

**CONSIDERATIONS ON SOME ASPECTS OF THE RELATIONSHIP
BETWEEN INTRINSIC AND EXTRINSIC TIME IN TWO
4-YEAR-OLD-CHILDREN'S AND ONE ADULT'S
SPEECH FOR DURATION IN BRAZILIAN PORTUGUESE**

**AGLAEL J.A. GAMA-ROSSI
(LIAACC/LAEL/ DEP. LINGÜÍSTICA/PUC-SP)**

RESUMO O trabalho tem como objetivo verificar as diferenças entre duas crianças de idade média de 4 anos e 5 meses e da professora delas na implementação do parâmetro acústico de duração. Crianças e professora são falantes do português brasileiro (PB) e foram gravadas num experimento de repetição de sentenças a partir de modelo fornecido pela pesquisadora. A duração tem se mostrado o principal parâmetro acústico na implementação do acento lexical no PB. Uma drástica redução dos segmentos acústicos em posição pós-tônica tem sido observada no contorno duracional da fala adulta. Entretanto, a literatura sugere, e os dados aqui apresentados vão nessa direção, que crianças abaixo de 6 anos ainda não reduzem tais segmentos como os adultos, principalmente devido ao processo, em curso, de maturação neuromotora. Os dados mostram como as duas crianças aqui estudadas implementam o contorno duracional do PB diferentemente da professora e também com diferenças entre elas, sendo que a questão é sempre como lidar com a redução dos segmentos acústicos nas posições não-acentuadas, a partir de suas capacidades neuromotoras quando da gravação. Com isso, põe-se em discussão a relação entre tempo intrínseco (ou da unidade dinâmica adotada na análise, o gesto articulatório) e tempo extrínseco (ou dos relógios no nível da sílaba e do sintagma ou frase entoacional que interagem com o tempo do gesto articulatório), considerando, para tanto, como as crianças implementam o parâmetro de duração no segmento, na palavra e no contorno duracional da sentença. É proposto que um relógio extrínseco no nível da frase é ajustado antes do relógio extrínseco no nível da sílaba, sem que seja necessário postular um relógio extrínseco no nível do acento lexical, uma vez que as palavras, principalmente na fala da criança menor aqui estudada, parecem estar sendo tratadas como frases entoacionais. Tudo isso pode ser acomodado num modelo dinâmico do ritmo como aquele proposto por Barbosa (2001).

1. INTRODUCTION

The aim of this paper is to discuss some findings relative to the implementation of the acoustic parameter of duration by two 4-year-old children and one adult, speakers of Brazilian Portuguese (Gama Rossi, 1999), in the light of dynamic models of speech production (Browman and Goldstein, 1989; 1990; Port, Cummins and McAuley, 1995, Barbosa, 2001). By collating data and some of the propositions of these theoretical models, we are mainly interested in finding out how children acquire or learn to deal with the relationship between the segmental and the prosodic sides of speech, or what, in dynamic models, designates respectively intrinsic and extrinsic time.

Duration has been considered the main acoustic parameter for implementing lexical stress in Brazilian Portuguese (Moraes, 1984; Massini, 1991; Barbosa, 1996; Gama Rossi, 1999). Intensity seems to be important too although, due to the difficulties concerning collection, measurement and normalization of intensity data, only duration has been studied in depth. We chose to study 4-year-old children because, at this age, children have acquired the greatest part of Brazilian Portuguese phonological distinctions, although sometimes lacking in some points relating to (1) the acquisition of the liquids, mainly the tap (/r/), which occurs in intervocalic position and, for some dialects, also in coda position, or (2) the acquisition of liquids /l/ e /r/ inside the consonantal clusters (made up by plosives and fricatives). Thus, 4-year-old children could be compared to adults for the most consonantal acoustic segments in the study of the acoustic parameter of duration.

2. METHODOLOGY AND BACKGROUND

We ran an experiment in which a *corpus* of 17 sentences was produced by the author and repeated three times by the subjects, two 4-year-old children (a child aged 4 years and 1 month (M) and a child aged 4 years and 9 months (E)), who belonged to the same kindergarten class in a school located in the city of São Paulo, and their teacher (S). The three subjects were São Paulo born females, as well as the author. We decided to use sentences, instead of words or minimal pairs inside carrier phrases (Smith, 1994), because sentences can be considered a minimal unit of sense, more easily repeated by children, from which many different linguistic and prosodic units (such as longer syllables, words/phonological words, noun phrases/intonational phrases, etc.) (Nespor and Vogel, 1986) can be detached, to be employed in the comparison of the subjects. All the sentences were relative to circus and they were linked to pictures, which described the content of each sentence. The author knew the link between sentences and pictures, so she could elicit each sentence from the corresponding picture. The aim was to avoid using the written form of the sentence in the picture, obviously because the children could not read, and so two different tasks, reading for the adult and repetition after model for the children, would be proposed. Pictures were shuffled before each presentation, so that sentence order was different at each repetition. The author tried to produce sentences at a self-comfortable speech rate, without focusing on any word. The models offered to the subjects for repetition were also recorded and the acoustic segmental duration of each sentence was measured as well, because we wondered if the model could be influencing the children's longer acoustic segments or, to put it in other words, we questioned whether the longer acoustic segmental durations presented in children's speech below the age of 6 (Kent and Forner, 1980) could be an artifact of the model offered to them, in a lower speech rate, since, as has been shown in the vast literature on motherese, the characteristics of adult speech directed to children are very distinctive from those of speech directed to other adults. Eventually, a Duncan test

comparing all the models offered by the author for repetition to the three subjects showed no significant difference between them.

Recordings were made in a silent, though not soundproof, room at the school, and technicians of a professional audio studio operated the equipment.

Duration measurements were made for all the acoustic segments in each sentence by using the software *CSRE 4.5*, with base on criteria set on sound waves and formant tracks generated for the digital speech signal.

In the analysis, we considered separately the different kinds of acoustic segments (such as vowel *vs.* consonant *vs.* constellation of acoustic segments similar to syllable length *vs.* constellation of acoustic segments similar to word length), in their respective stress positions, since these could stretch out or compress their intrinsic duration. Following the predictions for stress implementation in Brazilian Portuguese (Câmara, 1969), a drastic reduction is expected for the acoustic segments pertaining to the lexical post-stressed position. A pilot study (Gama-Rossi e Albano, 1996) showed that 4-year-old children could not reduce the acoustic segments in this position as adults did. Following this the acoustic segments of each sentence, corresponding to vowels, consonants, and syllables were demarcated for their position in the word: *pre-stressed position*, which was divided into 1 and 2, referring to the first and second pre-stressed syllables in the word, since we had 3 and 4-syllable words, in which stress was placed on the penultimate or the last syllable; *stressed position*, which included the highest prominence in the sentence, considered by Câmara (1969) as the last lexical stress on the right; and *post-stressed position* which mostly included just one syllable but that could include two syllables, as in words with 3 or more syllables, in which stress fell on the antepenultimate syllable. In the sentence the acoustic segments which belonged to the sentence's first and last syllables, which tend to be very stretched out (Klatt, 1976, Fougeron and Keating, 1997), were considered separately from those in the pre-stressed and post-stressed positions, respectively.

We made comparisons between the three pairs of subjects composed by the author and one of the three subjects: the child aged 4 years and 1 month (M), the child aged 4 years and 9 months (E), and their teacher (S). Inside each pair, subjects and the author were compared by means of the t test, as to the means and the standard deviations of the differences in duration, by considering each kind of acoustic segment (vowel, consonant, syllable) in each stress position (pre-stressed, stressed, post-stressed, beginning or end of the sentence), inside the same repetition (R1, R2 and R3). The mean of the differences in duration inside each pair was calculated according to the acoustic segmental duration in the subject's speech (M, E or S), minus the acoustic segmental duration in the author's speech (with control of the kind of segment, the stress position occupied by it, and the repetition). The hypothesis to be tested was whether the mean of the differences in duration would be statistically equal to zero. If it could, the subject and the author would have spent the same time in producing a given kind of acoustic segment in the same stress position.

3. RESULTS AND DISCUSSION

The Figures 1a to 1c show, in percentage, the amount of time that each acoustic segment of the sentence “A palhaçada dá muita risada”¹ (These clowns laugh a lot) occupies inside it, in each one of the three repetitions (R1, R2 and R3), for the three subjects: child M (4 years and 1 month), child E (4 years and 9 months) and their teacher (subject S).

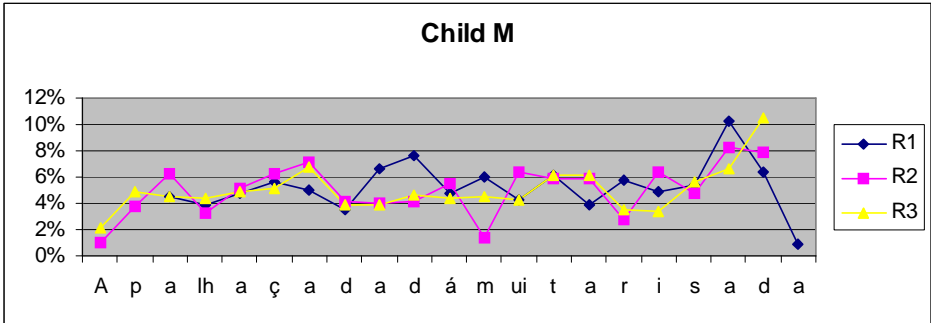


Figure 1a: Durational contours of the sentence “A palhaçada dá muita risada” for the subject M in percentage of the entire utterance. R1, R2 and R3 designate each repetition.

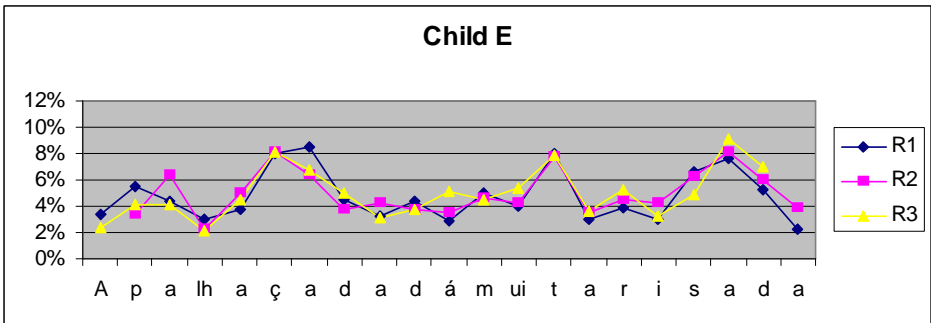


Figure 1b: Durational contours of the sentence “A palhaçada dá muita risada” for the subject E in percentage of the entire utterance. R1, R2 and R3 designate each repetition.

¹ The canonical phonetic transcription of the sentence “A palhaçada dá muita risada” is [apaʎa'sade'da'mũjtɛxi'zadẽ]. Note that in child M's speech the first “A” and the “p” of “palhaçada” in the first repetition (R1) could not be measured as well as the last “a” of “risada” in the second and in the third repetitions (R2 and R3, respectively). Also in child E's speech, the first “A” could not be measured in the second repetition (R2) as well as the last “a” of “risada” in the third repetition (R3).

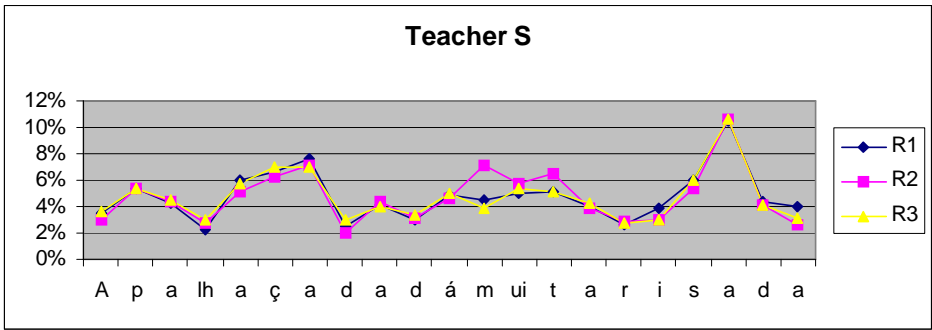


Figure 1c: Durational contours of the sentence “A palhaçada dá muita risada” for the subject S in percentage of the entire utterance. R1, R2 and R3 designate each repetition.

As we can see, the adult production is extremely regular, when comparing the three repetitions of the sentence. The only deviation observed on the acoustic segments of the word “*muita*” (a lot), in the teacher’s second repetition (R2), is due to an hesitation during her production of the /m/ of “*muita*”. The lesser regularity among children’s repetitions of the same sentence may be seen as a developing capacity of differentiating, tuning and mainly coordinating gestures in speech (Browman and Goldstein, 1989), which is not complete at the age of 4. It is worth noting that, although the age difference between the children is only 5 months, child E seems to be more advanced in controlling the acoustic segmental duration than child M. In what concerns the coordination or coarticulation of articulatory gestures (Folwer and Saltzman, 1993), we also observe that in child M’s speech this coordination is still precarious when we compare each pair (child M- author, child E-author, teacher S-author) by using the means and the standard deviations of the differences in duration for vowels, consonants, syllables and 2, 3, and 4-syllable words.

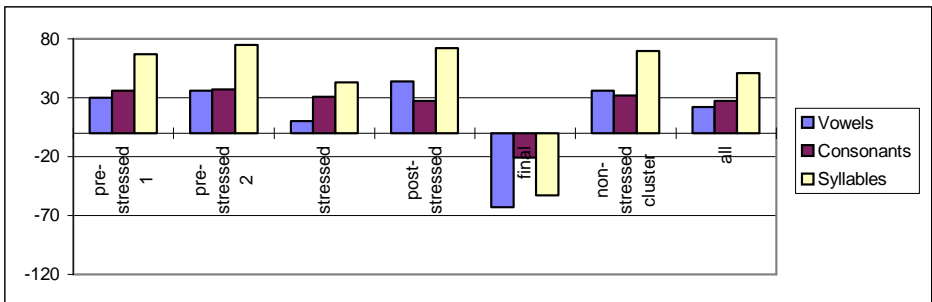


Figure 2a: Mean differences in duration (in milliseconds) for vowels, consonants and syllables, in all stress positions, for the pair child M-author.

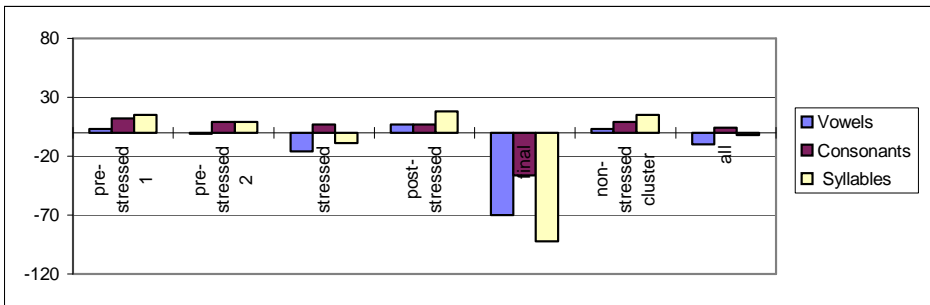


Figure 2b: Mean differences in duration (in milliseconds) for vowels, consonants and syllables, in all stress positions, for the pair child E-author.

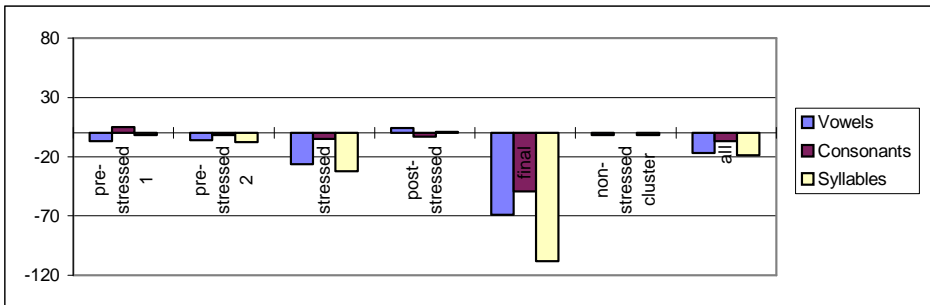


Figure 2c: Mean differences in duration (in milliseconds) for vowels, consonants and syllables, in all stress positions, for the pair teacher S-author.

In the Figures 2a to 2c, for each pair, the bars show the mean differences in duration for the acoustic segments (corresponding to vowels, consonants and syllables), in the different stress positions inside the word and in the last syllable of the sentence. The Figures 2a to 2c also consider the acoustic segments pertaining only to the non-stressed positions (pre-stressed 1, 2 and post-stressed positions altogether), since it is expected that acoustic segments in those positions must be reduced, because it is known that they constitute reference points to the rhythm (Fraisse, 1974, Barbosa and Bailly, 1994). We were interested in observing how the pairs child M-author, child E-author and teacher S-author would differ at non-stressed positions, since as we had seen before (Gama-Rossi and Albano, 1996), the acoustic segments belonging to those positions were not reduced in the 4-year-old children’s speech as it is in adult speech. Note that the cluster of each kind of acoustic segment in all stress positions is also presented as the last position in the Figures 2a to 2c.

In the Figure 2a referring to the pair child M-author, we can see that the bars go up in the positive direction, which means that the mean of the differences in duration of acoustic segments was longer in the child M’s speech than in the author’s speech for all kinds of acoustic segments and in all stress positions (exception made for the acoustic segments in the last syllable of the sentence and for the last syllable itself, which are consistently stretched out in the author’s speech inside the three pairs). Obviously,

longer acoustic durations in child M's speech were also observed for the acoustic segments at non-stressed positions (pre-stressed 1, 2, and post-stressed positions) and for the cluster of the acoustic segments when all stress positions were considered together.

Inside each sentence, syllable duration was calculated as the sum of the acoustic segments that made it up. Thus, the mean of the differences in duration for syllables, inside each pair, is the sum of the mean of the differences in duration for vowels and consonants. When we compare the three pairs productions of syllables, we can see that for the pair made up by the adults, we may assume a higher coordination between articulatory gestures (Browman and Goldstein, 1990) inside syllables, mainly for pre-stressed 1 and post-stressed positions, for which the mean of the differences in duration tends to zero (Table 1, below). The same can be said when we observe the cluster of the syllables that belong to all the non-stressed positions together.

Table 1: Results of the t test for the means and standard deviations of the differences in duration (in milliseconds) for syllables for the pair teacher S-author.

Stress Position	Number of Occurrence	Mean in ms	Standard Deviation in ms	t	P
First	51	14	28	3,4103	0,0013
Pre-Stress 1	98	-2	20	-1,114	0,2554
Pre-Stress 2	45	-8	15	-3,6663	0,0007
Stressed	172	-32	47	-8,9096	0,0001
Post-Stressed	112	1	19	0,4571	0,6485
Final	42	-108	59	-11,905	0,0001
Non-Stressed Cluster	258	-2	19	-1,8617	0,7478
All	523	-19	47	-9,2735	0,0001

*The significance level was previously fixed at 5%.

Table 2: Results of the t test for the means and standard deviations of the differences in duration (in milliseconds) for vowels for the pair teacher S-author.

Stress Position	Number of Occurrence	Mean in ms	Standard Deviation in ms	t	P
Pre-Stress 1	99	-7	14	-5,1808	0,0001
Pre-Stress 2	45	-6	12	-3,3222	0,0018
Stressed	175	-26	41	-8,6054	0,0001
Post-Stressed	113	4	16	2,4556	0,0156
Final	42	-69	33	-13,761	0,0001
Non-Stressed Cluster	260	-2	16	-2,4753	0,014
All	477	-17	35	-10,703	0,0001

*The significance level was previously fixed at 5%.

It is still worth noting that teacher and author disagreed significantly in regard to the implementation of vowel duration in all stress positions and clusters of stress positions,

as can be seen in Table 2. However this does not hold true for syllables, since no significant differences were found between teacher and author for the following stress positions: pre-stressed 1 ($p=0,2554$), post-stressed ($p=0,6485$) (see Table 1), cluster of non-stressed positions ($p=0,7478$).

Although this may not be sufficient to conclude in favor of a better coordination between articulatory gestures in adult speech than in children's, the observation of the standard deviation values for syllables in each stress position, inside each pair, shows a lesser value for vowels, consonants and syllables belonging to the adults to the same stress position for the children's speech (Table 3, below).

Table 3: Standard deviation values of the differences in duration (in milliseconds) for the acoustic segments corresponding to Vowels (V), Consonants (C) and Syllables (S) for the pairs child M-author, child E-author, teacher S-author

	Standard Deviation Values								
	Pair Child M-Author			Pair Child E-Author			Pair Teacher S-Author		
	V	C	S	V	C	S	V	C	S
First	-	-	41	-	-	28	-	-	28
Pre-Stress 1	28	37	45	26	27	38	14	19	20
Pre-Stress 2	33	26	50	25	24	34	12	13	15
Stressed	51	36	64	55	28	64	41	19	47
Post-Stressed	37	34	47	20	32	35	16	20	19
Final	67	62	63	41	45	60	33	51	59
Non-Stressed Cluster Without First	33	33	47	24	29	36	16	19	19
Non-Stressed Cluster Including First	-	-	51	-	-	35	-	-	22
All	48	41	61	44	33	56	35	28	47

Probably the higher accuracy in coordinating articulatory gestures in adult speech can be observed by noting that the standard deviation values of syllables do not grow markedly as compared to the values for vowels and consonants in the same stress positions, as they do for children.

The facts relative to the way in which the adults studied here had implemented vowel duration differences seem to signify that the vowel may be the major acoustic segment through which stress is reflected, and so the acoustic segment through which individual and stylistic differences can appear. On the other hand, at least in adult speech seen here, the syllables seem to be related to the rate at which articulatory gestures are produced, since, according to Barbosa (personal communication), the fact that there was no significant difference between the mean of the differences, in duration, for syllables belonging to pre-stressed 1 ($n=98$, $\text{mean}=-2\text{ms}$, $\text{standard deviation}=20\text{ms}$, $t=-1,144$, $p=0,2554$) and post-stressed positions ($n=112$, $\text{mean}=1\text{ms}$, $\text{standard deviation}=19\text{ms}$, $t=0,4571$, $p=0,6485$) may mean that the teacher has repeated the sentences at the same pace presented by the author if the succession of non-stressed

syllables is considered as a pacemaker. This coincides with the notion of an abstract speech rate given by a syllabic oscillator, as proposed by Barbosa (2001).

The precarious coproduction (Fowler and Saltzman, 1993) of the acoustic segmental duration in the children’s speech can be observed when whole words are considered. Notice that the values of duration for the pair made up by child E-author are as usual mid-way between those for the pair made up for child M-author and that for the pair made up by teacher S-author.

Table 4: Results of t test for the means and standard deviations of the differences in duration (in milliseconds) corresponding to words in non-final sentence position for the three pairs

Pair Child M-Author					
Word extension	Number of occurrences	Mean in ms	Standard Deviation in ms	t	p
2 syllables	15	87	84	4,0294	0,0012
3 syllables	12	174	110	5,4836	0,0002
4 syllables	10	213	64	10,576	0,0001
Pair Child E-Author					
2 syllables	25	27	75	1,8035	0,0839
3 syllables	20	53	78	3,0218	0,007
4 syllables	17	7	75	0,3669	0,7185
Pair Teacher S-Author					
2 syllables	26	-17	36	-2,3307	0,0282
3 syllables	31	-15	46	-1,8095	0,0804
4 syllables	20	-13	41	-1,4453	0,1647

*The significance level was previously fixed at 5%.

Although we are using acoustic duration to refer to the differences between children’s and adults’ speech in what concerns the differentiation, tuning and coordination of articulatory gestures (Browman and Goldstein, 1989), we know that acoustic duration and articulatory gestures are only indirectly related, since acoustic duration is the output of the formation and the release of a given constriction that takes place in the three-dimensional space of the vocal tract, which also takes time to occur, the so called articulatory gesture, a dynamic unit, because it has an intrinsic time or duration. The articulatory gesture is itself the input to the model of speech production in Articulatory Phonology (Browman and Goldstein, 1990). However, since gesture’s formation and release can be modeled as critical trajectories of a first-order mass-spring model, a higher stiffness (k parameter) of the spring will produce a lower displacement of it, as well as an output of reduced acoustic durations (Browman and Goldstein, 1985). So, hypothetically, we assume that the longer acoustic durations observed in child M’s speech, mainly for the acoustic segments that belong to the non-stressed positions, could be thought of as a matter of a developing neuromotor capacity that cannot yet change the stiffness of the tissues involved in forming and releasing a given constriction inside the vocal tract. However, we can observe that the neuromotor capacity of child

M during the data collection was already interacting with the linguistic knowledge of a competent Brazilian Portuguese hearer. When we consider all the different stress positions for vowel, in the pair made up by child M-author, the lesser mean of the differences in duration was met for the stressed position, as can be seen in Table 5, below.

Table 5: Results of the t test for the means and the standard deviations of the differences in duration (in milliseconds) for vowels for the pair child M-author

Stress Position	Number of Occurrence	Mean in ms	Standard Deviation in ms	t	p
Pre-Stress 1	94	30	28	10,32	0,0001
Pre-Stress 2	44	36	33	7,2181	0,0001
Stressed	158	10	51	2,558	0,0115
Post-Stressed	85	44	37	10,784	0,0001
Final	17	-63	67	-3,8484	0,0014
Non-Stressed Cluster	226	36	33	16,229	0,0001
All	401	22	48	9,0601	0,0001

*The significance level was previously fixed at 5%.

Since the author shows higher means in the difference in duration for the stressed vowels in the pairs made up by child E (n= 173, mean=-16ms, standard deviation=55ms, t=-3,775, p=0,0002) and the teacher (n=175, mean=-26ms, standard deviation=41ms, t=-8,6054, p=0,0001), we can see that, when in the position of not being able to produce reduced non-stressed acoustic segments, child M had very much stretched the vowel corresponding to lexical stress (which in the analysis here includes the higher prominence of the sentence) and, in doing this, she created peaks of stress, probably offering the listener a duration contour more similar to that of Brazilian Portuguese (as can be seen in Figure 1a). Of course, the relationship between the stiffness of the spring (an abstract oscillation and its implementation through the mass-spring equation, proposed by Browman and Goldstein, 1990) and the supposed stiffness of the tissues involved in forming and releasing and constriction inside the vocal tract are very different things, which do not have, as yet, a neurological model that can put them together.

Gesture stiffness (k parameter) in Articulatory Phonology can also be changed by the intervention of the rhythmic tier (Browman and Goldstein, 1990), which is put above the gestural score (the score of the gestures that compose a given lexical item) and works like an extrinsic clock, thus changing the intrinsic time of the constellations of gestures corresponding to a syllable. Accordingly, we can theorize that the constellations of gestures under primary stress have a lower stiffness than those under secondary stress (or belonging to non-stressed positions), since gestures with a lower stiffness can be more stretched out than those with a higher stiffness. Thus, the gestures under primary stress will present a longer duration in the acoustic output. So, it can be

assumed that the only extrinsic clock proposed by Articulatory Phonology (Browman and Goldstein, 1990) refers to the syllabic level, since it is defined by the vowels' onsets.

In a model for Brazilian Portuguese rhythm, Barbosa (2001) has recently proposed two coupled adaptive oscillators (Port, Cummins and Gasser, 1995), defined by Port et al. (op. cit.) as extrinsic clocks that can adjust their frequency of oscillation to the frequency of the input and that present some tolerance to input variation, since they can adapt to it in just a few input cycles. Adaptive oscillators can be distributed among the different levels of speech production and perception. Thus, in Barbosa's model, two coupled adaptive oscillators interact with the gestural score, one for the phrasal stress and the other for the syllable-sized units, the latter controlling the speech rate. Now we would like to evaluate Articulatory Phonology and Barbosa's proposals for the interaction between the extrinsic and the intrinsic time with base on the two 4-year-old children's speech data presented here.

The data presented above make it difficult to talk about an extrinsic clock at the syllable level since, as we saw - to a greater degree for child M's speech and lesser to child E's speech - these two children seem not to have constituted this extrinsic clock or an adaptive oscillator at the syllable level. The reasons for this assumption are that, in the children's speech, the syllable and also the word duration may vary highly, depending on the number of the acoustic segments they possess. Therefore, based on the data of children M's and E's speech, we can assume that an extrinsic clock at the syllable level is not constituted before the age of 4. However, those children are in fact implementing the Brazilian Portuguese durational contour, using mainly a stress for words or for phrases, a fact that can be observed in Barbosa's proposal of an extrinsic clock or adaptive oscillator at the phrasal level.

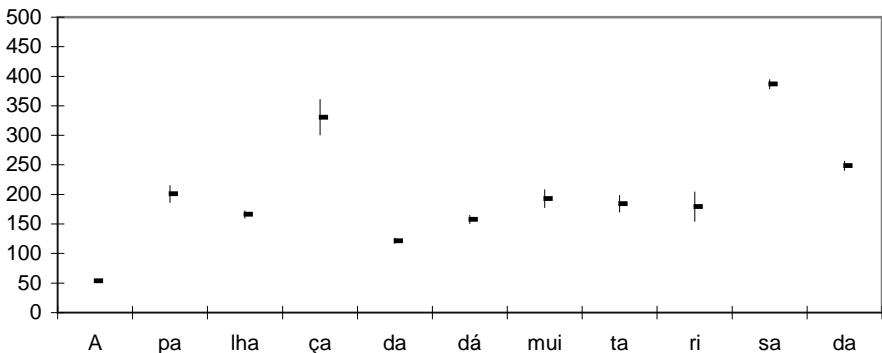


Figure 3a: Means and standard deviations (in milliseconds) of the models (three repetitions) offered by the author to child M's repetitions of the syllables in the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

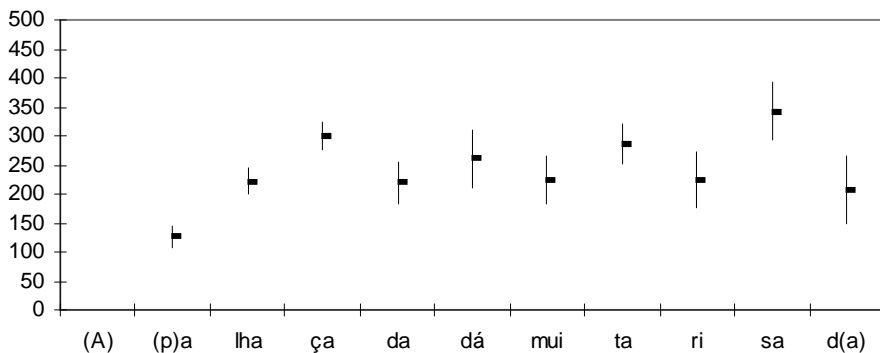
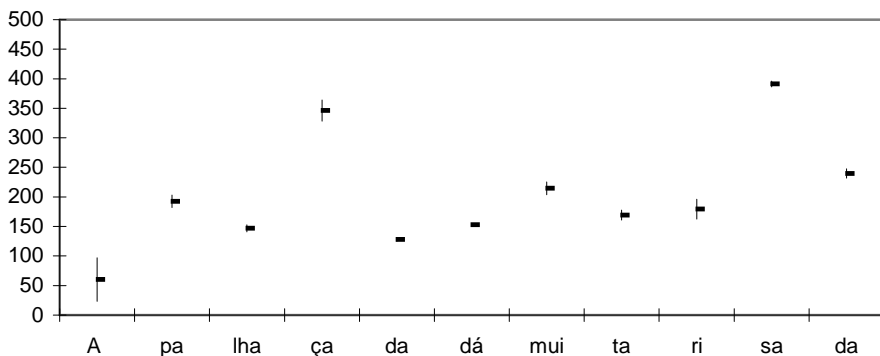


Figure 3b: Means and standard deviations (in milliseconds) of child M's three repetitions of the syllables in the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

In examining the above durational contours, made up of the means and the standard deviations for syllable duration values, both calculated for the three repetitions of the sentence "A palhaçada dá muita risada" (These clowns laugh a lot), and considering child M's speech and the models offered to her, we can observe that a word stress seems to be at work on child M's speech, since peaks can be seen on palhaÇAda², DÁ, muiTA (notice that here "TA" is actually the post-stressed syllable) and riSAda, although the last lexical stress on the right, on riSAda, presents a mean duration higher than the others, which could indicate that the child is making there the higher prominence of the sentence or putting focus on riSAda.



² Capital letters are marking the stressed syllable.

Figure 3c: Means and standard deviations (in milliseconds) of the models (three repetitions) offered by the author to child E's repetitions of the syllables in the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

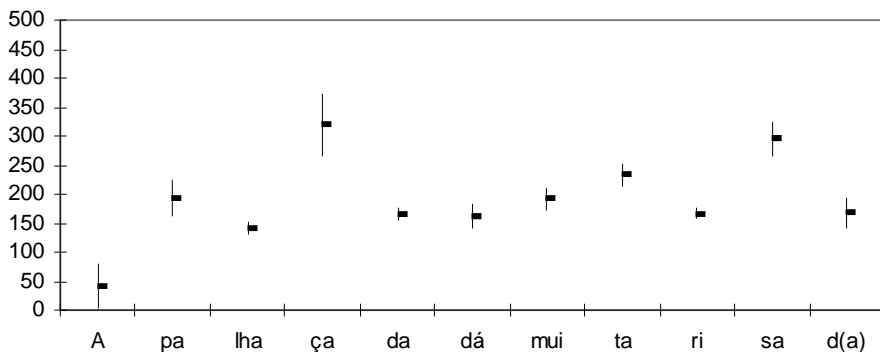


Figure 3d: Means and standard deviations (in milliseconds) of child E's three repetitions of the syllables in the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

In child E's speech, as we can see in the graph above, a phrasal stress seems to be at work, since the mean duration of the stressed syllable "ÇA" (or [sa] de "palhaÇAAda") and the mean duration of the last stressed syllable "SA" (or [za] in "riSAda") are almost at the same level (even though the mean of "ÇA" ([sa]) is a little higher, probably through the influence of the unvoiced fricative; also, its larger standard deviation may mean less accuracy in the production of the fricative). So, it seems that child E has divided the sentence in at least two phrases or intonational phrases: "A palhaçada" (noun phrase) "dá muita risada" (verbal phrase). Furthermore, because of the higher mean duration of the post-stressed syllable "ta" ([ta]) of "muita" (a lot), we could also wonder whether a third intonational phrase would be "dá muiTA" (actually the verb "dá" + the adverb "muita" pertaining to the next adverbial phrase: AvP[muita risada]), since it forms a rising, but lower, durational contour between the two major stresses on "palhaÇAAda" and "riSAda". The higher mean duration observed in the post-stressed syllable "ta" (muita), in both children's durational contour, seems to indicate that they introduced a pause between "muita" and "risada", probably focusing on "muita" (a lot) because of its semantic content. The last syllable acoustic segments of "muita" may have been stretched out. Thus, to explain the data of these two children's acquisition of the implementation of the duration acoustic parameter in Brazilian Portuguese, it seems necessary to propose an oscillator to the phrasal level, as Barbosa (2001) has done.

Actually we could assume that this oscillator is acquired before that for the syllable level, which probably requires, as seen in the adult pair data, a higher accuracy of the segmental duration. This can be again observed in the Figures 3e and 3f for the adults' productions of the same sentence, seen below.

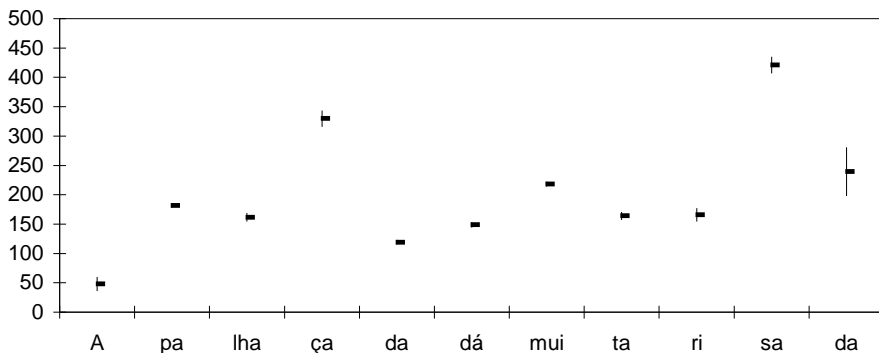


Figure 3e: Means and standard deviations (in milliseconds) of the models (three repetitions) offered by the author to teacher S's repetitions for the syllables of the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

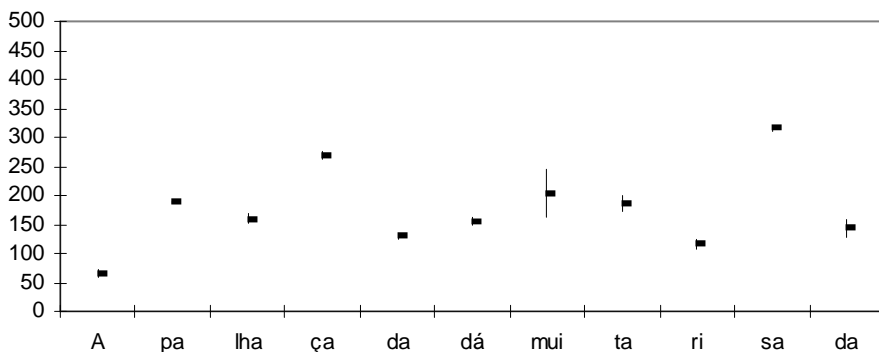


Figure 3f: Means and standard deviations (in milliseconds) of teacher S's three repetitions of the syllables in the sentence "A palhaçada dá muita risada" (These clowns laugh a lot)

In the Figures 3e and 3f, referring to the adults' production, we can see a contour on the verbal phrase "dá muita risada" with ascending peaks of stress, on "DÁ" (verb) "MUIta" (adverb) and "riSAda" (noun) respectively, the latter being ("SA" or [za] in "risada") the higher peak on the durational contour sentence for both adults. We can also observe that teacher S may have focused on the word "muita", because of the higher standard deviation presented by the post-stressed syllable "ta" in her speech, when compared to the author's speech. However, differently from the children, she did not stress "ta" more than "MUI", the canonical stressed syllable on the word "MUIta".

4. CONCLUSION

One might question whether it is necessary to also propose an adaptive oscillator at the word level. It does not seem to be necessary, firstly because Barbosa (2001) has already shown that two coupled adaptive oscillators, one at the phrasal and the other at the syllable level, can afford emerging pauses and differences in speech rate, and secondly because we can also assume that child M, at the prosodic level, may be treating words as phrases (or phonological words as intonational phrases) (Nespor and Vogel, 1986), mainly because of her difficulty in reducing non-stressed elements in speech, which leads her to constitute more peaks of stress.

In conclusion, the data here presented has shown how 4-year-old children, speakers of Brazilian Portuguese, can create, in accordance with their neuromotor possibilities, the differences of the Brazilian Portuguese durational contour, when they implement the acoustic durational parameter. In what concerns the relationship between intrinsic (articulatory gesture) and extrinsic time, it seems necessary to propose, for the children studied here, an adaptive oscillator or an extrinsic clock mainly at the phrasal level, which sustains Barbosa's proposal (2001).

ACKNOWLEDGEMENTS My heartfelt thanks to Dr. Eleonora Cavalcante Albano, Dr. Plínio Almeida Barbosa and Paulo Rehder, the latter for the statistical analysis. I also thank Prof. Ângela Levy for the revision and editing of the English version. This work was sponsored by CNPq, process number 140564/94-2.

REFERENCES

- BARBOSA, P.A. & BAILLY, G. (1994). Characterisation of rhythmic patterns for text-to-speech synthesis. *Speech Communication*, v.15, p.127-37.
- BARBOSA, P.A. (1996). At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation. *Cadernos de Estudos Lingüísticos*, v.31, p.33-53.
- _____. (2001). Generating duration from a cognitively plausible model of rhythm production. In: *Proceedings of the Seventh European Conference on Speech Communication and Technology (Eurospeech 2001)*. Aalborg, Denmark, September 3-7, v.2, 967-970.
- BROWMAN, C. & GOLDSTEIN, L. (1985). Dynamic Modeling of Phonetic Structure. In: FROMKIN, V. A. (ed.) *Phonetic Linguistics Essays in honor of Peter Ladefoged*. Orlando, Academic Press.
- _____. (1989). Articulatory gestures as phonological units. *Phonology*, v.6, p.201-51.
- _____. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: KINGSTON, J.; BECKMAN, M. E. (eds.) *Papers in laboratory phonology I. Between the grammar and physics of speech*. Cambridge, Cambridge University Press.
- CÂMARA, J.M. (1969). Problemas de lingüística descritiva. 2.ed. Petrópolis, Vozes.
- FOUGERON, C. & KEATING, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, v.101, p.3728-40.

- FOWLER, C. & SALTZMAN, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, v.36, p.171-95.
- FRAISSE, P. (1974). *La psychologie du rythme*. Paris, Presses Universitaires de France.
- GAMA-ROSSI, A.J.A. & ALBANO, E.C. (1996). Possible phonological influences on segment duration differences between adults and children: a pilot study. Poster presented at the *Fifth Conference on Laboratory Phonology (LabPhon)*. Evanston, Northwestern University.
- _____. (1999). Relações entre desenvolvimento lingüístico e desenvolvimento motor: A aquisição da duração no português brasileiro. Unpublished thesis. Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- KLATT, D. (1976). H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, v.59, p.1208-21.
- MASSINI, G. (1991). *A duração no estudo do acento e ritmo do português*. Unpublished dissertation. Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- KENT, R. D. & FORNER, L.L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics*, v.8, p.157-68.
- MORAES, J.A. (1984). *Recherches sur l'intonation modale du portugais brésilien parlé à Rio de Janeiro. Analyse acoustique, perceptive et fonctionnelle*. Unpublished thesis. Université de la Sorbonne Nouvelle Paris III, Paris.
- NESPOR, M. & VOGEL, I. (1986). *Prosodic Phonology*. Dordrecht, Foris Publications.
- PORT, R.; CUMMINS, F. & GASSER, M. (1995). A dynamic approach to rhythm in language toward a temporal phonology. In: DAINOR, A.; HEMPHILL, R.; LUKA, B.; NEED, B. and PARGMAN, S. (eds.) *Papers from the 31th Meeting of the Chicago Linguistic Society*. Chicago, Chicago Linguistic Society.
- SMITH, B. (1994). Effects of experimental manipulations and intrinsic contrasts on relationships between duration and temporal variability in children's and adult's speech. *Journal of Phonetics*, v.22, p.155-75.