

## TOWARDS A PHONETICS OF THE DISCOURSE

ANTÔNIO R.M. SIMÕES

The University of Kansas at Lawrence

### ABSTRACT

Experimental phonetics research had been characterized since its beginnings to very recent date, by studies in which the analyses were limited to the word level. Of course, these limitations were due to the instruments of analysis and not to the researcher. Studies in Experimental Phonetics beyond the word level are of recent date as a consequence of recent developments in speech technology. The study presented here comes from my dissertation (Simões: 1987) which deals with Experimental Phonetics at the discourse level. The realization of this type of study can be accomplished only if new techniques of segmentation and analysis are developed. For this reason, techniques of segmentation and analysis were developed for the purpose of the present study.

The present discourse analysis deals with the relationship between lower level linguistic components (phonetics/phonology) and higher level linguistic components (syntax, discourse) from the experimental analysis of the temporal organization of the three extreme vowels [i], [a], and [u]. D. Klatt's (1976) model, based on American English, is examined in order to know how it could be adapted to Brazilian Portuguese. Following this comparison, rules are proposed to predict sound-segment duration of Brazilian Portuguese in continuous speech for the particular case of the informant in this study. Data for this investigation were obtained from one speaker from Rio de Janeiro who read a text for children containing over one thousand words.

### 1. INTRODUCTION

Most analyses in phonetics use duration, intensity and frequency as basic parameters. This study concentrates on the analysis of duration. Although duration is not linguistically significant in BP, there are several reasons for concentrating so much effort in analyzing duration.

Fant (1970:224) remarks that "the simple and fundamental cue of duration deserves greater attention than is conventionally paid to it." The work of Klatt (1976) and many other scholars (Peterson and Lehiste: 1960; Lindblom: 1979; Lindblom et al.: 1981) have shown that temporal patterns contain important information regarding the hierarchy of linguistic components (phonology, morphology, syntax, semantics) and building of speech models as well as information of practical use for speech synthesis-by-rule, automatic verification and recognition of the speaker, speech pathology, diachronic and synchronic phonology, etc. Charles Read (personal communication) used Klatt's (1976) conclusions to improve listening comprehension of children, by giving extra duration to words at syntactic boundaries of recorded readings.

Studies on duration in BP, then, do not have immediate application to the teaching of BP. The teaching of BP will eventually benefit from this type of study, after other parameters, besides duration, have been studied. The major purpose of the present analysis is to provide BP with a model as effective as D.H. Klatt's (1976) model for American English and consequently develop a better understanding of BP at the discourse level.

The study of duration in BP has received the attention of three scholars, besides the present study on the phonetics of the discourse: Abaurre-Gnerre (1981), in a phonological analysis, and in both a phonological and experimental analysis there are the works of Major (1981, 1985) and Kelm (1989). These investigations use duration to analyze the rhythmic patterns of BP.

A survey of studies on duration of other languages (Martinet: 1949; Jakobson et al.: 1952; Fry: 1955; Miller and Nicely: 1955); Lisker and Abramson 1964; Kim and MacNeilage: 1973; Chafcouloff et al.: 1976; Ferrero et al.: 1979) shows that duration, at the phonetic and phonological level is affected by articulatory effort, namely articulatory effort will be greater during vocal cords vibration, stressing of a sound-segment, or in articulatory displacement (distance) when there is a passage from one point to another. As Lehiste (1970) remarks, the longer the displacement the longer the sound-segment.

*Segmental duration* is reflected as two linguisti phenomena: *reduction* and *expansion*. The goals of the present study are:

- [1] To uncover the factors or language components that affect sound-segment duration by both shortening and lengthening *unmarked* vowel duration.

- [2] To determine how these factors interact. Do they interact simultaneously or in series? If in series, is there an ordering process?
- [3] To determine vowel-duration change rules for the specific case of this BP speaker fitting D.H. Klatt's model (1976) to predict duration of the three vowels studied.

It is necessary to point out that in the present work the term "unmarked" is used synonymously with Klatt's use of the term "inherent" and Peterson and Lehiste's (1960) "intrinsic." Therefore, I would like to define the term "unmarked" in the following manner: "The *unmarked duration* of a given sound-segment  $\mu$  is the *median* duration of all occurrences of that sound-segment  $\mu$  in a text of over one-thousand words."

## 2. METHOD

### 2.1. Introduction

Spectrograms, since Liberman's et al. (1967) paper, have been basically left aside as a useless document for serious linguistic analysis. Cole et al. (1980) returned some credibility to the use of spectrograms through the use of spectrograms in their analyses. The consequences of the spectrogram regaining its position among speech researchers are such that even in state-of-the-art research these displays are most helpful. Cole et al. (1983) have found that derivation of acoustic correlates directly from spectrograms is an important trend in phonetics.

### 2.2. Experimental protocol

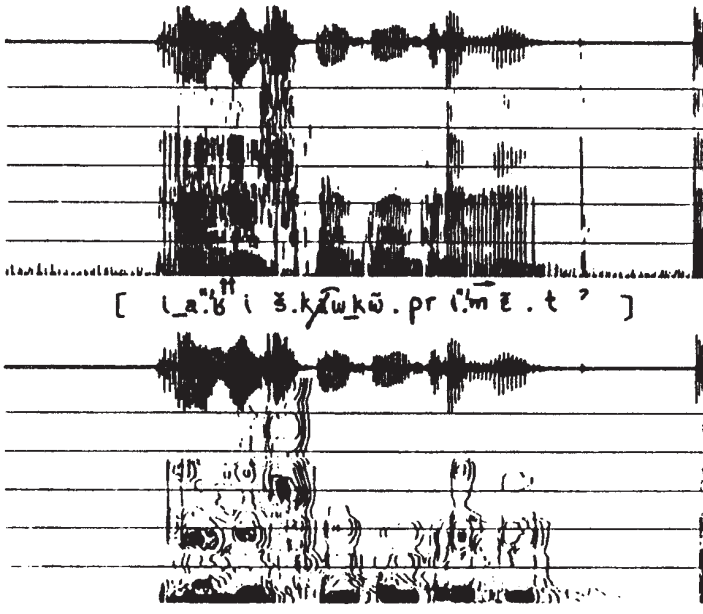
The experimental protocol is organized according to three main procedures: the production of the corpus (recordings), the production of the spectrograms for sound-segment segmentation, and data analysis.

A single speaker of BP, PM, 32 years old, from Rio de Janeiro, was recorded. PM read a 1286-word text for children in two recording sessions. The recording sessions took place one week apart. PM was asked to read the same text three times in each session, totalling six readings of the same story. The same reading of the second session, i.e. the sixth recording, is the one used in the analysis. The recordings were made at the language laboratory of the University of Texas at Austin, with the help of a laboratory technician. An acoustically isolated recording booth was used and the recording was done using an AKG dynamic, unidirectional cardioid microphone situated at 40 cm from the subject's mouth. Before

each recording the system was calibrated. The tape used is as AMPEX tape, 1/4", 1.5 mil, mylar.

Some 700 spectrograms were produced to observe and measure the 1286-word text and its sound-segments in this study. The spectrograms were obtained through a stereo tape deck TEAC A-2300SX coupled to a Digital Kay Sonagraph 7800. Since the use of the usual 3D broad band spectrogram would require too many arbitrary decisions in the analysis of continuous speech, a special technique was developed during the preliminary analyses elsewhere (Simões:1987). This technique combines two spectrograms for each sample analyzed. As figure 1 shows, one spectrogram has the regular broad band and a second one is made using the amplitude contour. Superimposed on the top of each spectrogram there is an oscillographic image.

**Figure 1:** Spectrograms of the sentence "...e arrisca um cumprimento:" (... and she tries to salute:") showing superimposition of different images from the Sona-Graph.



### 2.3. Segmentation procedures

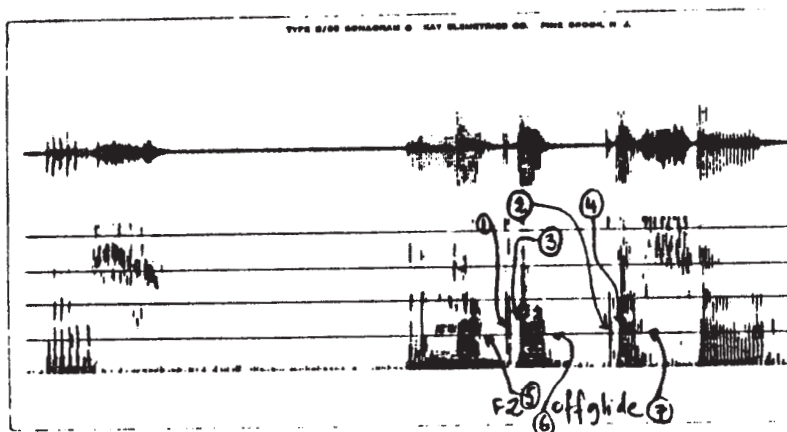
Studies of sound-segment are usually done in one of four ways: (1) using isolated sound-segments (Stevens:1960), (2) using a sentence context (Heinz and Stevens: 1961; Chafcouloff et al.:1976), (3) using both (Jassem:1965, 1968), and using connected speech in long context (Klatt:1976;Simões:1987).

Klatt's way of segmenting served as the base for my own segmentation rules or procedures. Besides using Klatt's way of segmenting I also borrowed the notions of onglides, offglides, steady state, and simple and complex nuclei found in Lethiste and Peterson (1961). The reason for centering visual observations on the second formant region is found in Lieberman (1977) and further reinforced by the notion of centers of gravity developed in a study Chistovich et al. (1979).

The manual segmentations are established according to the type of sounds under study, namely, the vowel. The visual cues which allowed quite reliable segmentation using the Digital Kay Sona-Graph are listed below in order of importance. When one procedure did not suffice, I would go the next or combine two or three procedures. The following pages contain the rules for manual segmentation that should suffice for manual segmentation. However, in case there are doubts, there is secondary information that can be used. For example, in the high frequency regions, namely regions above 4 kHz, there is usually a clear distinction between vowels and consonants.

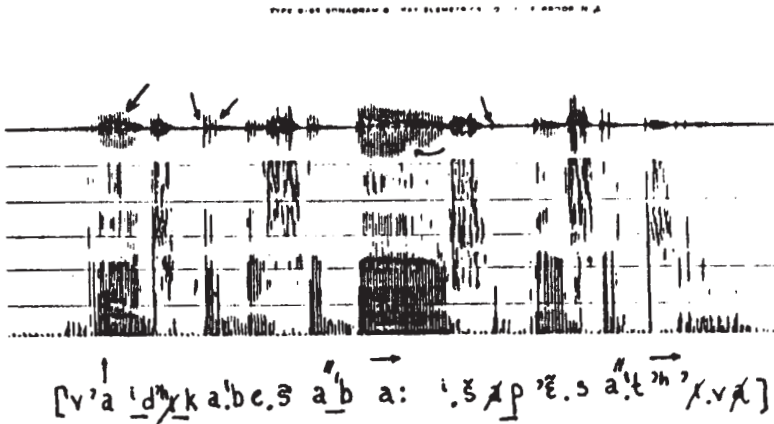
*Rule 1.* The point of departure is the onglide and the offglide (Lethiste and Peterson, 1961) of the second formant transition, i.e. when the second formant starts (most of the times preceded by a "blank"), and when it is interrupted. In other words, the whole formant transitions are included as part of the sound to which they belong. In case there is a burst or a glottal stop, the segmentation is done before the burst or glottal stop, still at the F2 level.

**Figure 2:** Illustration of the point of departure for the manual segmentation process: the onglides and offglides (Lehiste and Peterson, 1961) are to be segmented first. Arrow 1 and 2 indicate bursts at the F2 region since bursts and glottal stops are segmented as part of the following segment; arrows 3,4 indicate F2 onsets; and arrows 5,6,7 show F2 offglides.



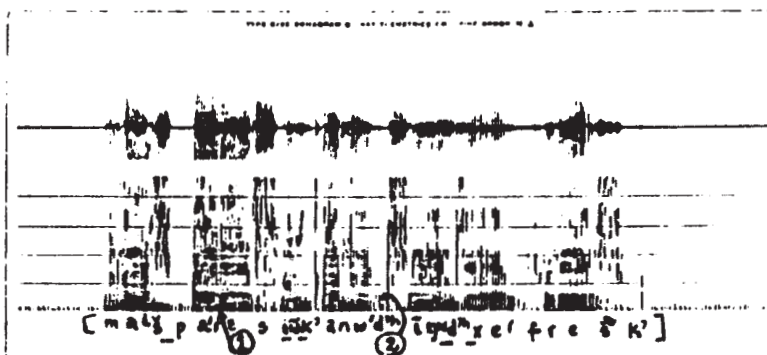
**Rule 2.** Use the waveform information provided by a superimposed oscillograph built in the Digital Sona-Graph. This sound wave image shows clear variations when the glottal pulses (vertical spikes) shorten vertically to become almost confused with the zero lines. Such variations are indicative of the end of sound-segment and/or the beginning of another. This is a consequence of damped oscillations. Any other sudden change in the amplitude of these soundwave images are potential indicators of sound boundaries, although these sudden changes are observed outside the zero line region.

**Figure 3:** The superimposed oscillograph image is mandatory in analyses using continuous speech. The document illustrates the sentence "Vai de cabeça baixa, pensativa" with sudden changes in the superimposed oscillagraphic image, indicative of sound segment boundary. The five arrows in the figure below point to phoneme boundaries in the oscillagraphic image.



**Rule 3.** Observe the changes in the relative intensity in the first formant region, reflected in the darkening of images. Any change in energy concentration might indicate a new sound. It is known that true consonants have less energy than true vowels. In terms of sound relative intensity (reflected in the darkening differences on the spectrogram), consonants have less intensity than vowels. This may be explained by the production of both classes of sounds. Vowels find no obstacle in their way out of the vocal tract and is realized with most of its energy from the glottal source. On the other hand, consonants are completely or partially obstructed in their realization creating a loss of energy.

**Figura 4:** The differences in the relative intensity of speech-sounds in the region of the first formants. Consonants are characterized by weaker energy (less darkening in the spectrograms) than vowels. Arrow 1 points the less darkened image of the liquid [r] in between the darker images of [a] and [ɛ] in the word [pa',rɛ.st] "parece" ("looks like"); arrow 2 shows [d] less darkened than the preceding [ω] in [ka.nω.'dʰɪ.ŋω], "canudinho" ("little straw").

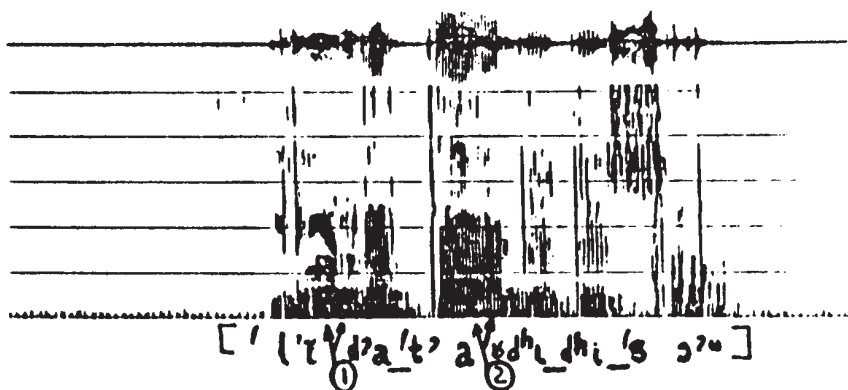


**Rule 4.** Take into consideration the lowering or rising of the fundamental frequency observed through spikes spacment (glottal pulses). Lowering of the  $F_0$  (fundamental frequency) happens when, relative to the glottal pulses of the preceding adjacent sound-segment, the glottal pulses are more apart; in rising of the  $F_0$  glottal pulses are closer, relative to the glottal pulses of the preceding adjacent sound-segment.



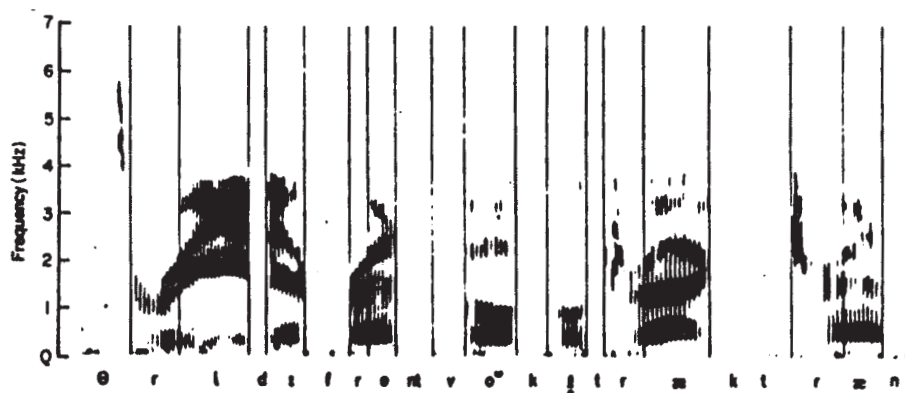
**Figure 5:** Spectrogram showing how rule 4 is to be applied. The arrows, as above, indicate the points described. Double-arrow 1 indicates pulses, on the left branching of it, closer (higher Fo) than on the right branching; double-arrow 2 shows a similar example with the left branching indicating rising of Fo, and the right branching lowering of Fo.

TYPE B/88 SONAGRAM © KAT ELECTRONICS CO PINE BROOK, N. J.



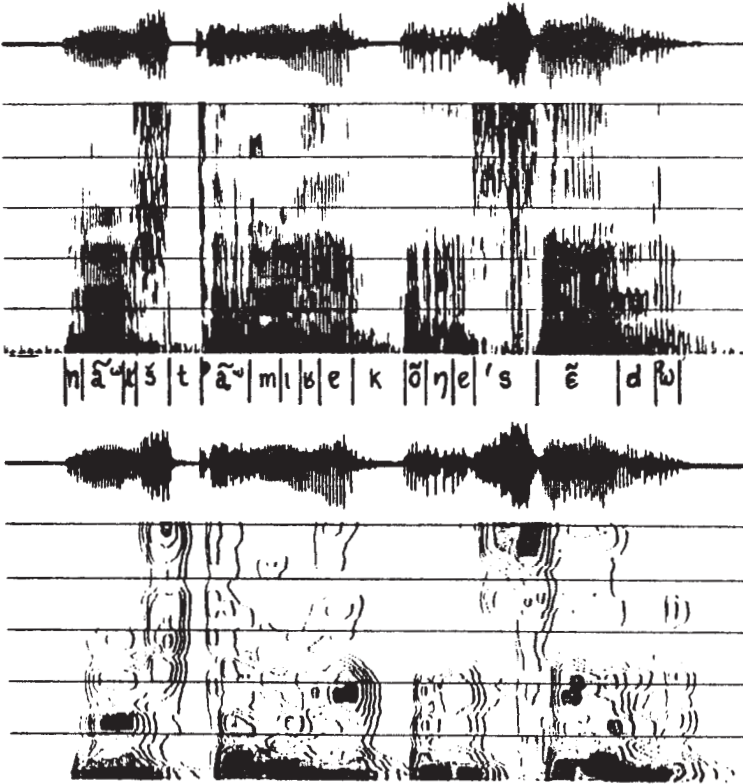
*Rule 5.* As Klatt (1975) suggested, the burst characteristic of stop consonants, is considered as part of the following sound-segment, not as part of the stop. This is helpful in treating vowel segmentation when vowels are not preceded by stop consonants. The spectrograms in this study show that a glottal stop (visually similar to a burst) precedes all vowels most of the times, especially the front low [a]. In addition, liquids are also preceded by a glottal stop. Finally, any aspiration, including that associated with palatalization is included in the vowel portion.

**Figure 6A:** Klatt's (1975) method for segmentation. This method is repeated in the present study as Figure B shows for comparison with Figure A.



**Figure 1** A broad-band spectrogram is shown of the sentence fragment "Three different vocal tract trans(fer) ...". Vertical lines have been drawn at segment boundaries according to criteria defined in the text.

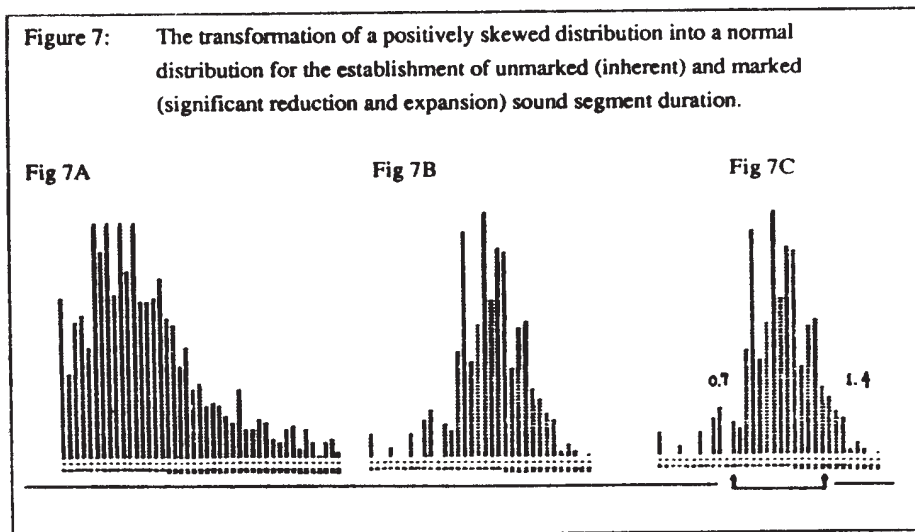
**Figure 6B:** An example of the segmentation done in the present study. The segmentation of the continuous speech of the speaker in this study is reliable only if both images using the regular 3D and the amplitude contour are superimposed.



### 3. DISCUSSION

Analyses of all data from these vowels in the corpus consistently show functions positively skewed, namely lower values concentrate on the left side of the abscissa, in such a way that higher values will spread rightward. A parameter such as duration is expected to be positively skewed. For this reason, the median was chosen as the measurement type for this parameter, instead of the mean. In figure 7A a common pattern of the data is shown. Besides the use of the median in the establishment of an unmarked duration, I have found that the common logarithm of the skewed function can be used as well.

Figures 7A-7C illustrate this additional technique. Common logarithms of positively skewed distributions were taken so that a normal distribution such as the one seen in 7B would be obtained. Figure 7C shows how significant lengthening can be determined by considering values greater than 1.4 times the median duration of the segment studied. Shortening processes appear on the left side of the abscissa and the cutoff point for significant shortening processes is 0.7.



The unmarked (inherent) values are obtained from the grand median, namely all stressed and unstressed positions. The grand medians (unmarked duration) are given in Table 1 and the measurements in stressed and unstressed positions are given in Table 2. Measurements from tables 1 and 2 are used in the application of the model (equation 1).

		items
i,a,u	60ms	1275
i	56ms	379
a	64ms	636
u	48ms	260

	+stress	items	-stress	items
i,a,u	120ms	239	52ms	1036
i	108ms	78	48ms	301
a	120ms	120	60ms	510
u	112ms	35	44ms	225

In the present study the speaker's speech is affected by phonological, word, syntactic and semantic factors in both lengthening and shortening processes. These remarks do not necessarily reflect how production processes take place. These conclusions might be related only to the model chosen here and to the purpose of synthesizing speech. These factors will be used in an adaptation of Klatts' (1976) model to predict and consequently to create rules for sound segment synthesis of duration in BP. The accuracy of Klatt's model is tested mainly around the linear equation  $D_o = K x (d_i - D_{min}) + D_{min}$  (eq 1) where  $D_o$  for duration output, is the duration sought at any given point in the text;  $K$  is a constant value for each phonological environment, each position in a word, each position in a sentence, and each type of semantic factor;  $D_i$  is the unmarked duration for each sound-segment; and  $D_{min}$  is the minimum reduction the inherent sound-segment duration can have (see Klatt, 1976:1215).

The value of  $K$  is established by operating on the original equation (1) making equation (2)  $K = (D_o - D_{min}) / (D_i - D_{min})$ . Due to possible complete reductions of sound segments in PM's speech at the end of a word and

especially at the end of a word at the end of a sentence, the rules proposed in this study attempt to reflect this phonological characteristic of PM, common in BP.

#### **4. THE APPLICATION OF THE MODEL**

The present work is based on results and analysis done exclusively at the acoustic domain. According to Klatt (1976:1208) "duration often serves as a primary perceptual cue in the distinctions between (1) inherently long versus short vowels, (2) voiced versus voiceless fricatives, (3) phrase-final versus non-final syllables, (4) voiced versus voiceless postvocalic consonants, as indicated by changes to the duration of the preceding vowel in phrase-final positions, (5) stressed versus unstressed or reduced vowels, and (6) the presence or the absence of emphasis." Conclusions (3), (5) and (6) are relevant to the speech of PM at the phonetic level. Conclusions (1), (2) and (4) do not apply to PM's speech because of the phonology and phonotactics of BP, namely BP vowels are not distinctive in terms of duration and most syllables in BP are open. Klatt suggests rules to reduce or expand sound-segments in English that apply recurrently from local to outer units according to the equation (1). While Klatt's work assesses the importance of the linguistic information by means of perceptual tests, the present work filters out significant linguistic information by means of statistical analysis. Klatt (1973b) already had preliminary attempts to apply a percent change model which applies several rules simultaneously, which demonstrated that this model failed. Lindblom and Rapp (1973:47) did propose a model for Swedish with recurrent rules going from outer units (sentences) into inner units (clauses and words). Lindblom et al. (1981) then moved to Klatt's approach going from inner to outer units. There are many more details in the Lindblom et al. (1981) model that make it quite attractive in testing for other languages as well. Because of the simplicity and effectiveness of Klatt's model, which uses only four rules and nine parameters to predict average vowel duration occurring in 56 situations (Klatt, 1976:1217), the present study tested this model first.

The rules of the present study are listed below. They were established for the prediction of shortening and lengthening of the three vowels [i, a, u] according to equation (1). Although I proposed at first to have equation (1) operating in series from domain 1 to domain 4, I do not have an argument against models in which all involved factors operate through one functional form, viz. simultaneously. I have not yet tested this possibility myself, but I hope to have an opportunity to run such rules in series and simultaneously in a laboratory. The value of  $K$  here is obtained

by means of equations (2) and the unmarked duration of each vowel initializes each process in any position within a word in the case of applying the rules in series. The subsequent  $D_i$  values are the outputs ( $D_o$ ) of the application of the rule just applied.

**Level 1 (Sound segment domain)** - Rule 1: Initialization. From a given sound duration inventory the unmarked sound segment duration is set.  $D_i$  of PM's vowels are: [i] 56ms, [a] 64ms, [u] 48ms; Rule 2: If the phonetic voiceless fricative [s] follow the vowel within the same syllable, shorten the vowel by  $K = -1.17$  (65% decrease). In case the consonant is [x], shorten the vowel by 25%, viz.  $K = .17$ . Rule 3: If a phonetic voiced consonant follows the vowel within the same syllable, no change,  $K = 1.0$ .

**Level 2 (Word domain)** - Rule 4: If the vowel is in postonic position, decrease the vowel by 25%, viz.  $K = .17$ ; Rule 5: If the vowel is in pretonic position, not preceded by a consonant, decrease it by 10%, i.e.  $K = .67$ . In case the vowel is preceded by a consonant, increase it by 42%, viz.  $K = 2.4$ ; Rule 6: If the vowel is in immediate pretonic position, increase it by 13%, viz.  $K = 1.43$ ; Rule 7: If the vowel is in stressed position, increase the vowel by 90%, viz.  $K = 4$ .

**Level 3 (Sentence domain)** - Rule 8: If the vowel is at the beginning of a sentence or a pause, no change,  $K = 1.$ ; Rule 9: If the vowel is in sentence medial position, decrease by 13%, viz.  $K = .57$ . Rule 10: If the vowel is in sentence final position without physical pause, increase the vowel by 20%,  $K = 1.67$ , but if the vowel is in a major final position, and a physical pause follows, increase the vowel by 32%,  $K = 2.07$ .

**Level 4 (Semantic domain)** - Rule 11: If vowel is within a focused word, increase the vowel by 60%, making  $K = 3$ ; Rule 12: If the vowel is in an exclamatory word, increase the vowel by 80% by making  $K = 3.7$

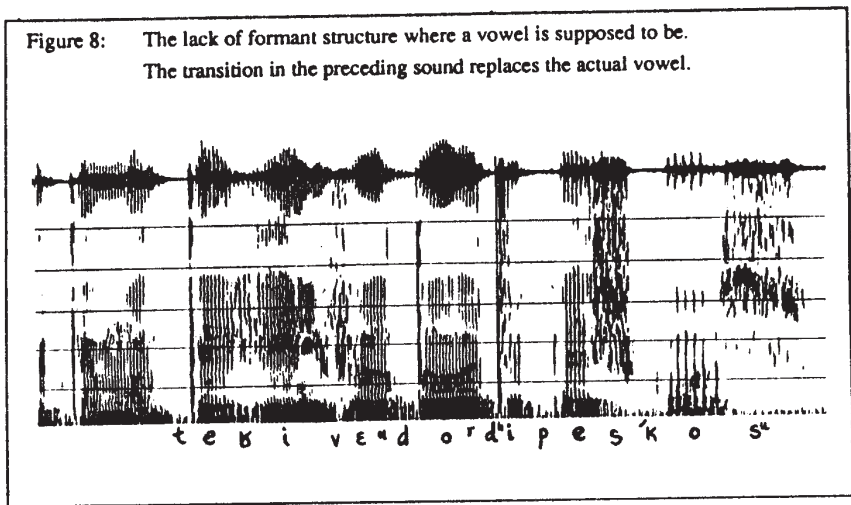
Rules 10-12 do not apply to vowels in postonic position and stressed vowels are the only ones with a minimum duration.

When comparing both sets of rules in Klatt (1976) and in the present work, one has to bear in mind that all but one rule in Klatt's model will *shorten* vowels for the inherent vowel duration in Klatt's model is the *longest* vowel (1976:1217). In the present model the phonological operations can either shorten to a complete reduction or lengthen the vowel segments.

## 5. CONCLUSION

This model and its duration changes rules are not intended to reflect mental processing. The preceding discussion gives an idea of the difficulties found in the first incursions into this type of analysis subsequent to proposition of a model. I have already mentioned the need

to use the median, instead of the longest vowel, because the BP syllables are normally CV sequences, especially in connected speech. Analysis of the signal from connected speech, as shown in Figure 8 will show that in connected speech, very often the whole formant structure of the nucleus disappears.



Thanks to a reanalysis of the spectrograms produced for this study and to an article by Parker and Diehl (1981), the interpretation of the temporal organization of BP vowels at the discourse level may acquire a new view. As depicted in Figure 2, vowel in connected speech commonly "enters" the preceding consonant in the speech of my informant. It may be necessary for BP, then, to eliminate the element  $D_{min}$  from the equation (1), making it a much simpler equation, viz.

$$\text{equation (3) } D_0 = K * D_j$$

or, if equation (1) is to be maintained, the duration of a vowel is not its nucleus, but instead the preceding consonant and whatever is left of the vowel (Parker and Diehl: (1981). In the latter option, the  $D_{min}$  is kept. A



model such as Klatt's (1976), can only account for the cases of complete reductions of the formant structures, if the duration of a vowel is considered as being the duration of both the vowel and the preceding sound. If the duration of a vowel is considered to be what is the formant structure, then equation (1) used in Klatt's work needs to be simplified by cancelling the element  $D_{\min}$ . In case this assumption is true, a simpler linear equation that does not include  $D_{\min}$  has to be used:  $D_0 = K * D_i$  (eq 3).

## REFERENCES

- ABAURRE-GNERRE, M.B.M. (1981). Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. In *Cadernos de Estudos Lingüísticos*, 2, 23-44. Campinas, Brazil: Instituto de Estudos da Linguagem, Unicamp.
- ABERCROMBIE, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- BECKER, R.A. and J.M. Chambers (1984). *S: an interactive environment for data analysis and graphics*. Belmont, Ca: The Wadsworth Advanced Book Program.
- CÂMARA, J.M., Jr. (1977). *Pra o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão.
- CHAFCOULOFF, M. et.al. (1976). Effects de la coarticulation sur les caracteristiques acoustiques des contoïdes fricatives du français. In *Travaux de l'Institut de phonétique d'Aix*, 3, 61-113.
- CHISTOVICH et al. (1979). Centres of gravity and spectral peaks as the determinants of vowel quality. In *Frontiers of speech communication research*, B. Lindblom and S. Ohman, eds. London: Academic Press.
- COLE, R.A., A.I. Rudnick, V.W. Zue and D.R. Reddy (1980). Speech as patterns on paper. In *Perception and production of fluent speech*, R.A. Cole, ed. Hillsdale, N.J.: Earlbaum.
- DELATTRE, P. (1966). *Studies in French and comparative phonetics. Selected papers in French and English*. London, Paris, The Hague: Mouton.
- DIGITAL *Sona-Graph, Model 7800, Instructional Manual* (1984). Pine Brook, NJ: Kay Elemetric Corp.
- FANT, G. (1970). *Acoustic theory of speech production*, 2nd. ed. The Hague: Mouton.
- FERRERO, F.E. et al. (1979). Fricative duration of Italian unvoiced fricatives. In *Frontiers of speech communication research*, B. Lindblom and S. Ohman, eds. New York: Academic Press, 159-65.

- FRY, F.B. (1955). Duration and intensity as physical correlates of linguistics stress. In *Journal of the Acoustical Society of America*, 27, 765-69.
- HEINZ, J.M. and K.N. Stevens (1961). On the properties of voiceless fricatives consonants. In *Journal of the Acoustical Society of America*, 33, 589-96.
- JAKOBSON, R., G. Fan and M. Halle (1952). Preliminaries to speech analysis, 8th ed. (1969). Cambridge, Ma.: MIT Press.
- JASSEM, W. (1968). Acoustical description of voiceless fricatives in terms of spectral patterns. In *Speech analysis & synthesis*, 1, 189-206.
- . (1965). The formants of fricative consonants. In *Language and speech* (Middlesex), 8, 1-16.
- KELM, O.R. (1989). Temporal aspects of rhythm which distinguish Mexican Spanish and Brazilian Portuguese, Ph.D. diss. (unpub.). Berkeley: University of California.
- KIM, J. and P.F. MacNeilage (1973). Voiceless consonant duration in VCV utterances. In *Journal of speech hearing research*, 17, 51-63.
- KLATT, D.H. (1974). The duration of [s] in English words. In *Journal of speech hearing research*, 17, 51-63.
- . (1973b). Interaction between two factors that influence vowel duration. In *Journal of the Acoustical Society of America*, 54, 1102-04.
- . (1976). Segmental duration in English. In *Journal of the Acoustical Society of America*, 59, 1208-21.
- . (1975a). Vowel lengthening is syntactically determined in a connected discourse. In *Journal of phonetics*, 3, 129-40.
- . (1975b). Perception of segment duration in sentence contexts. In *Structure and process in speech perception*, A. Cohen and S. Noteboom, eds. Heidelberg: Springer Verlag.
- KLOKER, D. (1975). Vowel and sonorant lengthening as cues to phonological phrase boundaries. In *Journal of the Acoustical Society of America*, 57, 533(A).
- KOZHEVNIKOV, V.A. and L.A. Chistovich (1965). Speech articulation and perception. *JPRS* 30. Washington, DC.
- LEHISTE, I. (1979). Perception of sentence and paragraph boundaries In *Frontiers of speech communication research*, B. Lindblom and S. Ohman, eds. New York: Academic Press, 191-201.
- . (1970). *Suprasegmentals*. Cambridge, Ma.: MIT Press.
- . (1973). Rhythmic units and syntactic units in production and perception. In *Journal of the Acoustical Society of America*, 54, 1228-34.
- LEHISTE, I. and G.E. Peterson (1961). Transitions, glides, and diphthongs. In *Journal of the Acoustical Society of America*, 33, 268-77.

- LIBERMAN, A.M., F.S. Cooper, D.P. Shankweiler, and M. Studdert-Kennedy (1967). Perception of the speech code. In *Psychology review*, 74, 431-61.
- LINDBLOM, B., B. Lyberg, and K. Holmgren (1981). *Durational patterns of Swedish phonology: do they reflect short-term motor memory processes?* Bloomington, Indiana: Indiana University Linguistic Club.
- LINDBLOM, B. and K. Rapp (1973). *Some temporal regularities of spoken Swedish*. Publication no. 21. Stockholm: Institute of Linguistics of the University of Stockholm (unpub.).
- LISKER, L. and A.S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. In *Word* 20, 384-422.
- LYBERG, B. (1979). Final lengthening - partly a consequence of restrictions on the speed of fundamental frequency change? In *Journal of phonetics*, 7, 187-96.
- . (1981). Some consequences of a model for segment duration based on F<sub>0</sub>-dependence. In *Journal of phonetics*, 9, 97-103.
- MAJOR, R. (1981). Stress-timing in Brazilian Portuguese. In *Journal of phonetics*, 9, 343-51.
- . (1985). Stress and rhythm in Brazilian Portuguese. In *Language*, 61, 2, 259-89.
- MENZERATH, P. and A. de Lacerda (1933). *Koartikulation, Steuerung und Lautabgrenzung*. Berlin.
- MILLER, G.A. and P.E. Nicely (1965). An analysis of perceptual confusion among some English consonants. In *Readings in acoustic phonetics*, 301-15.
- PARKER, E.M. and R.L. Diehl (1984). Identifying vowels in CVC syllables: effects of inserting silence and noise. In *Perception & psychophysics*, 36(4), 369-80.
- PETERSON, G.E. and I. Lehiste (1960). Duration of syllable nuclei in English. In *Journal of the Acoustical Society of America*, 32, 693-703.
- PIKE, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- SIMÕES, A.R.M. (1980). *Étude acoustique des consonnes [s] et [z] du Brésilien*. Mémoire de D.E.A., ms. Aix-en-Provence, France: Institut de phonétique.
- . (1990). Predicting sound segment duration in connected speech: an acoustical study of Brazilian Portuguese. In The First International Conference on Speech Synthesis, ETRW-GALF90, September 23-28. Autrans, France: Institut de la communication parlée.
- . (1987). *Temporal organization of Brazilian Portuguese vowels in continuous speech: an acoustical study*, Ph. D. diss. (unpub.). Austin, Texas: The University of Texas.

STEVENS, K.N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. in *Human communication: a unified view*, David E. and P. Denes, eds. New York: MacGraw-Hill, 51-66.

STREVENS, P. (1960). Spectra of fricative noise in human speech. In *Language and speech*, 3, 33-49.