrum of /s/ in rounded vowel environment for a Turkish male speaker.

The frequency locations and amplitudes of these four peaks, the duration of the consonant, and the amplitude of the first formant of the second vowel 20 ms after the consonant/vowel boundary were measured for each token of the utterances. To simplify the analysis and comparison, some parameters were defined.

Parameters

The following parameters are used in the analysis. Prominence criteria is based on the amplitude, i.e. what is referred to as the most prominent peak is the one with the highest amplitude. All amplitudes are measured in dB.

AP1: The amplitude of the most prominent peak below the 4kHz boundary in the male data and below 5kHz in the female data.

AP2: The amplitude of the second most prominent peak below 4kHz in the male data and below 5kHz in the female data.

AP3: The amplitude of the most prominent peak above 4Khz in the male data and above 5kHz in the female subjects.

AP4: The amplitude of the second most prominent peak above 4Khz for the male data and above 5kHz for female subjects.

AVHG: The arithmetic average of AP1 and AP2.

AVLW: The arithmetic average of AP3 and AP4.

DFRNC: The difference of AVHG from AVLW.

AV1: The amplitude of the first formant of the adjacent vowel 20 ms after the consonant/vowel boundary (averaged over a 20 ms window).

All of these parameters were measured for each utterance. The range of values of DFRNC in rounded vowel environments was compared to the range of values of DFRNC in unrounded vowel environments. A positive value of that parameter indicates that prominence is above the boundary, i.e above 4kHz for male and 5kHz for female data. A negative value implies that prominence was shifted to lower frequencies through lip rounding. Statistics of the distribution of DFRNC are calculated for better comparison.

3.3.2 Time Analysis for /s/ and /z/

This analysis was done on the utterances where /s/ or /z/ are surrounded by rounded vowels to provide an estimate of spreading of rounding throughout the consonant. The parameter DFRNC (as defined above) was measured at three time points during the duration of the consonant for one token of each wordlist. No averaging of spectra was utilized. The three time points at which measurements were made are as follows: i) 10 ms after the first vowel/consonant boundary; ii) the midpoint of the consonant and iii)10 ms before the consonant/second vowel boundary. The following parameters are used to represent the value of DFRNC at different time points.

DIFF_BEG: The value of DFRNC 5 ms after the vowel/consonant boundary. This parameter gives an idea whether there is a significant carryover of rounding form the preceding vowel.

DIFF_MID: The value of DFRNC at the midpoint of the consonant. This location is the least likely place to be effected by the rounding feature of the neighboring vowels. Thus, it serves as a good measure to test whether rounding is preserved throughout the entire duration of the consonant.

DIFF_END: The value of DFRNC 5 ms before the consonant/vowel boundary. This parameter will reflect effects of anticipation of the rounding feature of the upcoming vowel.

The variation in the value of DFRNC at these three points reflects how rounding feature of the surrounding vowels impacts the spectra of the consonant through time.

3.3.3 Analysis on /t/ utterances

A different averaging technique was used for analyzing /t/. The initial burst was averaged over 15 ms avoiding the aspiration region. The resultant average spectra was analyzed using the method described in average spectra analysis of /s/ and /z/. The method did not work very well because of the problems explained in Chapter 6.

3.3.4 Analysis of /sh/ utterances

There are two prominent peaks in the spectra of /sh/, one in low and one in high frequency region. The effect of rounding on these two peaks were analyzed independently. The spectra of /sh/ was averaged starting 10 ms after the first vowel/ consonant boundary and ending 10ms before the consonant/second vowel boundary. The frequencies and amplitudes of the two prominences were recorded in rounded and unrounded vowel environments for eight speakers. The following parameters were defined to simplify the analysis.

PEAK1: The prominent peak in low frequency region at around 2.5kHz,

FREQ1: The frequency of PEAK1,

AMP1: The amplitude of PEAK1,

PEAK2: The prominent peak in high frequency region at around 4.5kHz,

FREQ2: The frequency of PEAK2,

AMP2: The amplitude of the prominence at PEAK2.

These parameters were measured for utterances with unrounded vowels as controls. The expectation was to see these prominence occur at the same frequency regions for both languages in the unrounded vowel setting, i.e., have comparable values for PEAK1 and PEAK2 since the difference in coarticulatory organization comes into surface only in rounded vowel environment. Speakers are likely to employ the same organization of coarticulation as they produce a sequence of two rounded vowels separated by /sh/ as they have used in utterances with /s/ and /z/. Thus, a significant shift toward lower frequencies in at least one of PEAK1 and PEAK2 was expected in rounded environment for Turkish. Such a shift was not expected in English since previous analysis showed that speakers do not modify their production of the consonant depending on the rounding specification of the adjacent vowels.

Chapter 4

RESULTS FOR UTTERANCES WITH /S/

4.1 Results of the analysis on the average spectra for males

The value of DFRNC was calculated on average spectra of each utterance containing an /s/. The beginning hypothesis was to see a deviation in the way Turkish and English speakers employ labial coarticulation in production of a sequence of two rounded vowels separated by a rounding neutral consonant. As controls, initially the values of DFRNC in unrounded vowel environment are calculated for both languages. The languages behave similarly in an unrounded environment. Table 4.1 shows the results for Turkish /s/ utterances in between unrounded vowels.

Table 4.1: DFRNC for Turkish /s/ in unrounded environment, male speakers

utterances	Subject CS	Subject OO	Subject OK
hasan-1	**13.10**	**2.85**	**10.65**
hasan-2	**6.15**	**6.50**	**9.25**
hasan-3	**6.95**	**6.25**	**10.75**
hasan-4	**3.45**	**2.35**	**15.05**
hasan-5	**8.55**	**5.20**	**17.30**
karasin-1	**8.25**	**8.35**	**10.60**
karasin-2	**10.35**	**3.05**	**11.20**
karasin-3	**6.40**	**3.25**	**11.35**
karasin-4	**7.95**	**-1.45**	**9.35**
karasin-5	**9.55**	**3.90**	**8.95**
keser-1	**10.30**	**-6.55**	**11.20**
keser-2	**-0.75**	**5.50**	**11.35**
keser-3	**4.95**	**12.10**	**8.60**
keser-4	**14.10**	**7.00**	**10.90**
keser-5	**7.05**	**0.65**	**3.85**

A positive value of DFRNC implies that the prominence in spectra of the consonant is located above 4kHz. For Turkish males, this parameter is positive 93% of the time across all utterances as seen from Table 4.1. There is no lip protrusion that would have shifted the frequencies of the major peaks down. Since the surrounding vowels are rounding neutral, the speakers have no motivation to round their lips during the production of the alveolar consonant no matter what coarticulatory organization they employ.Table 4.2 shows the results for English male subjects.

Table 4.2: DFRNC for English unrounded environment, male speakers

utterances	KS	AM	KO
bessy-1	8.10	8.35	16.25
bessy-2	5.30	6.85	12.60
bessy-3	12.10	10.40	11.60
bessy-4	4.40	8.10	15.30
bessy-5	13.35	8.20	14.00
lissa-1	3.95	9.00	7.85
lissa-2	11.05	6.25	11.75
lissa-3	10.60	6.30	12.10
lissa-4	6.10	6.15	13.20
lissa-5	8.30	5.15	8.55
tissi-1	11.35	1.30	11.55
tissi-2	9.50	5.30	13.20
tissi-3	11.55	5.15	8.00
tissi-4	7.90	4.35	11.10
tissi-5	8.80	1.45	12.60

100% of the time the value of DFRNC is positive. The prominence is above 4kHz as in Turkish. As expected, both languages have dominant peaks at high frequencies. Figure 4.1 shows the mean and standard deviation for each speaker in the unrounded vowel environment.

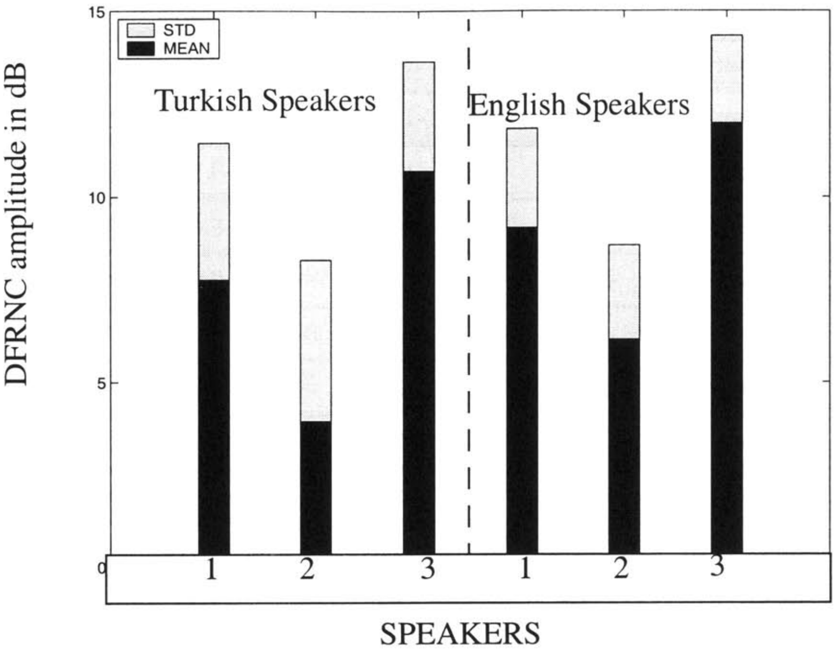

Figure 4.1: Statistics for /s/ in unrounded environment for male speakers.

The mean value of DFRNC is positive for all Turkish and English speakers. In terms of the standard deviation from the mean, the results of both languages are comparable, too.

The values of DFRNC in rounded vowel environments sheds light on whether the speakers preserve or terminate lip rounding during the production of the consonant. If DFRNC takes comparable values in rounded versus unrounded environments, then the implication is that the speaker terminates rounding as he moves onto the consonant. If DFRNC takes values that are significantly less than the values in unrounded vowel environment, then lip rounding is spread onto the consonant and shifts the prominent peaks down in frequency.

As seen from Table 4.3, there is a remarkable difference between the values DFRNC takes in rounded versus unrounded vowel environments in Turkish. 98% of the time, DFRNC is negative, i.e. the prominence which was at around 4.5 kHz in unrounded vowel environment is shifted below 4kHz.

Table 4.3: DFRNC for Turkish Males in rounded vowel environment

utterances	CS	OO	OK
korusun-1	-8.75	-11.20	-8.40
korusun-2	-10.45	-13.50	-5.10
korusun-3	-10.40	-12.00	-8.30
korusun-4	-10.30	-14.65	-8.65
korusun-5	-8.50	-8.75	-9.05
kusur-1	-10.85	-12.80	-10.55
kusur-2	-11.50	-13.85	-7.95
kusur-3	-7.95	-7.75	-4.20
kusur-4	-10.80	-9.40	-8.90
kusur-5	-10.95	-11.10	-11.10
tosun-1	-6.80	3.60	-3.60
tosun-2	-12.30	-12.95	-5.20
tosun-3	-8.80	-15.65	-7.05
tosun-4	-7.40	-19.45	-9.45
tosun-5	-8.55	-10.30	-10.00

In Chapter 2, it was discussed that lip protrusion pushes the prominence roughly down to 2900 kHz by elongating the front cavity by 1 cm. Since the consonant itself is rounding neutral, the accompanying lip protrusion is a consequence of the spreading of the rounding feature corresponding to the surrounding rounded vowels. These results are consistent with what Boyce had observed in her study. She had seen a plateau pattern of rounding in utterances with two rounded vowels separated by an alveolar consonant and concluded that Turkish speakers tend to preserve their lip protrusion during the consonant until the end of the entire utterance.

English behaves very differently than Turkish. There is no consistent shift of prominence down to lower frequencies in English data as can be seen from Table 4.4.

Table 4.4: DFRNC for English males in rounded environment

utterances	KS	AM	KO
husu-1	6.30	-0.40	6.15
husu-2	-1.05	9.35	1.75
husu-3	5.10	9.65	-0.30
husu-4	-1.05	-0.45	1.85
husu-5	-4.60	0.85	2.75
losso-1	0.25	8.40	2.05
losso-2	5.20	2.45	10.30
losso-3	-3.25	13.30	6.70
losso-4	2.05	2.85	2.90
losso-5	-3.80	5.60	2.45
russo-1	-0.50	3.95	4.60
russo-2	-1.45	-1.80	16.10
russo-3	5.80	-5.10	9.40
russo-4	4.35	2.10	11.85
russo-5	8.95	-4.80	11.45

Across all utterances, the value is positive 73% of the time which suggests that in general the prominence is not shifted to lower frequencies through lip protrusion. The negative values seen occasionally are of small absolute value compared to others which implies that in these cases average spectra are flat and not dominated by any frequency region. Since it is not possible for articulators to change configurations instantaneously, it is natural to see some increase in energy of lower frequencies due to rounding feature of surrounding vowels. However, this effect is not significant enough to shift the frequencies of the major peaks of the consonant below 4kHz. Since rounding activity is minimized after the first rounded vowel, it will be reinitiated prior to the onset of the second rounded vowel. Thus, in English rounding activity for an utterance with two rounded vowels separated by a rounding neutral consonant will show a double peak pattern. Boyce had reported a double peak, "trough", pattern in the rounding activity in her analysis of EMG signals. The results of this study are fully consistent with what she had observed.To better compare and contrast data, the statistics are computed for each speaker in rounded vowel environment.

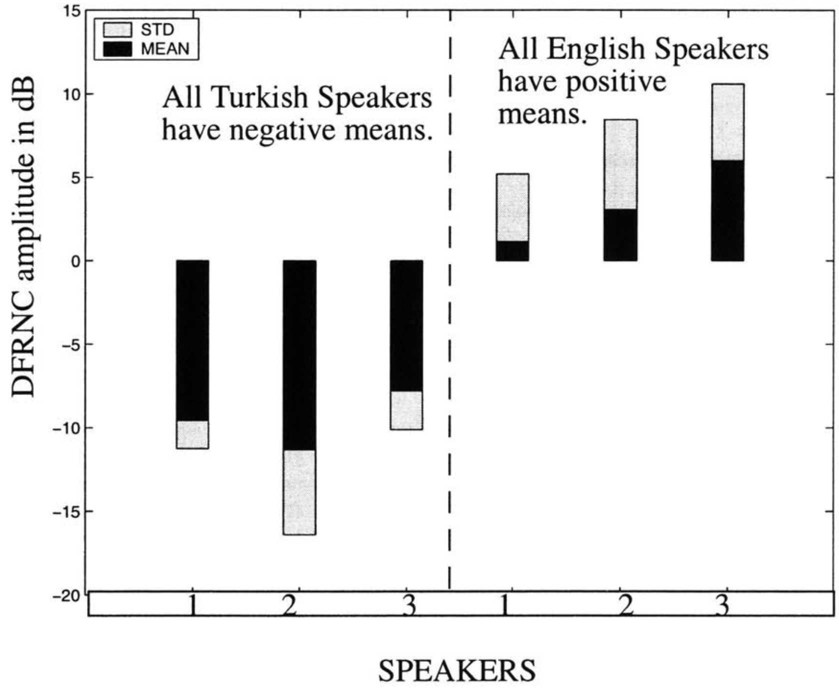

Figure 4.2: Statistics for rounded vowel environment.

Statistics for the rounded environment confirm that there is a remarkable difference in the behavior of Turkish and English speakers. Turkish speakers preserve rounding as they move onto the consonant in between two rounded vowels which shifts the prominence to frequencies lower than 4kHz and results in a negative value for DFRNC in rounded vowel environment. In English the prominence is above 4kHz in the spectra of /s/ in both rounded and unrounded environments since speakers terminate rounding during the consonant. Although the mean value for each speaker is significantly smaller in the rounded vowel environment, the high frequencies are still more prominent than low frequencies.

Boyce's study had shown the same difference in employment of labial coarticulation through an analysis of EMG signals. The results of this study suggest acoustic analysis is sufficient to detect this difference. To get a better idea of the effects of rounding on the spectrum of the consonant, a time analysis is used where the value of DFRNC is measured at three points throughout the duration of the consonant.

4.2 Time Analysis

This analysis is made on the utterances with /s/ in between two rounded vowels. The spectra of the consonant are not averaged this time. The value of the parameter DFRNC is measured at the beginning, the midpoint and the end of the consonant.The exact location of the measurements is discussed in Chapter 3. The measurements are referred to as DIFF_BEG, DIFF_MID and DIFF_END.

DIFF_BEG detects the effects of rounding spread from the first rounded vowel onto the consonant. A positive value indicates that there is no significant carryover of rounding from the preceding vowel that would have shifted major peaks to below 4kHz. A negative value, on the other hand, implies that there is some spread of rounding onto the neighboring consonant. DIFF_MID represents the value of DFRNC at the midpoint of the consonant. This location is the least likely to be affected by the rounding feature of the neighboring vowels. A negative value implies that lip protrusion is preserved through the entire duration of the consonant whereas a positive value implies rounding is terminated before the midpoint of the consonant. DIFF_END gives information about the location of the prominence in frication at the end of the consonant. A negative value means that there is significant anticipation of the +rounded feature of the upcoming vowel whereas a positive value suggests that the effects of the anticipation of rounding on the spectrum of the consonant are negligible. Thus DIFF_BEG and DIFF_END show two different types of spreading of features. The former one indicates carryover from the preceding rounded vowel, whereas the latter indicates the anticipation of the upcoming +rounded feature. Table 4.5, 4.6 and 4.7 show the results of this analysis for each Turkish male speaker individually.

Table 4.5: Results for Turkish Speaker 1

Turkish speaker CS	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	-6.75	-9.15	-7.80
rounded utterance 2	-8.60	-9.35	-17.20
rounded utterance 3	-8.15	-5.00	-10.10

Table 4.6: Results for Turkish speaker 2

Turkish speaker OO	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	-7.40	-8.75	-20.65
rounded utterance 2	-9.55	-10.05	-16.25
rounded utterance 3	6.15	6.25	-1.75

Table 4.7: Results for Turkish speaker 3

Turkish speaker OK	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	-8.40	-4.75	-16.30
rounded utterance 2	-10.40	-7.85	-14.65
rounded utterance 3	-7.85	-5.10	-8.00

For rounded vowel environments, the prominence is at frequencies lower than 4kHz even at the midpoint of the consonant (except for one case, the third utterance by speaker 2) confirming that lip protrusion is preserved throughout the entire duration of the consonant. The value of DIFF_BEG is not as negative as the value of DIFF_END, suggesting that the carryover of rounding from the first vowel is not as significant as the anticipation of the rounding corresponding to the upcoming vowel.

These results are consistent with the look-ahead model, i.e., the articulatory scanning mechanism moves toward a goal as early as possible as Boyce suggested before. Since there is another rounded vowel coming up, lip protrusion is not terminated at the end of the first rounded vowel.

Tables 4.8, 4.9 and 4.10 show how differently English speakers behave.

Table 4.8: Results for English Speaker 1

English speaker KS	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	0.95	0.95	-13.05
rounded utterance 2	0.95	7.40	-9.30
rounded utterance 3	-2.75	0.35	-4.90

Table 4.9: Results for English Speaker 2

English speaker AM	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	-4.55	3.40	-9.95
rounded utterance 2	8.75	8.10	0.20
rounded utterance 3	-12.90	2.35	9.85

Table 4.10: Results for English Speaker 3

English speaker KO	**DIFF_BEG (dB)**	**DIFF_MID (dB)**	**DIFF_END (dB)**
rounded utterance 1	8.60	13.15	-6.60
rounded utterance 2	-6.80	7.45	-8.40
rounded utterance 3	-4.60	8.55	-0.60

First of all, the value of DIFF_MID, which represents the prominence region at the midpoint of the consonant, is always positive. This confirms that although there may be some spreading of rounding from the consonant at V_1C boundary and CV_2 boundary, it is terminated by the time the speaker reaches the midpoint of the consonant. Thus, the trough pattern of rounding reported in the literature for English speakers is replicated in this study.

DIFF_BEG is negative half of the time, indicating that there may be some carryover of rounding onto the upcoming segments. However, there is no consistent pattern.

DIFF_END is negative most of the time, suggesting that there is anticipation of rounding feature corresponding to the second vowel in advance of the onset of that vowel. This is consistent with the time-locked model since it suggests that there may be some anticipation a specific time period ahead. That is, although the preceding consonant is compatible with lip rounding, the articulators will not move toward the rounding feature before that specific time window. To get a better representation of the variation in DFRNC throughout the consonant, DIFF_BEG, DIFF_MID and DIFF_END was averaged across all male speakers for the two languages. Figure 4.3 shows the average values of these parameters in the two languages.

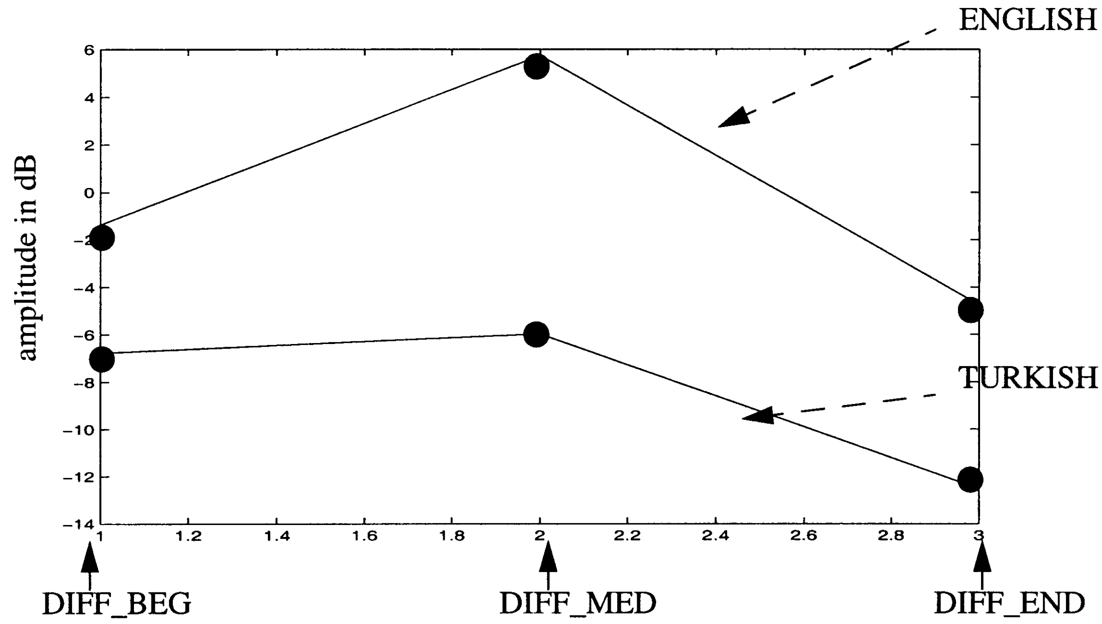

Figure 4.3: Average Values for the Parameters Used for the Time Analysis

4.3 Summary of Results and Discussion

The results of the acoustic analysis on average spectra of male data supports Boyce's findings from EMG signals and measurements of lip protrusion. The variation of DFRNC in Turkish and English confirms that the first has a "plateau" and the second a "trough" pattern of rounding in production of two rounded vowels separated by a non-labial consonant. Acoustic analysis was powerful enough to detect the variation in labial coarticulatory organization of the two languages.

The second type of analysis gave some idea about the time course of rounding. The comparison of the location of peaks in the midpoint of frication sheds light onto whether rounding is preserved through the entire duration of the consonant. In both languages, there seems to be some anticipation of the upcoming rounding feature. Both the time-locked and the look-ahead model acknowledge that anticipation. The difference between the models is in the mechanism behind the anticipation. Look-ahead model suggests a scanning mechanism that moves toward a goal as early as possible, whereas the time-locked model suggests that anticipation is context-independent and is not initiated before a specific time period prior to the onset of the feature to be anticipated. Turkish is more consistent with look ahead since the rounding feature is spread onto the entire duration of the consonant. English is in harmony with time locked model since the +rounded feature of the adjacent vowels affects the spectra only at the endpoints, as the time analysis show.

4.4 Results for female data

Three female Turkish and three female English speakers participated in this experiment. As in analysis of male data, the average spectra were studied initially. However, this time a boundary of 5kHz is used to divide the spectra into high and low frequency regions. A time analysis identical to the one explained before was conducted on female data. The results of the average spectra analysis is discussed first.

Table 4.11 shows the results of DFRNC in unrounded environment for Turkish females. The value of DFRNC is positive most of the time. There are only five cases out of fifteen in which the value is negative.

Table 4.11: DFRNC in Unrounded Environment for Turkish

utterances	Subject BD	Subject ID	Subject HO
hasan-1	6.40	9.60	9.30
hasan-2	3.70	1.25	4.00
hasan-3	4.70	7.95	0.10
hasan-4	5.30	3.80	0.25
hasan-5	7.70	8.75	1.00
karasin-1	14.50	7.50	6.45
karasin-2	8.95	-4.85	-2.85
karasin-3	9.50	3.80	-2.50
karasin-4	10.50	10.20	5.10
karasin-5	8.00	6.40	-2.95
keser-1	5.00	8.30	8.25
keser-2	1.15	9.40	-5.65
keser-3	8.10	5.75	1.90
keser-4	5.90	8.40	6.00
keser-5	3.90	-1.95	2.35

Table 4.12 shows the results for English.

Table 4.12: DFRNC in Unrounded Environment in English

utterances	subject MC	subject MM	subject KP
bessy-1	12.70	4.10	7.15
bessy-2	14.30	8.70	12.35
bessy-3	17.75	10.15	10.40
bessy-4	15.85	7.15	11.35
bessy-5	16.05	8.95	11.85
lissa-1	14.30	5.75	7.65

Table 4.12: DFRNC in Unrounded Environment in English

utterances	subject MC	subject MM	subject KP
lissa-2	15.25	-2.00	10.10
lissa-3	15.60	8.10	5.70
lissa-4	12.05	5.65	15.35
lissa-5	16.75	3.15	13.10
tissi-1	12.55	4.15	12.30
tissi-2	18.35	3.25	11.80
tissi-3	17.90	4.75	6.45
tissi-4	14.55	4.60	11.60
tissi-5	13.20	5.50	17.20

The prominence in the spectrum of frication is above 5kHz consistently. As expected, the two languages display similar results in unrounded environment which is included to serve as a control. Figure 4.3 shows the statistics for unrounded vowel environment. The mean and standard deviation are calculated for each speaker.

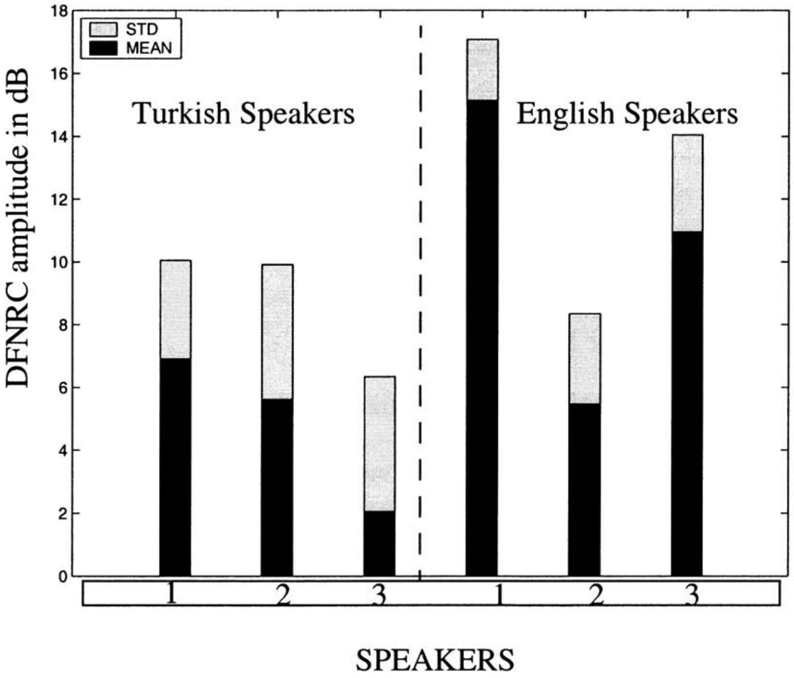

Figure 4.4: Statistics for Females in unrounded environment.

In an unrounded environment, the mean value for DFRNC is positive for each speaker. The mean values are higher for English speakers than the values for Turkish speakers. This indicates that spectra for Turkish speakers tend to be flatter, whereas the spectra for English speakers are dominated by more prominent peaks in the high frequency region.

The variation of labial coarticulatory organization is expected to appear on in the rounded vowel environment.Table 4.13 shows the results from the Turkish data in rounded vowel environment. The value is consistently negative, i.e the prominence is shifted to below 5kHz through lip protrusion. However, the absolute value of results is much smaller compared to the results of male data in this environment, suggesting that the spectra are flatter than before. In male data, there was more consistency across subjects. Here all of the subjects have major peaks shifted down in frequency, but the prominences of the peaks vary more than they did in the male data.

Table 4.13: DFRNC in Rounded Environment for Turkish females

utterances	Subject BD	Subject ID	Subject HO
korusun-1	**0.30**	**-1.10**	**-7.75**
korusun-2	**-1.20**	**-5.30**	**-10.40**
korusun-3	**0.10**	**-5.05**	**-7.20**
korusun-4	**-0.30**	**-3.00**	**-8.90**
korusun-5	**1.45**	**-2.35**	**-11.15**
kusur-1	**-1.80**	**-5.05**	**-15.45**
kusur-2	**0.15**	**-4.00**	**-10.95**
kusur-3	**-1.65**	**-2.75**	**-16.40**
kusur-4	**0.20**	**0.40**	**-17.40**
kusur-5	**-1.95**	**-8.40**	**-12.45**
tosun-1	**-4.45**	**4.55**	**-5.60**
tosun-2	**-0.60**	**-0.35**	**-9.50**
tosun-3	**-5.70**	**-0.30**	**-13.50**
tosun-4	**-6.60**	**-0.40**	**-9.45**

Table 4.13: DFRNC in Rounded Environment for Turkish females

utterances	Subject BD	Subject ID	Subject HO
tosun-5	-6.90	-2.20	-8.45

Table 4.14 shows the results from the English data in this environment. DFRNC takes positive values that are comparable with the values it had taken in the unrounded case. That is, the spectra of /s/ are not affected significantly by the rounding feature of the surrounding vowels. In other words, the rounding feature is not carried onto /s/ but terminated at the end of the first vowel and reinitiated at the beginning of the second vowel.

Table 4.14: DFRNC in Rounded Environment for English females

utterances	subject MC	subject MM	subject KP
husu-1	12.40	4.10	10.25
husu-2	12.95	3.10	9.90
husu-3	11.05	-0.65	-6.05
husu-4	13.95	1.90	9.80
husu-5	13.45	4.65	7.55
losso-1	8.50	5.05	10.10
losso-2	11.90	5.40	13.40
losso-3	4.60	11.45	7.65
losso-4	18.60	15.05	9.05
losso-5	17.15	7.90	14.55
russo-1	13.50	11.05	-0.70
russo-2	11.35	8.10	-8.10
russo-3	7.35	13.70	13.60
russo-4	9.10	9.70	21.90
russo-5	6.30	9.25	14.70

Figure 4.4 shows the statistics for rounded vowel environment.

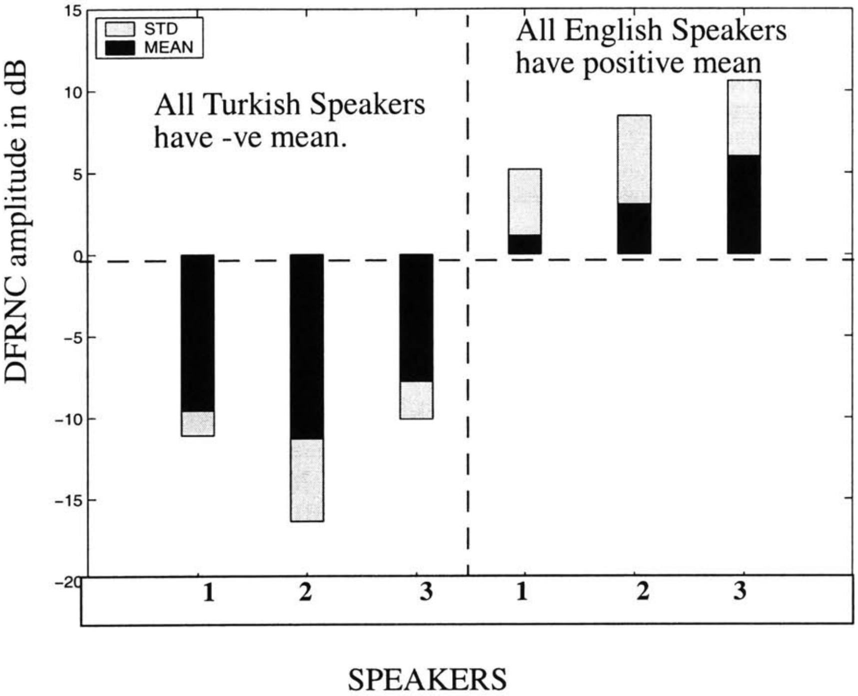

Figure 4.5: Statistics for Rounded Environment for Females

In summary, as in male data, there is a deviation in the behavior of the speakers of the two languages in the rounded vowel environment. Turkish speakers preserve lip rounding, which shifts the prominence in frication to lower frequencies and results in negative values for DFRNC. English speakers, on the other hand, terminate lip protrusion during the consonant and therefore have positive values for DFRNC in both rounded and unrounded environments.

4.5 Time Analysis

To get a better idea of how rounding affects the spectrum of the consonant, DFRNC was measured at three points through the duration of the consonant. Table 4.15-4.17 show the results for Turkish females.

Table 4.15: Turkish Female Speaker 1

Turkish speaker BD	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-1.70	3.10	-5.55
rounded utterance 2	-6.95	-2.75	-11.00
rounded utterance 3	-5.05	-2.35	-6.70

Table 4.16: Turkish Female Speaker 2

Turkish speaker ID	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-1.40	-2.50	-5.90
rounded utterance 2	-4.55	-3.60	-10.40
rounded utterance 3	14.70	10.00	-1.25

Table 4.17: Turkish Female Speaker 3

Turkish speaker HO	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-1.35	-8.10	-8.85
rounded utterance 2	-16.10	-15.10	-14.80
rounded utterance 3	-9.00	-5.25	-13.00

DIFF_MID is negative in seven out of nine cases suggesting that rounding is preserved through the entire duration of the consonant. DIFF_BEG and DIFF_END measure effects of rounding on the consonant in close vicinity of the rounded vowels. They are negative in general. In the male data, DIFF_BEG was consistently of a greater absolute value than DIFF_END. Such a pattern is not evident in Turkish female data. Table 4.18-4.20 display the results for English speakers.

Table 4.18: English Female Speaker 1

English speaker MC	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-4.75	11.15	7.25
rounded utterance 2	-9.95	5.10	5.85
rounded utterance 3	-8.85	4.95	8.95

Table 4.19: English Female Speaker 2

English speaker MM	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-2.15	10.00	-3.20
rounded utterance 2	-3.65	2.65	-7.00
rounded utterance 3	-0.50	9.70	1.50

Table 4.20: English Female Speaker 3

English speaker KP	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	2.55	9.80	-1.30
rounded utterance 2	-5.05	11.30	6.25
rounded utterance 3	-9.45	6.75	3.00

DIFF_MID is positive in all cases suggesting that the lip protrusion is terminated by the midpoint of the consonant. DIFF_BEG and DIFF_END take positive and negative values interchangeably, i.e., they do not have a consistent pattern. In English male data, it was observed that the absolute value of DIFF_END was greater than that of DIFF_BEG, suggesting that the effects of the anticipation of rounding feature prior to a rounded vowel is greater than that of the carryover of rounding from a preceding rounded vowel. With the female speakers of this study, this pattern does not hold.

To see the difference in behavior of the two languages, DIFF_BEG, DIFF_MID and DIFF_END were averaged across all speakers. Figure 4.6 shows the average values of these parameters for Turkish and English.

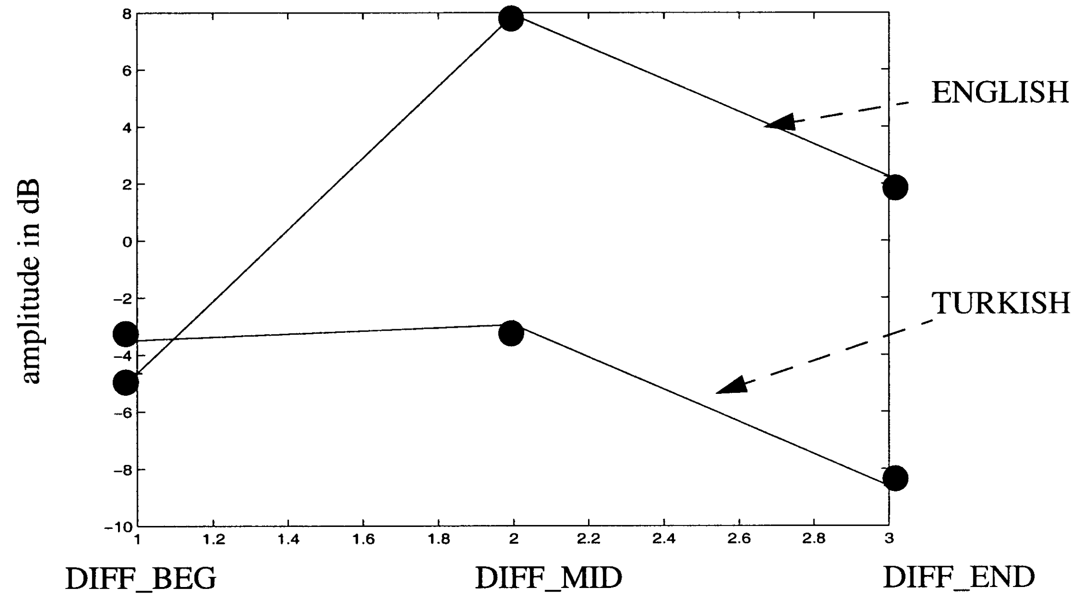

Figure 4.6: Average Values of the Parameters in Time Analysis for the Two Languages.

The results of the acoustic analysis on average spectra of female data support Boyce's findings from EMG signals generated by labial movements. As before, acoustic analysis was powerful enough to detect the differences in labial coarticulatory organization of the two languages.

Chapter 5

RESULTS FOR UTTERANCES WITH /Z/

5.1 Average Spectra Analysis For Males

The production of /z/ is almost identical to that of /s/ except for the accompanying voicing. The expectation is a coarticulatory pattern similar to the one in /s/ utterances since the voicing is not likely to impact the coarticulation mechanism employed by the speaker. As before, initially DFRNC is calculated for utterances with unrounded vowels as a control. Table 5.1 and 5.2 display the values this parameter takes in the unrounded environment for Turkish and English respectively.

Table 5.1: DFRNC for unrounded vowel environment in Turkish, male speakers

utterances	Speaker CS	Speaker OO	Speaker OK
bezek1	9.10	5.70	5.90
bezek2	4.85	-0.95	2.60
bezek3	5.95	6.55	7.10
bezek4	4.50	3.15	10.40
bezek5	3.35	0.85	4.85
ezin1	5.55	0.25	8.35
ezin2	4.85	-8.00	9.60
ezin3	7.75	6.40	3.95
ezin4	7.45	-3.85	12.65
ezin5	6.50	-0.95	5.45
kazi1	6.75	4.25	6.60
kazi2	0.10	7.85	14.40
kazi3	2.75	1.10	5.50
kazi4	-1.45	4.10	5.75
kazi5	11.35	4.30	7.75

DFRNC is positive most of the time with occasional negative values of small absolute value. This is what was observed in Turkish /s/ utterances. Since the surrounding vowels are unrounded, there is no reason for the speaker to protrude lips at any point in the utterance. The prominence in the frication is at high frequencies. English behaves similarly since in the unrounded vowel environment, the differences in labial coarticulatory organization of the two languages do not come onto surface.

Table 5.2: DFRNC for unrounded vowel environment in English, male speakers

utterances	KS	AM	KO
buzzy1	**5.65**	**-7.60**	**10.1**
buzzy2	**12.40**	**-6.15**	**2.90**
buzzy3	**5.90**	**-6.20**	**-0.25**
buzzy4	**2.05**	**4.20**	**14.75**
buzzy5	**7.75**	**-3.85**	**8.70**
lizzy1	**1.65**	**9.10**	**7.50**
lizzy2	**14.30**	**3.40**	**4.70**
lizzy3	**9.70**	**4.95**	**5.20**
lizzy4	**8.40**	**4.60**	**6.65**
lizzy5	**13.20**	**6.05**	**-3.35**
tazzi1	**13.45**	**-0.70**	**12.60**
tazzi2	**10.55**	**-2.70**	**9.55**
tazzi3	**16.60**	**6.40**	**11.65**
tazzi4	**9.75**	**5.35**	**5.35**
tazzi5	**7.85**	**4.65**	**5.75**

In English, DFRNC takes a positive value in 37 out of 45 cases. A positive DFRNC implies that the prominence in spectra of the frication is above 4kHz. The greater the value of DFRNC, the more significant the prominence is. As in /s/ utterances, the two languages behave similarly in an unrounded vowel environment and have energy concentrated at higher frequencies.

To get a better representation of the variation of DFRNC across the subjects, statistics are calculated. Figure 5.1 displays the mean and the standard deviation for each subject individually in an unrounded vowel environment. Once the range of values of DFRNC in an unrounded vowel setting is known, it can be determined whether lip rounding is spread into the consonant in rounded vowel environment by comparing the values of DFRNC in a rounded environment to the values in the unrounded environment.

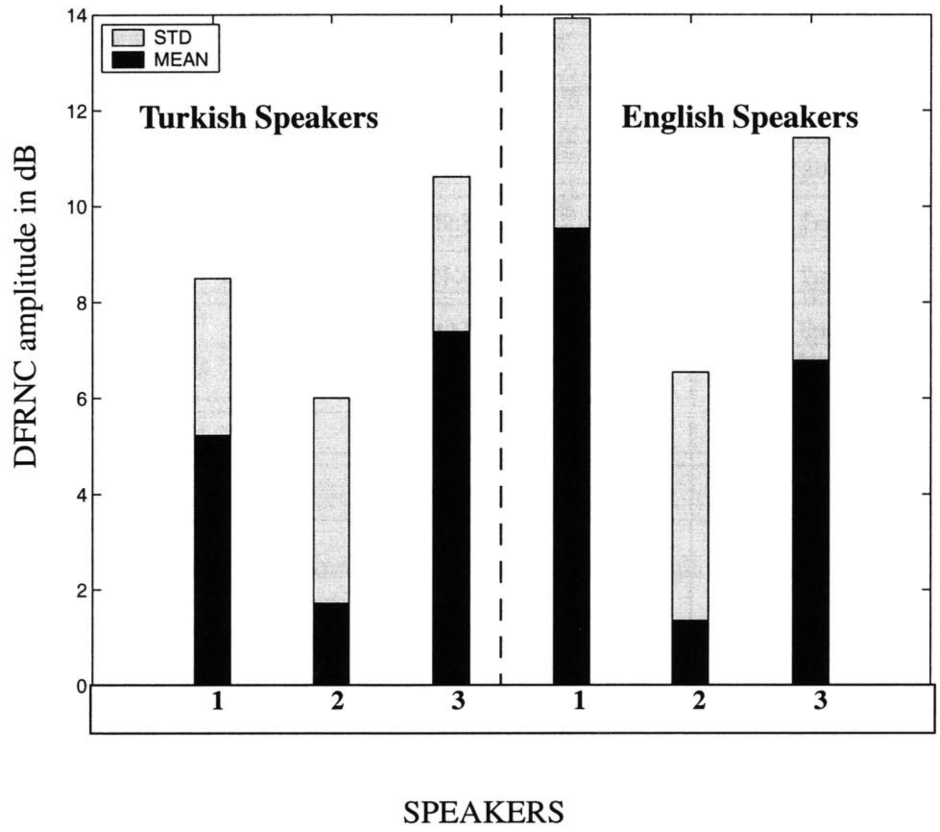

Figure 5.1: Statistics for Unrounded Vowel Environment

As seen from the figure, both languages have positive means and comparable variation in the results in this environment. For all subjects, the mean value is smaller than the mean value for /s/ utterances in unrounded environment which is expected. DFRNC represents the energy at high frequency region compared to the energy at low frequency region. Since /z/ is voiced whereas /s/ is not, there is more energy located at low frequency region in spectra of /z/ than that of /s/. Thus the difference of the energies located at high and low frequency regions is smaller in /z/ utterances.

The variation in the coarticulatory organization of the two languages come to surface in the rounded vowel environment. Table 4.3 shows the results for Turkish.

Table 5.3: DFRNC for rounded vowel environment in Turkish, male speakers

utterances	**CS**	**OO**	**OK**
bozuk1	**-20.30**	**-10.10**	**-12.65**
bozuk2	**-18.15**	**-12.95**	**-5.25**
bozuk3	**-19.20**	**-12.20**	**-2.45**
bozuk4	**-10.90**	**-6.00**	**-7.25**
bozuk5	**-16.60**	**-7.30**	**-7.95**
kuzu1	**-12.20**	**-5.90**	**-7.35**
kuzu2	**-16.50**	**-9.45**	**-10.85**
kuzu3	**-14.10**	**-14.00**	**-4.90**
kuzu4	**-18.10**	**-13.50**	**-4.55**
kuzu5	**-11.55**	**-6.95**	**-7.95**
ozon1	**-17.85**	**-4.05**	**-10.20**
ozon2	**-15.95**	**-6.30**	**-7.85**
ozon3	**-9.15**	**-7.00**	**-3.65**
ozon4	**-14.55**	**-14.10**	**-6.15**
ozon5	**-11.00**	**-12.25**	**-8.55**

For each speaker, the prominence in spectra of /z/ is shifted to lower frequencies 100% of the time. As in /s/ utterances, lip protrusion is preserved during the production of the consonant. The results support the “plateau” pattern of rounding in production of two rounded vowels separated by a non-labial consonant which was reported earlier by Boyce. Table 5.4 displays the results for English.

Table 5.4: DFRNC for rounded vowel environment in English, male speakers

utterances	KS	AM	KO
bozzo1	-3.25	-4.90	-1.10
bozzo2	-0.40	1.05	15.15
bozzo3	-8.65	0.90	0.35
bozzo4	-1.30	-10.25	7.95
bozzo5	-1.05	-1.25	-0.30
kuzzu1	-4.95	8.20	3.45
kuzzu2	-1.40	0.30	4.05
kuzzu3	-5.10	7.60	0.70
kuzzu4	-2.45	10.70	4.15
kuzzu5	1.35	-5.10	2.65
tozzu1	1.55	-7.95	2.40
tozzu2	6.60	6.70	0.80
tozzu3	1.45	-4.30	n/a
tozzu4	-1.70	6.75	1.85
tozzu5	-1.25	2.40	2.90

There is a lot of variation between speakers. Speaker 1 has a prominence below 4kHz 73% of the time whereas speaker 2 has a prominence at high frequencies 60% and speaker 3, 87% of the time. Although in general the results for /z/ utterances in rounded vowel environment are consistent with the hypotheses, they are not as clear as the results for /s/ utterances. One possible reason is that the duration of /z/ is much shorter than the duration of /s/ for all subjects. Table 5.5 displays the average duration for /z/ versus /s/ for each male subject over all utterances with these consonants. It is known that there will be some anticipation of rounding prior to the onset of the vowel no matter what pattern of labial coarticulation the speaker utilizes. In /s/ utterances, this anticipation was of very short duration compared to the entire duration of the consonant in English, so when the spectra were averaged, it did not change the spectra significantly. The average spectra were only affected if the subject preserved lip rounding during the entire duration of the consonant. However, since the duration of /z/ is much shorter than the duration of /s/, the anticipation of rounding at the two ends of the consonant has a greater effect when the spectrum is averaged. Thus, we see more negative values for English /z/ utterances than English /s/ utterances. The duration of /z/ is shorter than the duration of /s/ in Turkish, too. However since speakers preserve lip protrusion during the entire duration of both consonants, the difference in average duration does not affect the results.

Table 5.5: Comparison of the Average Durations of /s/ and /z/ for each subject

Subjects	**Average Duration of /z/ (ms)**	**Average Duration of /s/ (ms)**
English Subject KS	**70.43**	**110.80**
English Subject AM	**71.80**	**120.57**
English Subject KO	**52.33**	**76.70**
Turkish Subject CS	**51.00**	**66.43**
Turkish Subject OO	**38.83**	**69.43**
Turkish Subject OK	**56.20**	**80.50**

Figure 5.2 displays the statistics for each speaker in the rounded vowel environment.The mean value is negative for all Turkish subjects whereas it is positive for two out of three English subjects. Standard deviations are less for all Turkish subjects than English subjects. Clearly, rounding affects the spectra of /z/ more than it did the spectra of /s/ for reasons discussed earlier.

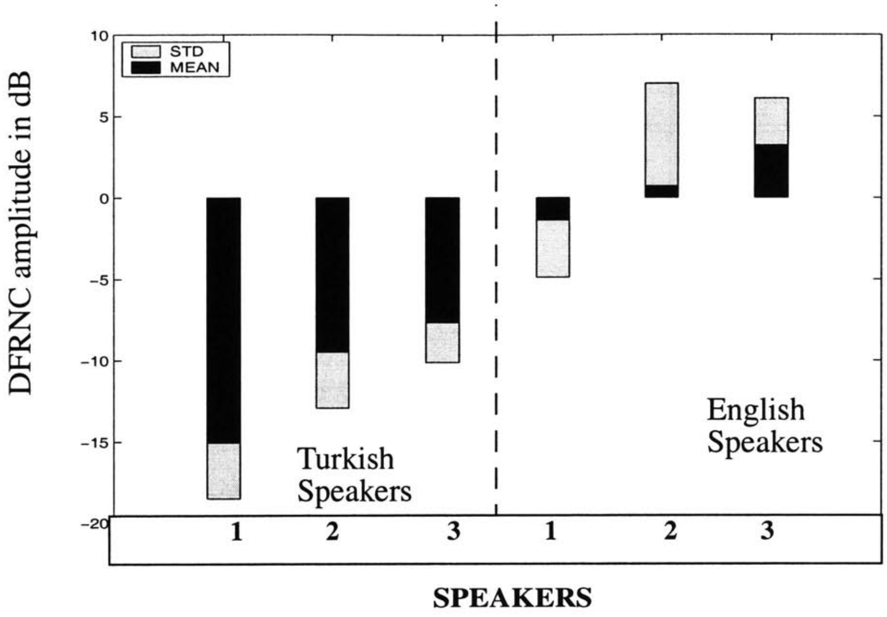

Figure 5.2: Statistics for Rounded vowel Environment

5.2 Time Analysis

To get an idea about how the rounding feature of the adjacent vowels affect the spectrum of an alveolar consonant throughout its duration, a time analysis was conducted where the value of DFNRC was measured at the onset, the midpoint and the end of the consonant.

The value at the onset is referred to as DIFF_BEG, the midpoint as DIFF_MID and the end as DIFF_END as before. Table 5.6, 5.7 and 5.8 show the results of this analysis for Turkish male speakers.

Table 5.6: Variation of DFRNC for Turkish Subject 1

Turkish speaker CS	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-17.20	-15.65	-18.50
rounded utterance 2	-14.75	-8.90	-17.90
rounded utterance 3	-10.75	-8.45	-15.85

Table 5.7: Variation of DFRNC for Turkish Subject 2

Turkish speaker OO	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-6.10	-7.00	-15.20
rounded utterance 2	-3.45	-8.10	-13.70
rounded utterance 3	-6.90	-7.50	-6.05

Table 5.8: Variation of DFRNC for Turkish Subject 3

Turkish speaker OK	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-8.05	-12.85	-10.45
rounded utterance 2	-12.40	-13.95	-11.65
rounded utterance 3	-13.60	-16.75	-20.00

The value of DFRNC is negative at everywhere even at the midpoint of the consonant which is the location that is the least likely to be affected from rounding. This is what we would expect to see for look ahead model of coarticulation since the model suggests that subjects preserve their lip protrusion throughout the entire consonant and have prominence at low frequencies in their spectra.

Tables 5.9-5.11 display the results for English. English subjects behave significantly different from the Turkish subjects.

Table 5.9: Variation of DFRNC for English Subject 1

English speaker KS	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-5.00	1.85	-9.50
rounded utterance 2	-4.30	1.25	-3.10
rounded utterance 3	-1.7	5.90	-5.20

Table 5.10: Variation of DFRNC for English Subject 2

English speaker AM	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-7.75	3.25	-3.70
rounded utterance 2	-5.70	9.35	-0.35
rounded utterance 3	-1.85	2.40	-14.05

Table 5.11: Variation of DFRNC for English Subject 3

English speaker KO	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-14.75	4.10	-2.20
rounded utterance 2	-7.65	18.05	-3.35
rounded utterance 3	-1.70	4.65	-12.55

The results of the time analysis are clearer than those of the average spectra analysis. Since no averaging is done, the anticipation in the close vicinity of the vowels do not impact the entire duration of the consonant. The value of DIFF_MID is always positive, implying that rounding is terminated by the time the midpoint of the consonant is reached. The rounding feature of the vowels shift the prominence to lower frequencies at the two ends of the consonant so DIFF_BEG and DIFF_END are always negative. The trough pattern of rounding in production of two rounded vowels separated by an alveolar consonant reported in the literature for English speakers is replicated in the /z/ utterances of this study, too.

To see the difference in variation of DFRNC throughout time clearly, average values of DIFF_BEG, DIFF_MID and DIFF_END are computed and the results are summarized in Figure 5.3.

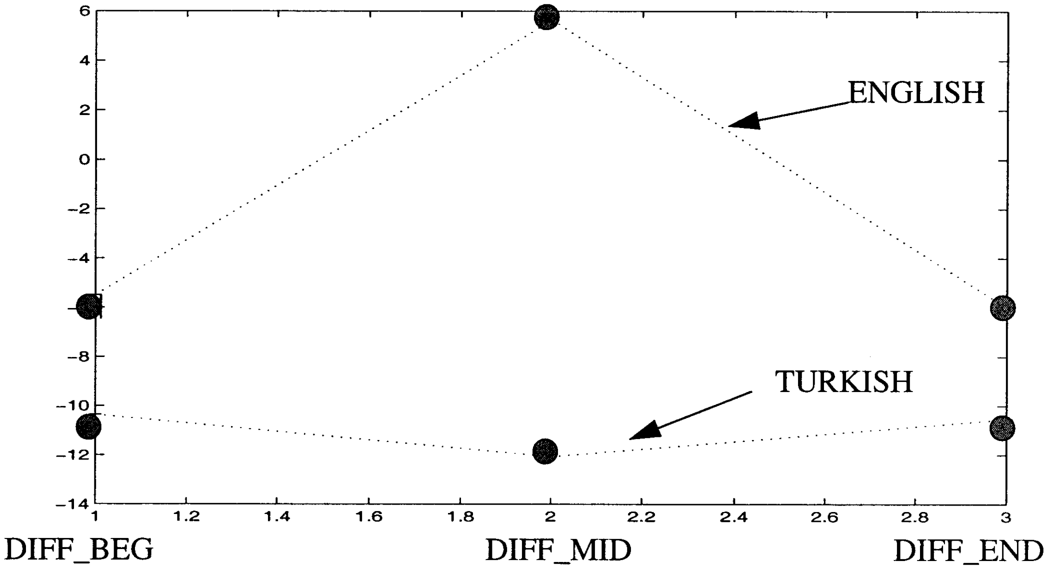

Figure 5.3: Average values of Parameters used in Time Analysis

5.3 Summary of Results and Discussion

The results of the acoustic analysis on average spectra of male data for /z/ utterances support Boyce's findings from EMG signals generated by labial movements. The variation in the range of values DFRNC takes in Turkish and English are consistent with the hypothesis that the first has a “plateau” and the second a “trough” pattern of rounding in production of two rounded vowels separated by a non-labial consonant. Acoustic analysis was powerful enough to detect the variation in labial coarticulatory organization of the two languages. In English, the anticipation of the +rounded feature of the adjacent vowels at the two ends of the consonant affected the average spectra of /z/ utterances more significantly than it did those of /s/ utterances since the duration of /z/ was consistently shorter than /s/.

The second type of analysis gave some idea about the time course of rounding. DIFF_MID was always positive for English and negative for Turkish which confirms that rounding is preserved in Turkish and terminated in English.

5.4 Results for female subjects

The average spectra were studied initially in female data. A boundary of 5kHz was used to divide the spectra into high and low frequency regions. The value of DFRNC was measured at utterances with unrounded vowels to serve as controls as before. Table 5.12 shows the results for Turkish females.

Table 5.12: DFRNC in unrounded vowel environment for Turkish females

utterances	BD	ID	HO
bezek1	-1.15	6.85	1.85
bezek2	6.15	7.00	4.05
bezek3	12.00	7.45	7.25
bezek4	6.35	3.60	-2.85
bezek5	4.45	6.70	2.90
ezin1	5.95	8.35	3.20
ezin2	4.80	10.30	-0.45
ezin3	6.35	5.45	6.95
ezin4	-2.20	8.70	0.50

Table 5.12: DFRNC in unrounded vowel environment for Turkish females

utterances	BD	ID	HO
ezin5	4.45	10.75	-1.10
kazi1	4.25	4.50	4.00
kazi2	9.50	-2.60	3.05
kazi3	-7.80	9.10	4.95
kazi4	5.70	5.70	-1.40
kazi5	2.00	8.50	-2.20

As in /s/ utterances in this environment, DFRNC takes positive values with a few exceptions. English is expected to behave similarly in this vowel environment. Table 5.13 shows the results for English females.The prominence in spectra of fricative is above 5kHz consistently. Since there are no rounded vowels, the difference in coarticulatory organization of the two languages does not show up.

Table 5.13: DFRNC in unrounded vowel environment for English females

utterances	MC	MM	KP
buzzy1	8.90	9.70	6.65
buzzy2	15.15	5.50	7.60
buzzy3	18.15	7.55	8.00
buzzy4	15.50	4.80	4.65
buzzy5	17.60	3.75	12.90
lizzy1	15.75	4.75	7.05
lizzy2	18.05	6.65	-0.95
lizzy3	13.90	5.60	6.15
lizzy4	15.05	9.90	10.80
lizzy5	13.90	10.40	4.95
tazzi1	14.30	-0.90	7.15
tazzi2	17.60	6.40	6.35
tazzi3	17.25	5.60	12.40
tazzi4	12.50	11.55	11.40
tazzi5	11.15	9.00	14.95

Figure 5.4 shows the statistics for this environment. The mean is positive for all speakers.The values are smaller in Turkish than the mean values for male subjects. In English they are comparable.

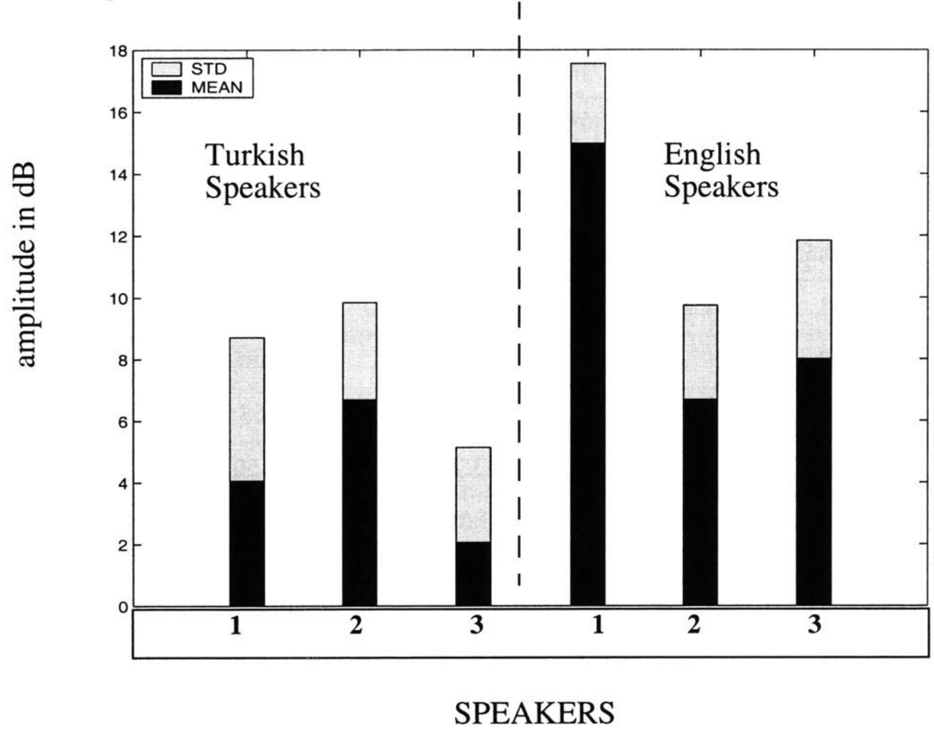

Figure 5.4: Statistics for Females in Unrounded Vowel Environment

The difference in coarticulatory patterns of the two languages comes up in rounded vowel environment. Table 5.14 shows the results for Turkish. Turkish females preserve their lip protrusion, so have the prominence at much lower frequencies than in the case of unrounded vowel environment. English females behave totally differently and have prominence at high frequencies at both rounded and unrounded vowel environments. Table 5.15 shows the results for English female subjects.

Table 5.14: DFRNC in Rounded Environment for Turkish

utterances	BD	ID	HO
bozuk1	**-9.75**	**-5.85**	**-19.90**
bozuk2	**-6.85**	**n/a**	**-11.30**
bozuk3	**-11.20**	**-4.00**	**-16.30**
bozuk4	**-3.15**	**-2.70**	**-12.90**
bozuk5	**-9.85**	**-7.25**	**-10.05**
kuzu1	**-2.85**	**-1.35**	**-10.15**
kuzu2	**-0.90**	**-4.00**	**-10.85**
kuzu3	**-9.65**	**-7.05**	**-12.65**
kuzu4	**-7.30**	**-3.70**	**-12.50**
kuzu5	**-3.60**	**-3.00**	**-13.80**
ozon1	**-16.35**	**10.45**	**-19.30**
ozon2	**-7.75**	**-1.75**	**-14.85**
ozon3	**-13.80**	**-4.20**	**-12.31**
ozon4	**-8.30**	**2.65**	**-15.60**
ozon5	**-8.15**	**-3.80**	**-18.85**

Table 5.15: DFRNC in Rounded Vowel Environment for English

utterances	MC	MM	KP
bozzo1	**8.50**	**11.95**	**9.75**
bozzo2	**13.05**	**6.15**	**5.75**
bozzo3	**17.45**	**7.10**	**4.05**

Table 5.15: DFRNC in Rounded Vowel Environment for English

utterances	MC	MM	KP
bozzo4	12.60	9.25	11.25
bozzo5	13.35	6.75	17.35
kuzzu1	14.55	8.60	9.45
kuzzu2	15.45	6.50	9.15
kuzzu3	20.60	5.95	-1.00
kuzzu4	15.75	6.45	7.55
kuzzu5	12.20	3.35	10.65
tozzu1	6.75	13.70	10.35
tozzu2	18.20	5.35	10.90
tozzu3	11.60	10.10	0.10
tozzu4	8.75	9.05	5.45
tozzu5	2.95	5.75	12.40

To get a better representation of the difference in the behavior of the two languages, mean and standard deviation is calculated for each subject. The mean value is positive for every Turkish subject and negative for every English subject as expected. In terms of the standard deviation from the mean, the results of the two languages are comparable, too. Figure 5.5 displays the statistics for rounded vowel environment.

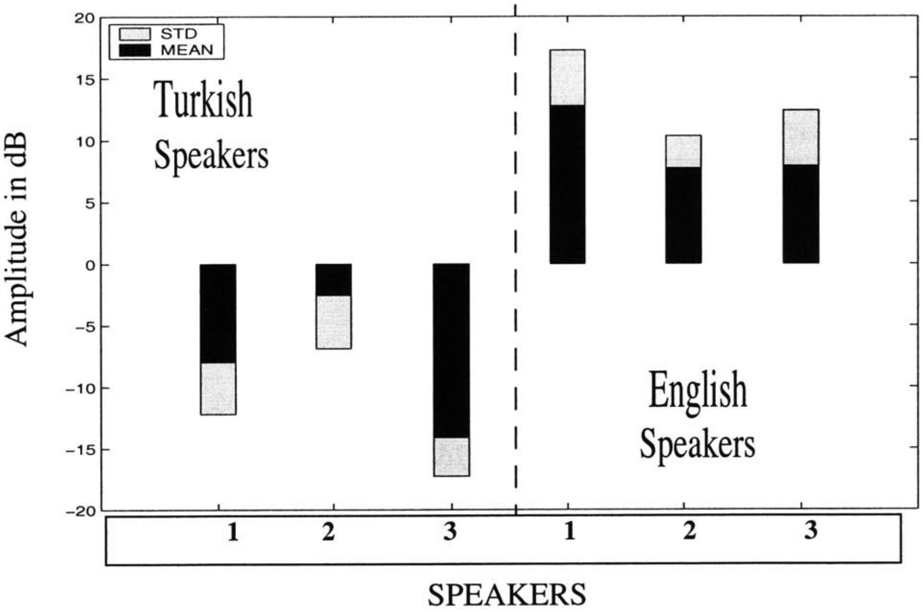

Figure 5.5: Statistics for Rounded Vowel Environment

5.5 Time Analysis

As before, a time analysis is conducted to get a better idea of the effects of rounding on spectra. Tables 5.16, 5.17 and 5.18 display the results for each Turkish female speaker

Table 5.16: Time variation of DFRNC for Turkish speaker 1

Turkish speaker BD	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-10.80	-5.75	-1.05
rounded utterance 2	-3.45	-7.90	-5.85
rounded utterance 3	-10.50	-14.15	-10.00

Table 5.17: Time variation of DFRNC for Turkish speaker 2

Turkish speaker ID	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-5.30	-5.55	-10.15
rounded utterance 2	-6.00	-2.45	-2.65

Table 5.17: Time variation of DFRNC for Turkish speaker 2

Turkish speaker ID	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 3	-7.15	7.85	-4.10

Table 5.18: Time variation of DFRNC for Turkish speaker 3

Turkish speaker HO	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-9.55	-12.55	-20.25
rounded utterance 2	-9.40	-6.35	-9.85
rounded utterance 3	-23.55	-9.70	-15.70

DFRNC is negative even at the midpoint of the consonant confirming that Turkish displays a plateau pattern of labial activity in production of two rounded vowels separated by an alveolar consonant.Tables 5.19-5.21 show how differently English females behave.

Table 5.19: Time variation of DFRNC for English speaker 1

English speaker MC	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-15.15	5.30	-5.00
rounded utterance 2	-3.30	15.20	5.35
rounded utterance 3	-7.80	19.20	-3.50

Table 5.20: Time variation of DFRNC for English speaker 2

English speaker MM	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-3.70	14.45	-1.40
rounded utterance 2	0.65	7.20	-1.25
rounded utterance 3	1.15	16.3	1.05

Table 5.21: Time variation of DFRNC for English speaker 3

English speaker KP	DIFF_BEG (dB)	DIFF_MID (dB)	DIFF_END (dB)
rounded utterance 1	-1.05	6.65	-3.85
rounded utterance 2	-1.50	16.55	-3.85
rounded utterance 3	-5.45	2.25	-13.40

In all three speakers, the value of DFRNC increases significantly at the midpoint of the consonant's spectra. Since the speaker terminates rounding at the end of the first rounded vowel, the prominence moves to high frequencies by the time the midpoint is reached. DIFF_BEG and DIFF_END take negative values most of the time since the +rounded feature of the vowels impact the spectra in the close vicinity and shift the prominence down to lower frequencies.

As before, the values at three measurement points used were averaged to get a better comparison of the behavior of the two languages.

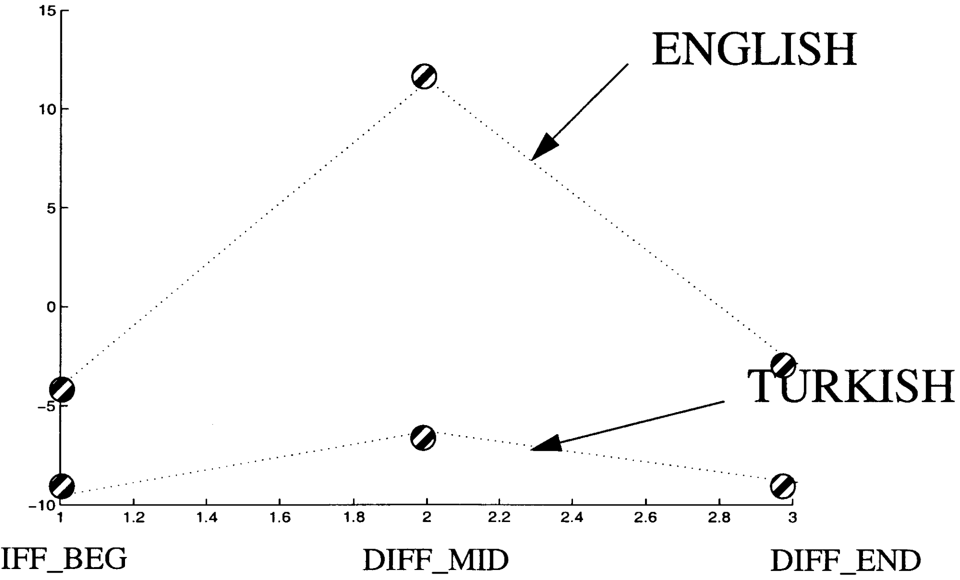

Figure 5.6: Average values for the parameters used in the time analysis

5.6 Summary and Discussion

There are no differences in the behavior of female and male speakers. The results of the acoustic analysis on average spectra of female /z/ data supports Boyce's findings from EMG signals generated by labial movements. The variation of DFRNC in Turkish and English female data confirms that the first has a "plateau" and the second a "trough" pattern of rounding in production of two rounded vowels separated by a non-labial consonant. As before, acoustic analysis was powerful enough to detect the differences in labial coarticulatory organization of the two languages.

Chapter 6

RESULTS FOR /T/ AND /SH/

Acoustic analysis confirmed Boyce's findings for utterances containing /z/ and /s/. The next step was to check whether the same coarticulatory patterns hold for utterances with the stop /t/ and palatoalveolar /sh/ as well. The analysis for these two consonants was done on data from eight speakers. In the previous parts, analyses for male and female speakers were conducted separately. No differences were observed in the results, so there was no advantage for doing the analyses separately. Analyses of data from both genders is done simultaneously in this part.

6.1 Analysis of /t/ utterances

The initial burst was averaged over 15ms avoiding the aspiration as explained in Chapter 3. A boundary of 4kHz for males and a boundary of 5Khz for females is used to separate the spectrum of /t/ into low and high frequency regions. The same method for analyzing /s/ and /z/ utterances was utilized once the average spectra were calculated. Utterances with unrounded vowels are included to serve as controls. Since the method for the average spectrum analysis has been discussed and displayed in detail before, only the statistics calculated from the results of this part are presented to avoid repetition.

Table 6.1 shows the statistics for each Turkish speaker in unrounded and rounded vowel environments. The value of DFRNC is approximately 10dB lower for rounded environment compared to unrounded environment, consistent with the hypothesis that Turkish speakers preserve lip rounding that shifts the prominence to lower frequencies. However, there is a problem in the results from /t/ utterances that was not observed in /s/ and /z/ utterances. In the previous cases, DFRNC took positive values in unrounded environments, indicating that the prominence was at high frequencies. Here, it took negative values in unrounded vowel environment. This observation suggests that the spectral characteristics of a Turkish /t/ may be different than that of an English /t/. Turkish /t/ is known to be dental whereas the English /t/ is alveolar. The beginning assumption was that dental and alveolar consonants share similar spectral characteristics and both can be characterized with a predominance of high frequency energy. Although the spectrum of the burst for the dental /t/ was expected to be concentrated at lower frequencies than the alveolar /t/, it was not expected to be at below 4kHz in unrounded vowel environment.

Table 6.1: Statistics for DFRNC in the Two Vowel Environments for Turkish

Speakers	**Mean in unrounded setting**	**standard deviation**	**Mean in rounded setting**	**standard deviation**
CS Turkish male	-11.59	5.52	-25.42	1.99
OK Turkish male	-9.37	6.60	-20.38	3.59
BD Turkish female	-9.81	4.19	-18.10	3.50
ID Turkish female	-2.39	5.39	-12.06	8.70

Table 6.2 shows the mean and standard deviation of DFRNC in unrounded and rounded vowel environments for each English speaker. The results are not as consistent as the results from /s/ and /z/ utterances. On average, there is a 4dB drop in value of DFRNC when going from unrounded to rounded vowel environments. This drop is 6dB less than the drop in Turkish. There is a lot of variation across speakers. In three out of four subjects, the prominence is at high frequencies in an unrounded vowel environment. In a rounded vowel environment, the prominence is shifted to low frequencies in two out of the four English speakers. The rounding feature of the adjacent vowel is expected to affect the average spectra of /t/ more than that of /s/ or /z/ since averaging is done only over the

burst, which is at the consonant edge, in the vicinity of the vowel, whereas the averaging was done over the entire duration of /s/ and /z/. This may be the reason of getting negative values for DFRNC in /t/ utterances when we did not get them in /s/ and /z/ utterances in English.

Table 6.2: Statistics for DFRNC in the Two Vowel Environments for English

Speakers	Mean in unrounded setting	standard deviation	Mean in rounded setting	standard deviation
KS English male	-8.44	9.12	-13.84	4.39
KO English male	3.19	5.18	-7.53	5.40
MM English female	4.34	5.33	7.82	4.69
KP English female	-6.72	5.39	-7.30	3.81

Another possible reason for the high variation in results is the difficulty of averaging the burst of the stop consonant. The method of averaging the burst over 15 ms did not always eliminate aspiration. English speakers flapped their /t/'s approximately 50% of the time even though they were warned not to do so prior to the recording. Bursts for flaps are known to be highly variable in English. Turkish speakers do not flap their /t/'s in any setting so the results from Turkish data had comparatively more consistency.

Another problem was the choice of 4 kHz as the boundary for males and 5kHz for females, especially for Turkish subjects. These choices worked well with /s/ and /z/ utterances since the prominence in an unrounded vowel setting was above 4kHz for males and 5kHz for females. Preservation of rounding during the consonant shifted the prominence below the boundary. Here in /t/ utterances, the prominences in the spectra seemed to be at a lower frequency than 4Khz for some of the males and below 5kHz for most of the females even in neutral vowel setting.

In summary, results from /t/ utterances do not serve as evidence of the variation in coarticulatory patterns of Turkish and English because the spectra of the consonant are significantly different even in an unrounded vowel environment in Turkish and English.

6.2 Results for palatoalveolar /sh/

Analysis on Average Spectra

The impact of rounding on the two prominent peaks were analyzed individually as discussed in Chapter 3. The effects of the rounding feature on the peak at around 4.5kHz, Peak2, were not significant. However, the effects of rounding on the low frequency prominence, Peak1 were easy to identify. In Turkish the +rounded feature of the adjacent vowels shifted this peak to lower frequencies. In English, the location of the peak did not vary significantly across rounded versus unrounded vowel environments.

Table 6.4 shows the frequency and the amplitude of Peak1, the prominence around 2.5 kHz, in unrounded vowel environment for four speakers of Turkish.

Table 6.3: Location and Amplitude of the Low Frequency Prominence in Turkish Unrounded Vowel Environment

utterances	CS male PEAK (Hz)	CS male AMP (dB)	OK male PEAK (Hz)	OK male AMP1 (dB)	BD female PEAK (Hz)	BD female AMP (dB)	HO female PEAK (Hz)	HO female AMP (dB)
bashak1	2436	41.0	2520	42.4	2709	35.1	2993	32.9
bashak2	2520	35.8	2394	43.9	2741	41.5	3150	34.5
bashak3	2583	38.2	2457	41.8	2772	37.5	3087	34.6
bashak4	2615	42.0	3056	46.3	2520	39.3	3182	30.8
bashak5	2426	36.4	2489	36.0	2993	41.7	3056	29.9
besher1	2706	36.3	2741	42.7	2898	36.1	3150	33.2

Table 6.3: Location and Amplitude of the Low Frequency Prominence in Turkish Unrounded Vowel Environment

utterances	CS male PEAK (Hz)	CS male AMP (dB)	OK male PEAK (Hz)	OK male AMP1 (dB)	BD female PEAK (Hz)	BD female AMP (dB)	HO female PEAK (Hz)	HO female AMP (dB)
besher2	2678	41.1	2457	39.4	2961	43.1	3182	32.3
besher3	2678	36.7	2552	35.5	3056	39.3	3182	29.3
besher4	2646	38.6	2678	32.0	2930	40.8	3119	27.6
besher5	2709	37.5	2615	39.4	3087	36.9	2552	30.9
bitishik1	2741	39.2	2835	46.9	3308	39.8	3182	35.2
bitishik2	2741	41.3	2835	45.8	2457	37.7	3182	36.3
bitishik3	2709	36.9	2646	44.7	3056	41.5	3182	31.2
bitishik4	2709	40.8	2709	43.9	2993	47.1	3213	36.1
bitishik5	2741	41.0	2741	43.9	3024	43.1	3213	29.6

On average, Peak1 was located at 2675Hz in male data and had an amplitude of 40.2dB. In female data, it was located at 3004Hz and had amplitude of 36.2dB.

Table 6.4 shows the results for Turkish speakers in rounded vowel environment.

Table 6.4: Location and Amplitude of the Low Frequency Prominence in Turkish Rounded Vowel Environment

utterances	CS male PEAK (Hz)	CS male AMP (dB)	OK male PEAK (Hz)	OK male AMP1 (dB)	BD female PEAK (Hz)	BD female AMP (dB)	HO female PEAK (Hz)	HO female AMP (dB)
burushuk1	2205	33.2	1701	39.0	1827	41.8	2709	30.0
burushuk2	2583	35.3	1701	38.8	2394	37.6	2709	26.5
burushuk3	2111	36.2	2205	41.4	1922	35.1	3308	30.8
burushuk4	1733	36.3	1733	39.3	1733	37.6	2772	33.9
burushuk5	1733	33.6	1733	38.8	2426	35.7	2709	28.3
kosu1	1796	37.3	1764	39.3	1701	39.2	2835	32.1
kosu2	1764	40.0	1827	39.8	1638	37.9	2772	27.4
kosu3	1733	37.3	1827	39.0	1796	43.4	2867	35.8

Table 6.4: Location and Amplitude of the Low Frequency Prominence in Turkish Rounded Vowel Environment

utterances	CS male PEAK (Hz)	CS male AMP (dB)	OK male PEAK (Hz)	OK male AMP1 (dB)	BD female PEAK (Hz)	BD female AMP (dB)	HO female PEAK (Hz)	HO female AMP (dB)
kosu4	1701	39.1	2142	41.7	1764	41.3	3402	31.0
kosu5	1733	37.3	1764	39.4	1890	39.0	2867	29.9
kusum1	1701	39.1	1733	42.8	1796	41.5	2772	30.7
kusum2	1733	42.4	1827	44.0	1890	37.2	2835	28.6
kusum3	1670	35.5	1733	40.7	1701	38.5	3465	30.8
kusum4	1733	36.8	1764	35.9	1827	38.9	2772	31.9
kusum5	1764	39.4	1733	39.6	1827	41.1	2772	27.3

All of the speakers preserved their lip protrusion as they produced the /sh/, shifting the prominence to lower frequencies. In rounded environment, Peak1 was located at 1829 Hz and had an amplitude of 38.6 dB in male data on average. In female data, the average frequency and amplitude of Peak1 was 2355Hz and 34.7 dB respectively. Peak1 is shifted down by 816 Hz in male data and 649 Hz in female data. The amplitude of the peak did not change significantly across the two vowel environments.

Tables 6.4 and 6.5 show the results for English subjects in unrounded and rounded vowel environments.

Table 6.5: Location and Amplitude of the Low Frequency Prominence in English Unrounded Vowel Environment

utterances	KS male PEAK (Hz)	KS male AMP (dB)	KO male PEAK (Hz)	KO male AMP1 (dB)	MM female PEAK (Hz)	MM female AMP (dB)	KP female PEAK (Hz)	KP female AMP (dB)
audition1	2552	39.1	2804	33.1	2741	39.2	3402	40.3
audition2	2709	38.0	2709	29.3	2861	41.6	3056	32.9
audition3	2646	40.0	2741	29.7	2835	42	3056	39.0
audition4	2615	41.1	2709	32.5	2867	45.8	2993	33.4

Table 6.5: Location and Amplitude of the Low Frequency Prominence in English Unrounded Vowel Environment

utterances	KS male PEAK (Hz)	KS male AMP (dB)	KO male PEAK (Hz)	KO male AMP1 (dB)	MM female PEAK (Hz)	MM female AMP (dB)	KP female PEAK (Hz)	KP female AMP (dB)
audition5	2615	43.6	2709	24.8	2772	39.7	2646	27.0
mission1	2646	40.5	2741	32.3	2867	36.8	3056	34.3
mission2	2646	38.3	2741	26.7	2930	37.7	3056	34.2
mission3	2583	36.0	2772	29.2	2898	40.7	3119	38.9
mission4	2615	40.8	2804	33.4	2835	47.1	3497	39.4
mission5	2615	42.3	2709	32.5	2867	39	3056	29.3
station1	2678	40.9	2709	24.3	2835	37.3	3024	36.4
station2	2646	39.9	2678	27.7	2835	40.9	2993	34.1
station3	2741	39.4	2709	28.4	2835	41	2993	39.9
station4	2646	40.3	2741	29.0	2835	42.6	3024	36.6
station5	2741	42.0	2867	24.6	2835	40.7	3150	32.9

The average location and amplitude of Peak1 is 2695Hz and 34.7 dB in the male data, 2959Hz and 38.0 dB in the female data. The similar behavior of the two languages in the unrounded vowel environment may be interpreted that there is no significant deviation in the place of articulation for /sh/ that affects the location of PEAK1. In English /sh/ is generally rounded independent of vowel environment.Turkish /sh/ may be different than the English one, however the results here suggest that the difference does not affect Peak1.

Thus the difference that comes into picture in rounded vowel environment can be attributed to the variation in labial coarticulatory organization.

Table 6.6: Location and Amplitude of the Low Frequency Prominence in English Rounded Vowel Environment

utterances	KS male PEAK 1 (Hz)	KS male AMP1 (dB)	KO male PEAK 1 (Hz)	KO male AMP1 (dB)	MM female PEAK 1 (Hz)	MM female AMP1 (dB)	KP female PEAK 1 (Hz)	KP female AMP1 (dB)
hushu1	2394	38.3	2961	42.1	2646	37.8	3371	46.6
hushu2	2300	38.3	2615	31.5	2678	42.7	2930	36.2
hushu3	2300	37.0	2583	31.3	2615	42.2	2993	37.4
hushu4	2331	37.6	2583	33.1	2615	36.5	2961	41.5
hushu5	2363	36.2	2520	33.2	2678	35.8	2678	33.6
kushu1	2426	39.8	2552	30.3	2583	38.6	2835	32.7
kushu2	2457	39.8	2646	34.6	2615	34.7	3402	39.8
kushu3	2426	38.1	2615	27.7	2678	40.1	2835	33.1
kushu4	2394	40.4	2552	30.1	2835	41.6	2867	35.9
kushu5	2363	41.3	2615	35.2	2898	34.8	3339	39.8
moshu1	2300	33.3	2804	32.9	2552	38.9	2772	38.3
moshu2	2898	39.2	2457	30.0	2646	41.2	2867	37.9
moshu3	2772	36.5	2583	31.0	2772	36.2	2772	38.0
moshu4	2426	38.8	2583	30.0	2804	38.5	2835	37.8
moshu5	2489	42.4	2615	34.3	2709	38.0	2835	34.6

In rounded vowel environment, the average frequency of Peak1 is 2531Hz and amplitude is 35.5dB for males and 2821Hz and 38.0dB for females. Thus, there is no significant drop in the frequency of the peak in a rounded vowel environment compared to an unrounded vowel environment. The peak is shifted down only by 164 Hz in the male data and by 138 Hz in the female data. In addition, the +rounded feature of the vowels did not change the amplitude of Peak1 as well. As in previous parts, the impact of +rounded feature of the adjacent vowels on the spectra of the consonant is negligible compared to the impact seen in Turkish.

Table 6.7 summarizes the average location of Peak1 in two environments for both languages and illustrates how differently the two languages behave in rounded vowel environment.

Table 6.7: Average Values for the Location of the Peak in two Environments for English

Summary of Results	**Unrounded environment (Hz)**	**Standard Deviation (Hz)**	**Rounded Environment (Hz)**	**Standard Deviation (Hz)**	**Average drop in frequency due to rounding (Hz)**
TURKISH MALE SUBJECTS	2645	145	1829	207	816
TURKISH FEMALE SUBJECTS	3004	221	2355	543	649
ENGLISH MALE SUBJECTS	2695	71	2531	170	164
ENGLISH-FEMALE SUBJECTS	2959	180	2821	220	138

Results from /sh/ utterances are consistent with the previous observations for /s/ and /z/. As before Turkish speakers preserved their lip protrusion during the consonant whereas English speakers did not when producing two rounded vowels separated by a nonlabial consonant. Peak1 shifted down to low frequencies in Turkish and stayed at around the same frequency in English.

Chapter 7

GENERAL DISCUSSION

Boyce had shown that there is a significant difference in the way Turkish and English speakers organize coarticulation, at least in the way they use lip protrusion for segments with multiple rounded vowels separated by a nonlabial consonant. Her study was based on analysis of EMG signals and lip movements. The main question behind this thesis was whether it would be possible to observe the difference Boyce reported through an acoustic analysis. The data presented here suggested that it is. Acoustic analysis is powerful enough to display the widely cited "trough" pattern of rounding for English and the "plateau" pattern for Turkish.

The results from different types of consonants had varying consistency. The analysis on average spectra gave best results for the alveolar fricatives /s/ and /z/ since the method was developed initially for them. Turkish and English /s/ and /z/ had the same spectral characteristics in unrounded vowel environment, so it was concluded that there is no apparent deviation in the place of articulation for these fricatives across the two languages. The deviation in spectra of the consonants in rounded vowel environment was a consequence of the difference in coarticulation patterns employed by the two languages.

The analyses of utterances with /s/ and /z/ were done separately for male and female subjects. The motivation behind this was to make sure that there is no gender specific difference in the coarticulation patterns used. Since no such differences were observed, the analyses for both genders were done simultaneously for the other consonants. The difference in average vocal tract lengths of the two genders was taken into consideration when locating the resonances in the spectra.

The time analysis confirmed the conclusion that Turkish speakers preserve and English speakers terminate lip rounding during the consonant that intervenes between two rounded vowels. this finding is consistent with the results derived earlier from the average spectra analysis. In addition, the time analysis suggested that there is some anticipation of rounding in English. However it is limited to the close vicinity of the rounded vowels. Although the prominence shifts to low frequencies in these regions, i.e., DFRNC becomes negative, it shifts back to high frequencies by the midpoint of the consonant, i.e., DFRNC becomes positive. The anticipation is not spread onto the entire duration of /s/ or /z/. Although the production of the consonant is not completely context independent, there is no “look ahead” mechanism comparative to the one in Turkish that leads the speaker to move forward to an articulatory goal as early as possible.

Voicing did not have a significant impact on coarticulatory organization, as expected. The results from /s/ and /z/ utterances were similar. The duration of /z/ was much shorter than that of its voiceless cognate in both languages. In Turkish, the difference in duration did not have any effect on the results since lip protrusion is preserved during the entire duration of /s/ and /z/ when they are surrounded by two rounded vowels. In English, anticipation of rounding occurs only in the close vicinity of the rounded vowels and when the duration of the consonant is shorter, the impact of the regions in which anticipation takes place was greater on the average spectra. Although in English /z/ utterances, the prominence in the consonants’ spectra stayed at high frequencies, the magnitude of DFRNC was smaller than for /s/ utterances.

There were problems in the results from /t/ utterances. The languages behaved differently even in unrounded vowel environment. The difference in spectra of Turkish and English /t/ in a rounded environment could not be attributed only to the variation in coarticulation.

The results from /sh/ utterances supported the previous observations. In Turkish, the prominence that is located at around 2.5kHz in a neutral vowel environment was significantly shifted to lower frequencies. In English, the peak was located at around the same frequency in both environments.

With regard to the competing "look-ahead" and "time-locked" models of coarticulation, a collective interpretation of these results is that Turkish uses a look ahead model of coarticulation since the model predicts a plateau pattern of rounding for sequences of rounded vowels separated by a non-labial consonant and Turkish speakers display this pattern consistently. On the other hand, English speakers are more likely to utilize a time-locked model of coarticulation since they terminate rounding at the consonant that intervenes the two rounded vowels as suggested by the model.

From these findings it can be concluded that different languages employ different articulation strategies. Although humans are equipped with identical speech apparatus, the combination of phonology, lexicon and syntax leads them to utilize that apparatus in different organizations. Boyce had attracted attention to the fact that the vowel harmony existent in Turkish may be the reason why Turkish deviates from English. Languages similar to English in their freedom of combining rounded and unrounded vowels freely were reported to display the trough pattern. To increase the validity of this interaction between phonology and articulation, further studies in languages with constraints such as Turkish are necessary.

One problem researchers faced studying the articulation patterns was the difficulty of data collection. Most of the studies reported in literature were based on EMG signals and labial movement measurements. Obtaining these measurements were cumbersome so studies used only a limited number of speakers. Naturally, this made it difficult for the scientist to make generalizations about a language. The present study showed that it is possible to study articulation patterns through acoustic analysis. Since the method of data collection is much easier, more speakers can be utilized in studies. This will increase the validity of conclusions and enable scientists to examine more languages and hence to look for universal patterns in speech production. Once this is done, the efficiency of current speech recognition and synthesis technologies will increase significantly.

7.1 Directions for Future Work

This study confirmed that languages with different grammatical constraints utilize different patterns of coarticulation through an acoustic analysis. Determining the specifics of the coarticulatory organization of Turkish or English was beyond the scope of this study. Turkish was observed to be more consistent with the "look ahead" model. In all of the utterances, a single consonant intervened the two rounded vowels. Looking at the results, we can not answer the question whether the same plateau pattern of rounding will be replicated in cases where clusters of consonants separate the two rounded vowels. Future studies can repeat the acoustic analysis utilizing utterances with different numbers of consonants in between the vowels.

The results of this study supported the conclusion that English speakers are likely to use a time locked model, and thus their production of consonants is less context dependent than Turkish. The time analysis showed that there is some anticipation prior to onset of the vowel. However, it is not possible to determine specifics of this anticipation by looking at the results of this study.

The results supported Boyce's observation that different constraints in languages can give rise to different mechanisms for uttering a sequence of sounds with the same feature specifications. More cross-linguistic studies are needed to support or disprove this hypothesis.

REFERENCES

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

Boyce, S.E.,(1990). "Coarticulatory organization for lip rounding in Turkish and English," J.Acoust.Soc.Am. 88, 2584-2595.

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.

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

Gay,T.J.(1979). "Coarticulation in Some Consonant-Vowel and Consonant Cluster-Vowel Syllables," in Frontiers of Speech Communication Research, edited by B.Lindblom and S.Ohman (Academic, New York), pp. 69-76.

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.

Halle,M. and Stevens, K.N.,(1991)."Knowledge of language and the sounds of speech," in Music, Language, Speech and Brain, edited by J.Sundberg, L.Nord, and R. Carlson (MacMillan, Basingstoke, Hampshire, England), pp. 1-19.

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

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

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.

Kozhevnikov,V.A.,and Chistovich, I. A. (1965). Speech Production and Perception (Joint Publication Research Service,Washington, DC).

Lubker, J.F. (1981). "Temporal aspects of speech production: Anticipatory labial coarticulation," Phonetica 38,51-65.

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

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

Shadle, C.H.(1985)."The acoustics of fricative consonants," Research Laboratory of Electronics, Technical Report 506, (MIT,Cambridge,MA).

Stevens, K.N.,(1971). "Airflow and turbulence for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. 50, 1180-1192.

Stevens, K.N.,(1993). "Modelling affricate consonants," Speech Communication 13, 33-43.

Stevens, K.N.,(1998). Acoustic Phonetics, Cambridge, MIT Press.

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.