CLINICAL LINGUISTICS & PHONETICS, 2000, VOL. 14, NO. 8, 587–601

Acoustic characteristics of trill productions by groups of Spanish children

GLORIA CARBALLO and ELVIRA MENDOZA

Department of Personality, Evaluation and Psychological Treatment, University of Granada, Spain

(*Received 15 January 2000; accepted 15 June 2000*)

Abstract

Contrary to the English /r/, which has been extensively analysed, there is very little information on the trilled consonants in Spanish. These sounds are in general difficult to produce for young Spanish children and occur later (than other consonant sounds) in normal development. This paper describes acoustic measurements made on the trill productions in Spanish children with varying degrees of speech intelligibility. The spectral (frequency, amplitude, C–V ratio) and temporal characteristics (duration of the trill, number of open and closed periods—apertures and occlusions—and duration of these periods) were studied in 45 children (Granada, Spain) between 3.0 and 9.6 years old, divided into 5 groups. The results reveal differences in spectral and temporal acoustic correlates of trill /r̄/ among the five speaker groups (e.g., the duration of the first aperture period was longer for the articulatory disordered group than for the normal control group). They seem to indicate, in particular, that children who have trouble learning to make the trill sound do little more than tap the alveolar ridge. Moreover, these children seem to use more of their tongues to make these gestures; also, they do not make many trills or make them very quickly.

Introduction

The acquisition of data on the acoustic correlates of speech sounds in children across languages is very important for the understanding of phonological development. The aim of the present study is the analysis of the trilled sounds by means of a comparison of spectral and temporal measures of this sound as produced by different groups of Spanish children.

The acoustic study of the English /r/ has received the attention of researchers for many years (e.g. Delattre, 1965; Klein, 1971; Powers, 1971; Dalston, 1975; Strange and Broen, 1981; Sharf and Benson, 1982; Hoffman *et al.*, 1983; Sharf and Ohde, 1984; Lindau, 1985; Ohde and Sharf, 1988; Ohde *et al.*, 1989; Powell, Elbert and Dinnsen, 1991; Espy-Wilson, 1994; Espy-Wilson and Boyce, 1994; Shuster,

Address correspondence to: Gloria Carballo, Departamento de Personalidad, Evaluación y Tratamiento Psicológico, Facultad de Psicología, Universidad de Granada, 18071 Granada, Spain. E-mail: gloriac@platon.ugr.es

Clinical Linguistics & Phonetics
ISSN 0269-9206 print/ISSN 1464-5076 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals

1996; Boyce and Espy-Wilson, 1997; Espy-Wilson and Boyce, 1999). To the best of our knowledge, similar analyses have not been performed with the Spanish /r̄/. Spectrographic analyses of these sounds have concentrated on the specification of its average spectral characteristics in adults (Quilis, 1981; Massone, 1988; Mota, 1990; Recasens, 1991; Martinez-Celdrán, 1994). In regard to the temporal characteristics, these same studies have been limited to the analysis of the number and average duration of the periods of apertures and occlusions.

The phoneme /r̄/ is difficult for children to produce. It is acquired late in language development (Bosch, 1983; Jimenez, 1987; González, 1989), being frequently substituted by other sounds. It is also a phoneme which demands much clinical attention in the daily practice of language therapy. The acoustic characteristics of the trill /r̄/ in the productions of children differing in age and production abilities are interesting to explore because they could provide useful information concerning general development of speech motor control, and of the problems encountered by children with unintelligible speech. In fact, the results of this particular investigation are important not only for models of motor control in normal and disordered speech (Kent, 1999) but also for phonological descriptions of Spanish.

The trilled nature of these sounds comes from the short interruptions of the flow of air as it passes through the linguo-alveolar constriction. The Spanish trill /r̄/ consists typically of three rolls/trills in the initial syllable (Alarcos-Llorach, 1986; Navarro-Tomás, 1990; Quilis and Hernández, 1990; Martinez-Celdrán, 1994). During each contact of the tongue with the alveoli, the airflow is momentarily interrupted, resulting in a rapid series of short explosions (Navarro-Tomás, 1990). The term used to refer to this trilled sound is 'trill', but it does not have a direct component in the English phonological system, and it does not seem appropriate to refer to it as a multiple 'tap' (Recasens, 1991).

The sonogram of a Spanish trill produced by adults shows that it presents successive trilled movements formed by periods of closure or silence (interruptions), and by periods of opening/aperture, in which formants are observed. A typical example is given in figure 1, which shows a spectrogram of the word '*rata*' (rat) spoken by a native Spanish female speaker. The interruptions or periods of silence are also usually described as occlusions because they correspond to the articulatory contacts between the apex of the tongue and the alveoli. Often the vocalic elements, which correspond to regions of formants clearly differentiated between two occlusions, are simply called apertures because they correspond to the moments of separation of the tongue apex and the alveoli.

In this work, we have studied the spectral and temporal acoustic characteristics of the trill /r̄/ in five groups of children in Granada described below in order to see if the values of these variables are similar in all the groups and/or if the values can be compared with those in adults, given the lack of similar data in the child population. This work complements and extends the perceptual analyses of the intelligibility of the /r̄/ of Carballo *et al.* (1997). Its emphasis is not so much on the phonetic/linguistic nature of the consonant but on the developmental progress on the acquisition of the /r̄/ by Spanish children. Moreover, we want to contribute to the topic concerning whether or not children with speech disorders fail to produce some segments properly because their underlying representations are flawed or because they just plain can not move their mouths the right way to produce them. This work seems to suggest the latter explanation, at least for trilled /r̄/.

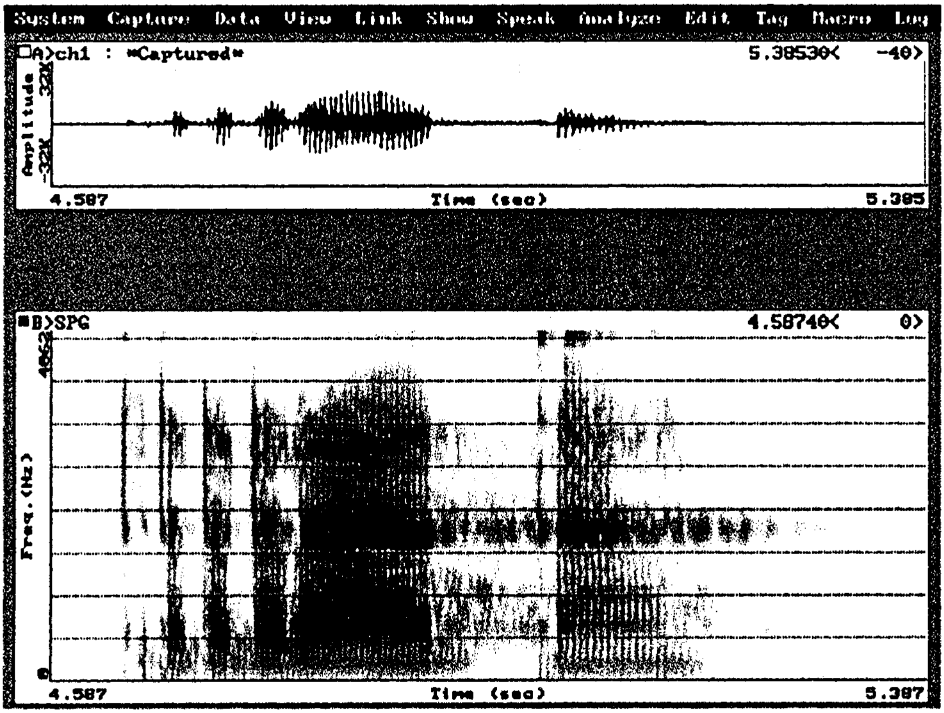

Figure 1. *Spectrogram of the word* 'rata' *(rat) by a native spanish female speaker.*

Method

Participants

Five experimental groups with nine subjects each participated in the study (see table 1). The 45 children were attending school centres in Granada. Twenty-seven of the subjects (13 boys and 14 girls) were between 3.0 and 6.6 years old. This age range was selected based on descriptive studies (Bosch, 1984; Serra, 1984; González, 1989; Carballo, 1995) in which it was shown that children within this range are in the process of acquiring and developing the trill sound. Thus, some of these children articulate the /r̄/ normally (group 1), and others have limited errors (group 2 and 3). For this reason, it was possible to establish groups of developing acquisition of /r̄/, as it is detailed in the description of the groups in table 1. The fact that at these

Table 1. *Distribution of subjects in the experimental groups*

	Group 1 HI	Group 2 MI	Group 3 LI	Group 4 IP	Group 5 CG
No. boys	5	5	3	5	4
No. girls	4	4	6	4	5
Total No.	9	9	9	9	9
Aver. age	485	478	446	770	808

HI = High Intelligibility of /r̄/; MI = Medium Intelligibility of /r̄/; LI = Low Intelligibility of /r̄/; IP = Incorrect Pronunciation; CG = Control Group.

ages a child does not correctly articulate the trill sound does not mean that his/her speech is defective or that clinical intervention is necessary, but that the child is in the process of normal development. The remaining 18 subjects (nine boys, nine girls) were between 7.0 and 9.6 years old. These subjects have already acquired the Spanish trill in a distorted (group 4) or normal (group 5) way. A minimun chronological age of seven years was chosen because the sound errors of the children younger than this may due to developmental factors.

The five subject groups were formed as follows:

Group 1 (High Intelligibility of /r̄/. HI): nine subjects between 3.0 and 6.6 years old who pronounced the trill normally.

Group 2 (Medium Intelligibility of /r̄/. MI): nine subjects between 3.0 and 6.6 years with a incompletely trilled production.

Group 3 (Low Intelligibility of /r̄/. LI): nine subjects between 3.0 and 6.6 with a production of /r̄/ similar to an intervocalic /r/ (i.e. the tap).

Group 4 (Incorrect Pronunciation. IP): nine subjects between 7.0 and 9.6 years old who produced a distorted /r̄/. There were no subjects in this group who substituted another sound for that studied here.

Group 5 (Control Group. CG): nine subjects between 7.0 and 9.5 years old whose pronunciation of the /r̄/ was totally normal.

The assignment of a subject to groups 1–3 were based on an auditory perceptual criterion of intelligibility of the consonant production: the accuracy of the children's articulation of the trill sound according to their productions of the consonant for the articulation stimuli (Carballo *et al*., 1997). Additionally, the following criteria were observed when the subjects of the different groups were selected: (i) Spanish was the only language spoken in the home; (ii) the subjects in the first three groups had *not* been previously diagnosed as having deficiencies in speech or language; (iii) the subjects of the IP group presented incorrect production of /r̄/ in the first syllable of the six words (and correct production of all the other phonemes) and no child was receiving treatment for any language problem, and (iv) the subjects in CG correctly produced the trill. The analyses involved judgements by 12 experienced speech-language pathologists on the auditory perception of the intelligibility of /r̄/ for the subjects. In particular, they indicated the differentiation of the first three aforementioned groups according to a rating scale in which the initial syllable of the word [r̄a] was completely incorrect (0) or correctly pronounced (10). The full range of scores (0 to 10) could be used, depending on the listeners judgement of intelligibility. See Carballo *et al.* (1997) for further details.

Stimuli

The following six target words were used as stimuli: *rabo* (tail), *rama* (branch), *rayo* (ray), *rana* (frog), *rata* (rat), *raya* (line). These stimuli had two characteristics: (i) words in current linguistic use found in the children's basic vocabulary (Casanova and Rivera, 1989), and (ii) words with a bisyllabic structure that starts with [r̄a]. The material was presented to each child by means of drawings on separate laminates. Each child was asked to name each of the six stimuli. No child needed stimulation to produce the words.

Measurements and procedure

270 words produced by the five groups of children were divided and measured (54 words per group). An acoustic analysis of the stimuli was performed, and the following variables were studied:

Spectral measures

(i) Frequency (in Hz) and amplitude (in dB) of the first three formants.

(ii) Consonant–vowel intensity ratio (C–V ratio), which refers to the difference in decibels between either the power or the energy of a consonant and that of an adjacent vowel (Freyman and Nerbone, 1989). This measure is usually taken in order to explain differences in intelligibility among talkers. In addition, with vowel level held constant, C–V intensity ratio determines the absolute level of the consonants.

The formants were measured by hand since formant trackers do not generally work well on children's speech. The formant frequencies were measured across several apertures when the children produced the trilled /r̄/ correctly.

Temporal measures

(iii) Consonant duration, corresponding to the time in ms from the beginning of the sound production of the consonant to the start of the transition to the vowel /a/.

(iv) Number of occlusions (interruptions) and number of apertures of the phoneme /r̄/.

(v) Duration of the first period of aperture (i.e. the aperture between the first and second occlusions) of the sonographic signal of the phoneme /r̄/.

Durational measurements have been argued to be used to characterize a child's developmental progress in attaining adultlike speech motor control (Kent and Forner, 1980). Moreover, developmental patterns in the control of duration are a necessary substrate for research on the acquisition of phonological processes (Smith, 1978).

The recordings were made in an isolated room of the school centers. Each participant spoke into a high-quality microphone, placed 20 cm in front of the child. The productions of all subjects were recorded on a digital audio-tape using a Sony 77Es Cassette-recorder, and they were subsequently processed and acoustically analyzed using a KAY Sonograph, model 5500 (frequency range: 0–8.0 kHz, transform size: 100 points, bandwidth: 300 Hz, window type: Hamming).

For each production, a broad-band spectrogram, the waveform and a visual display of the average power spectrum of the acoustic signal were simultaneously viewed and analysed. The consonant boundaries were fixed on visual examination of the waveform and the spectrogram, and on auditory cues made available through playback of the segments over a loud speaker. Segmentation was performed through placement of adjustable cursors on the monitor. The end of the consonant was identified with the beginning of the transition to the vowel [a]. The duration of the consonant was taken from the begining of the first period of aperture in all cases, whichever the number of apertures of the sonographic signal of the sound [r]. The first aperture was taken to start following the first occlusion. A similar process was done to determine the corresponding characteristics of the vowel [a] and to obtain the differences (in dB) of the maximal amplitudes of the consonant and the vowel.

Experimental analysis

The following analyses were performed:

- ANOVA for a between-groups unifactorial design to study if there were significant differences in the corresponding variable among the five subject groups (HI, MI, LI, IP and CG).
- Posteriori Analysis of comparisons (Tukey test, p set at 0.05). Tukey Honestly Significant Difference (HSD) *post hoc* tests were performed at the 0.05 level of significance to determine which groups differed significantly among themselves in the given variables.

Results

Spectral variables

Formant Frequencies F1, F2, and F3
Table 2 shows the means (in Hz) and SDs for the five groups of children for each of the spectral variables: frequency of the first formant (F1), of the second (F2), and of the third (F3). Comparisons of the acoustic characteristics of the /r̄/ stimulus produced by the five subject groups revealed significant mean differences for F1 measures ($F(4,40) = 6.354$: $p < 0.001$). Results of the Tukey HSD *post hoc* tests

Table 2. *Mean, SD and significance level of the values of the following analysed variables in the five children groups: formant frequencies (F1, F2, F3), in Hz, formant amplitudes (AF1, AF2, AF3), in dB, C–V ratio, consonant duration (Conson. duration), in ms, number of apertures (No. of apertures), number of occlusions (No. of oclusions) and duration of the first period of aperture (1st. per. ap. dur.)*

Variable	M and SD	G1 (HI)	G2 (MI)	G3 (LI)	G4 (IP)	G5 (CG)	Significance
F1	M	622.59	620.00	604.45	519.99	517.04	$p < 0.001$
	SD	15.55	30.24	33.61	11.59	15.31	
F2	M	1953.70	1966.67	2144.44	2074.81	1957.03	$p < 0.05$
	SD	34.32	31.14	66.03	21.49	36.18	
F3	M	3898.50	3941.91	4168.89	3912.67	3542.96	$p = 0.72$
	SD	115.75	213.12	115.54	44.99	127.62	
AF1	M	36.24	37.78	36.03	38.91	41.22	$p < 0.01$
	SD	1.33	1.93	2.09	0.68	0.79	
AF2	M	48.52	49.80	50.35	52.37	47.72	$p = 0.07$
	SD	2.12	0.79	2.27	1.33	1.17	
AF3	M	62.54	62.80	63.65	62.44	64.26	$p = 0.25$
	SD	1.73	1.79	1.19	1.60	1.73	
C–V ratio	M	4.37	4.55	4.39	6.67	8.39	$p < 0.01$
	SD	1.22	1.52	1.05	1.38	0.98	
Conson. duration	M	124.17	127.83	100.33	163.50	115.67	$p < 0.01$
	SD	12.76	14.27	7.89	10.60	3.88	
No. of apertures	M	3.57	2.41	1.07	1.40	3.31	$p < 0.001$
	SD	0.41	0.38	0.05	0.12	0.16	
No. of oclusions	M	2.63	1.87	0.20	0.41	2.31	$p < 0.001$
	SD	0.43	0.47	0.13	0.13	0.16	
1st per. ap. dur.	M	20.02	41.38	94.38	141.60	21.90	$p < 0.001$
	SD	1.68	9.98	18.16	8.41	0.67	

indicate that the greatest differences in the F1 variable occurred in the comparison of each of the evolving/developing groups of intelligibility of /r̄/—high, medium and low—with the groups of older children.

The means and SDs of the F2 measures, described in table 2, show significant differences ($F(4,40) = 3.335$: $p < 0.05$). The largest differences are found between the low intelligibility of /r̄/ group and the control group. *Post hoc* Tukey HSD comparisons predominantly showed that the LI group had a significantly larger mean compared to CG; this means that F2 was significantly higher for the LI group. No significant F ratios were obtained for the scores in the variable F3.

Formant amplitudes AF1, AF2 and AF3
The values of means (in dB) and SDs for the variable amplitude of the first formant AF1 are shown in table 2. Significant differences were seen in the variable AF1 among the groups ($F(4,40) = 4.307$: $p < 0.005$). Results of the Tukey HSD *post hoc* tests indicate that the largest differences in AF1 occurred between group 5 (CG) and the groups HI and LI. ANOVAs conducted with the data of the variables AF2 and AF3 showed no significant differences.

Consonant–vowel (C–V) ratio
The means and SDs of the different subject groups in this variable are shown in table 2. Significant differences are seen ($F(4,40) = 4.323$: $p < 0.005$) between groups. In particular, results of the Tukey HSD *post hoc* tests indicate that the greatest differences in C–V ratio appear between the control group and the three developing groups of intelligibility of the consonant. Moreover, the means in the C–V ratio are highest in the control group, almost twice as high as in the three groups of younger children. Finally, we find minimal differences for this variable among the three evolving/developing groups.

Temporal variables

Consonant duration
Table 2 presents the data from the analysis of this variable. It can be seen that (i) the highest mean values were reached by group 4 (IP) for the six stimuli used and were much higher than in the other groups, and (ii) the subjects who presented a shorter duration of /r̄/ belonged to the group 3 (LI). The duration of the consonant /r̄/ established significant differences among the 5 groups of subjects ($F(4,40) = 5.094$: $p < 0.01$). In the a posteriori analysis, it was found that the most notable differences were between group 4 (IP) and the groups 3 (LI) and 5 (CG).

Number of apertures and number of occlusions
The results of the analysis performed for number of periods of aperture and of occlusion are also presented in table 2. The groups that best pronounced the /r̄/ (groups 1 and 5) have a similar mean number of aperture periods, around 3. This number is lower in the group 2 (MI) in the stimuli, and the number is close to 1 in the groups whose pronunciations are farther away from the trill /r̄/. Elevated significant statistical differences were obtained among all the groups ($F(4,40) = 18.649$: $p < 0.001$). The HSD comparisons revealed that the groups 1 (HI) and 5 (CG) presented the greatest differences in the number of periods of aperture compared with the groups that pronounced the consonant worst (groups 3 and 4).

Regarding the number of occlusions, this variable is related to the periods of aperture and reults in the same scores in relation to the distribution in the different groups. Highly significant differences were obtained among the five groups ($F(4,30) = 8.805$: $p < 0.001$). Not all the subjects in groups 3 (LI) and 4 (IP) and even in group 2 (MI), presented periods of occlusion with the given stimuli. The comparisons in the Tukey test revealed the greatest differences between group 3 (LI) and groups 1 (HI) and 5 (CG).

Duration of the first period of aperture

Average values in the last temporal variable for all the stimuli were lowest in the groups that pronounced the /r̄/ best, groups 1 and 5, and were five times as great in the groups that pronounced the /r̄/ worst, groups 3 and 4 (see table 2). The duration of the first aperture period was significantly different between groups ($F(4,40) = 34.197$: $p < 0.001$). *Post hoc* comparisons showed that these differences existed between all groups, except between the groups 1 (HI) and 2 (MI) with group 5 (CG).

The first occlusion period data were not analysed because this period did not occur in every subject nor with every stimuli. This means, in particular, that some subjects only had one occlusion.

Discussion

The three groups of developing acquisition of /r̄/, HI, MI and LI show both acoustic and perceptual characteristics (Carballo *et al.*, 1997) that are clearly different.

Spectral variables

The results obtained for the variable frequency of the first formant (F1) showed the lowest values in the control group, 517.0 Hz (average age 8 years old), and the highest values, 622.6 Hz, for the HI group (3.5 to 6.6 years old). The groups of developing acquisition of /r̄/ (HI, MI, LI) which have different levels of intelligibility, have similar F1 values. On the other hand groups IP and CG, which are supposedly at either extreme of the intelligibility scale but similar in age, cluster together and show practically equal F1 values.

If these values are compared with those found in the literature about the Spanish trill in adults, Massone (1988) points out that F1 appears in frequential values around 500 Hz; Mota (1990) who differentiates between speech in the laboratory and continuous speech of an adult male, showing values somewhat lower for the frequency of this formant, 470.9 Hz and 496.7 Hz. To continue with works in Spanish, we mention that Quilis (1981) demonstrates a value of 557 Hz for F1 although it is not specified the kind of speakers (children, adults).

The results in the present study fit within the general tendency shown in the literature in English that the values of F1 frequency are higher in children than in adults (e.g. Eguchi and Hirsh, 1969; Fant, 1970; Kent, 1976; Kent and Forner, 1980; Pentz, Gilbert and Zawadski, 1979; Kent and Murray, 1982; Nittrouer, Studdert-Kennedy and McGowan, 1989, etc.), which is related to differences in the size of the vocal tract.

The lower values for the frequency of the second formant (F2) come from the group 1 (HI), 1957.0 Hz. The highest values come from the group 3 (LI), 2144.0 Hz.

The following values for the Spanish trill /r̄/ for adults in frequency of the second formant are: 1193 Hz (Quilis, 1981); 1089.9 Hz in laboratory speech and 1240.8 Hz in spontaneous speech (Mota, 1990); and between 1200 and 1600 Hz, varying according to the vowel (Massone, 1988).

This difference in the values of the frequencies of F2 found in samples of children and adults, in which higher values were found in children and decreased with age (Eguchi and Hirsh, 1969; Fant, 1970), have already been pointed out in studies of the English language. According to the data in our study, however, the difference in the values of F2 can be better related to the correct or incorrect pronunciation of the phoneme /r̄/, given that, although these differences are highly significant between the five subject groups, the lowest values were found in the two groups that produced the /r̄/ with the highest clarity. The highest values were reached in the groups that are in the other extreme, with deficient articulation of the phoneme. The incorrect pronunciations of certain speakers may be due to a tongue shape that is incompatible with the aerodynamic requirements for a trill. The higher F2 for the less proficient trillers could come from a more fronted tongue body, which has been hypothesized to impede trill production, or to limit the number of occlusions as Ladefoged and Maddieson (1996: 221) have pointed out. We support that F2 is associated with how fronted the tongue body is and F1 is associated with how close the tongue body is to the palate. We think that F1 is more an indicator of whether the tongue blade and tongue apex is raised or not.

The mean values for the frequency of the third formant, (F3) vary between 3542.9 Hz for the control group and 4168.9 Hz for the group 3 (LI). The groups of developing acquisition of /r̄/ show for this variable a pattern similar to the corresponding values seen in the second formant. However, the data for the F3 variable were not significantly different among the groups in this study. This is probably due to the fact that the young children have a greater difficulty in emitting the trill or simply their recordings have non sufficient acoustic power and spectral bandwidth to define three resonant frequencies. This phenomenon has also been observed in the production of the liquid sound /r/ and the glide /w/ in American English (Hoffman *et al.*, 1983). It should be noted that the data from our study cannot be compared with any data from other studies on the Spanish /r̄/, given that the only reference to the variable frequency of the third formant comes from Massone (1988: 23), 'for the two masculine adult speakers, the third formant is found in the region of 2200 Hz'. Our results show some values that are higher than those of her study due to the use of a child sample in the present study.

Regarding the formant amplitudes, the results of the analysis allow us to affirm that the amplitude of the first formant (AF1) is significantly different in all groups (see table 2). These differences are found between the control group 5 (CG) and the groups 1 (HI) and 3 (LI). On the other hand, given the present results, it cannot be concluded that the amplitudes of the second and third formant (AF2 and AF3) depend on the speaker groups. The control group showed the highests values in the amplitudes of formants F1 and F3. The mean values of the amplitudes of the second and third formants increased progressively in the groups 1, 2 and 3 (high, medium and low intelligibility), although their ages were similar. This parameter could explain part of the differences that exist between the three developing groups in their pronunciation of the trill /r̄/. Another possible explanation would be that the very low acoustic power of a child's recording, clearly shown for the variable F3, may account for the values relative to the amplitude of the formants.

These data seem to show that the F1 amplitude can indeed be used as a significant variable to differentiate among groups that incorrectly pronounce the trill sound. However, the amplitudes of the other two formants, F2 and F3, cannot do this.

Due to the scarcity of literature about the variable formant amplitude that exists in Spanish, it is difficult to compare the data from our study with others in Spanish children and adults. Mota (1990) states that spectral differences exist between continuous and laboratory speech, and the former type of speech favours the concentration of energy at high frequencies. However, the data in this variable should be considered with some caution as there are not available comparative references. Further studies would be needed on this topic, with more control over this variable and with data on continous speech.

The variable consonant–vowel ratio (C–V) is in direct relation with the variables of amplitude aforementioned, and thus, the above considerations can be applied here. In accord with the present results, this variable significantly differentiated between groups. The control group presented values that were almost double to those of the developing groups. However, the same as was mentioned regarding the variable AF1 of the consonant, the lack of comparative data on the Spanish /r̄/ in the literature, applies here. It can be supposed from the data that the greatest difference in amplitude of the consonant–vowel in the control and the incorrect pronunciation groups with respect to the three developing groups could be in part related to age.

The spectral measures of our study, specially the F2 and F3 measures, suggest that children who do not produce the trill /r̄/ correctly (groups 3 and 4) use a contriction with a different place and/or shape. Moreover, the pattern that we see is that the children who can not produce satisfactory trill sounds seem to use more of their tongues to make these gestures. As a result, they do not make many trills or make them very quickly. These conclusions are partly because both F2 and F3 measures are higher for the poor trillers than for the good trillers. For narrow constrictions it has been provided (Fant, 1970) good evidence that F2 is a resonance of the cavity in front of the constriction. If both these formant frequencies are higher in the trill, both cavities are smaller, and so the constriction itself must be longer (i.e. in all likelihood it is specifically the shape of the constriction that is different, not the place). Therefore, we conclude the poor trillers are trying to use more tongue. Probably because of this approach, the poor trillers do not have as many occlusions as the good trillers, and they have a long aperture duration. In addition, our data shows that the C–V ratio is low in all the non-controls, regardless of their mastery of the sound.

Temporal variables

Examination of the temporal parameters relative to the initial consonant duration has revealed important differences in the five subject groups and the two following results deserve mention: (i) the children in Group 3 (LI) produced the shortest tokens (mean = 100.3 ms) since they consisted of only one occlusion, and (ii) the children in Group 4 (IP) produced the longest tokens (mean = 163.5 ms).

When interpreting these results, it must be remembered that several factors intervene in the production of the consonant: (i) the number of occlusions and apertures (the duration increases with the occlusion and aperture periods), (ii) the age of the subject (on average, the duration decreases with age), and (iii) the presence or absence of deficiencies in the subject's speech (the duration is greater in subjects with poor /r̄/ articulation).

The control group (average age of 8.08 years, table 2) presents durations lower than the groups 1 (HI) and 2 (MI) (average age of 4.85 years and 4.78 years, respectively) of the consonant. This agrees with the tendency observed in normal subjects who have lower durations with increasing age, as pointed out by many authors (Kent and Forner, 1980; Kent, Netsell and Abbs, 1979; Kubaska and Keating, 1981; Rimac and Smith, 1984; Robb and Saxman, 1990; Smith, 1978; 1992; 1994; Smith *et al.*, 1983).

Nevertheless, the group 3 (LI), similar in age to the other two developing groups (groups 1 and 2), showed durations of the consonant considerably shorter. This can be due to a purely articulatory effect, derived from the earlier mentioned fact, according to which the duration depends on the number of occlusions and vocalic elements of the corresponding sound in the production of the consonant. The children in group 1 produced the trill by successive movements of occlusion and aperture, usually three or four, while those of group 2 (MI) produced it with some fewer periods and variability. Lastly, the children in group 3 (LI) produced the /r̄/ with one sole aperture, and, thus, its duration is notably shorter (100 ms).

In regard to the group 4 (IP), the results suggest that in children who misarticulated the /r̄/, production was characterized by consonants of longer duration (163 ms) than in the control group. This difference may be due to the poorer motor control of the children with bad articulation. This is the same line of discussion that Weismer and Elbert (1982) and Maxwell and Weismer (1982) follow when explaining the greater variabilility of the children who erroneously articulate /s/ in comparison with normals.

An absolute comparison reveals that all the types of segments produced by the children in group 4 (IP) were of a mean duration between 35.7 and 63.2 ms greater than those of the four other groups. This longer duration is quite distinct, perhaps, as mentioned before, because the children with an age between 7 and 9.6 years in this study already have sufficiently fixed erroneous speech patterns. The same does not occur in the group 3 (LI), due to the same explanation previously expounded. The /r̄/ of the third group is a sound closer to the intervocalic /r/, with one sole aperture period, though of longer duration, but whose consonant duration has mean values that are smaller than in the rest of the groups.

In the case of the Spanish trill sound in adults with which the comparison is made, it can be highlighted that the mean values of duration of the Spanish /r̄/ vary between 58.6 and 103 ms depending on the number of periods of occlusion and aperture (Mota, 1990; Quilis, 1981; Recasens, 1991).

The mean temporal values found in the present study, though in complete accord with those found by other researchers, are greater in the children and, above all, are much greater in the group of incorrect pronounciation of /r̄/. One possible interpretation is that the children who incorrectly pronounced the phoneme might have been less precise than the children who articulate normally because of the more deficient motor control. Let us mention that, there may exist perceptual processes in the children with normal speech development and with deficient development that could explain in part these differences, a concept that should be studied in depth.

Finally, some authors (e.g Eckert and Eichorn, 1977; Surwillo, 1977; Smith 1978; Kent and Forner, 1980) found that the children differ in duration variability with respect to adults, and this may reflect different abilities of motor control of speech.

If these differences are studied either within or between age groups, one might think that some children who incorrectly pronounce /r̄/ might have speech motor control that is less precise than in adults, and that this would be the cause of their deficient articulation; for example, difficulties in tongue control at the moment of the vibrations of the tongue tip against the alveolar ridge. If the erroneously articulated speech is a manifestation of a developmental delay or a deficient motor behaviour (Kent, 1976), the deficit could be reduced or eliminated as the child progresses through maturation.

The results from the variables number of occlusions and number of apertures, seen in table 2, demonstrate that both variables were significantly different between groups, varying between an average of 3.57 and 1.07 apertures per group, and between 2.63 and 0.20 occlusions. The two variables depend totally on the mode of articulation of the consonant. The groups that pronounced the trill /r̄/ best (groups 1 and 5) have the highest number of both periods. This number is more constant for all the stimuli in the control group, due to a more automatic pronunciation of the consonant. For this reason, the production of the phoneme is more similar to that of the /r̄/ by adults.

In an inverse manner, the groups 3 and 4 that pronounced the /r̄/ more diffusely, articulation-wise, usually presented one unique very long aperture period, or two at the most, but of different sizes, the second being much longer in duration. Concerning the periods of occlusion, the subjects in these two groups either did not present occlusions or presented one sole period.

Let us briefly compare the present data with those existing for adults in Spanish (Quilis, 1981; Massone, 1988; Recasens, 1991; Mota, 1990). Massone (1988) found a mean of three aperture periods in the initial syllable; Quilis (1981) found three occlusions or interruptions and two apertures; Recasens (1991) talks about an intermediate aperture and two closures, and Mota (1990) indicates that laboratory speech can be produced with one or two apertures and two or three occlusions, depending on whether the consonant has three or five periods. The children who pronounced the trill /r̄/ in the present work presented on average three or four apertures and two to three occlusions. In the case where the /r̄/ did not reach vibration of the trill, this number was signficantly reduced in the groups studied.

The most interesting observations from our results of the variable duration of the first aperture are the following: (i) its high statistically significant difference between the groups of incorrect and correct pronounciation, as seen in table 2, and (ii) the fact that the children with defective pronounciation of /r̄/ can be characterized by having an aperture period of a duration notably longer than the normal-speaking children.

Regarding the factor of age, it can not be definitively stated that all the children in this study pass through the three stages of intelligibility of /r̄/ during the acquisition process of the phoneme. Looking at the three developing groups (HI, MI, and LI), it is seen that all are of similar ages, 3.0–3.6 the youngest and 6.0–6.6 years, the oldest and thus the lower intelligibility of /r̄/ cannot be associated with the younger ages of the children. Yet, it can be related to the greater or lesser motor control of speech precision (Kent, 1976) that some children reach before others. To test the possibility that all children go through the three periods in the learning process of /r̄/ would require performing an extensive longitudinal study.

In the literature on the duration of the first period of aperture in Spanish-speaking adults, one finds a duration that wavers between 15 and 20 ms (Quilis, 1981; Mota, 1990; Recasens, 1991). The values in the present study with normal-speaking children are a little higher, 20.0 ms in group 1 (HI) and 21.9 ms in the control group. As the /r̄/ is progressively pronounced less intelligibly, the duration becomes considerably greater (MI, LI, and IP, see table 2). Other studies dealing with the acoustic and temporal development of the phoneme in children along the period of acquisition in the Spanish language, with which to compare data do not exist to the best of our knowledge.

The most remarkable result of our study concerns the trills themselves. The temporal measures show that there is a great difference between those subjects that produced a trilled /r̄/ and those who did not. Sometimes between age 3 and 6.6, Spanish speaking children normally acquire adult-like trill production. This conclusion is based on the findings that the numbers of occlusions and apertures is the same for the young, normal group as for the older normal group. Moreover, the duration of the first aperture is similar for the two normal groups.

Finally, we would like to highlight the following finding: children who have trouble learning to make the trill /r̄/ simply can not do the trill. Children in both the young, impaired and the older, impaired group do little more than tap the alveolar ridge. The older, impaired children seem to be trying to include more trills, but are not very adept at doing so. One indication of this, according to our acoustic measurements, is the inordinately long aperture duration.

In summary, this study describes acoustic measurements made on the trill productions of Spanish children differing in age and production abilities. This sound was choosen because it is one of the most difficult sounds to correct according to the clinicians reports and because of its clinical frequency. Although we have two ages (under 6.6 and over 7.0 years) and two general ability levels (normal and impaired), we have classified the 45 subjects of our study into five groups (see table 1) to gain more insight and more detailed information about the phonological development. Indeed, this investigation may provide insight as to what characterizes disordered speech (deficits in underlying representations or deficits in speech motor control). Our results point to differences in the production of the trill /r̄/ across age groups and levels of intelligibility for the sound.

References

Alarcos-Llorach, E., 1986, *Fonología Española* (Madrid: Gredos).

Bosch, L., 1983, El desarrollo fonológico infantil: Una prueba para su evaluación. *Anuario de Psicología*, **28**, 87–114.

Boyce, S. and Espy-Wilson, C. Y., 1997, Coarticulatory stability in American English /r/. *Journal of the Acoustical Society of America*, **101**, 3741–3753.

Carballo, G., 1995, *Estudio de las adquisiciones fonológicas. Análisis acústico de la evolución del fonema /r̄/*, unpublished doctoral dissertation (University of Granada: Spain).

Carballo, G., Mendoza, E. and Valencia-Naranjo, N., 1997, Interobserver agreement of perceived intelligibility of /r̄/ in children. *Perceptual and Motor Skills*, **84**, 1099–1104.

Casanova, A. and Rivera, M., 1989, *Vocabulario básico de EGB* (Educación General Básica) (Madrid: Espasa Calpe).

Dalston, R. M., 1975, Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. *Journal of The Acoustical Society of America*, **57**, 462–469.

Delattre, P., 1965, *Comparing the Phonetic Features of English, French, German and Spanish: an interim report* (Heilderberg: Julius Groos Verlag).

Eckert, H. M. and Eichorn, R. E., 1977, Developmental variability in reaction time. *Child Development*, **48**, 452–458.

Eguchi, S. and Hirsh, I. J., 1969, Development of speech sounds in children. *Acta Oto-Laryngologica* (Suppl.), **257**, 5–48.

Espy-Wilson, C. Y., 1994, A feature-based semivowel recognition system. *Journal of the Acoustical Society of America*, **96**, 65–72.

Espy-Wilson, C. and Boyce, S., 1994, Acoustic differences between 'bunched' and 'retroflex' variants of American English /r/. *Journal of the Acoustical Society of America*, **95**, 2823(A).

Espy-Wlson, C. and Boyce, S., 1999, A simple tube model for American English /r/. *Proceedings of the International Congress of Phonetic Sciences*, 2137–2140.

Fant, G., 1970, *Acoustic theory of speech production* (The Hague: Mouton).

Freyman, R. L. and Nerbonne, G. P., 1989, The importance of consonant-vowel intensity ratio in the intelligibility of voiceless consonants. *Journal of Speech and Hearing Research*, **32**, 524–535.

Gonzàlez, M. J., 1989, Análisis del desarrollo fonológico en sujetos malagueños. *Infancia y Aprendizaje*, **48**, 7–24.

Hoffman, P., Stager, S. and Daniloff, R. G., 1983, Perception and production of misarticulated /r/. *Journal of Speech and Hearing Disorders*, **48**, 210–215.

Jimenez, B. C., 1987, Acquisition of Spanish consonants in children aged 3–5 years 7 months. *Language, Speech and Hearing Services in Schools*, **18**, 357–363.

Kent, R. D., 1999, Motor control: Neurophysiology and functional development. In A. J. Caruso and E. A. Strand (Eds) *Clinical Management of Motor Speech Disorders in Children* (New York: Thieme).

Kent, R. D., 1976, Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. *Journal of Speech and Hearing Research*, **19**, 421–447.

Kent, R. D. and Forner, L. L., 1980, Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics*, **8**, 157–168.

Kent, R. D. and Murray, A. D., 1982, Acoustic features of infant vocalic utterances at 3, 6 and 9 months. *Journal of The Acoustical Society of America*, **72**, 353–365.

Kent, R. D., Netsell, R. and Abbs, J. H., 1979, Acoustic characteristics of dysarthria associated with cerebellar disease. *Journal of Speech and Hearing Research*, **22**, 627–648.

Klein, P. R., 1971, Acoustic analysis of the acquisition of acceptable r in American English. *Child Development*, **42**, 543–550.

Kubaska, C. A. and Keating, P. A., 1981, Word duration in early child speech. *Journal of Speech and Hearing Research*, **24**, 615–621.

Ladefoged, P. and Maddieson, I., 1996, *The Sounds of the Worlds Languages* (Oxford: Blackwell).

Lindau, M., 1985, 'The story of /r/'. In V. Fromkin (Ed.) *Phonetic Linguistics: Essays in honor of Peter Ladefoged* (New York: Academic Press).

Martinez-Celdràn, E., 1994, *Fonética* (Barcelona: Teide).

Massone, M. I., 1988. Estudio acústico y perceptivo de las consonantes nasales y líquidas del español. *Estudios de Fonética Experimental*, **III**, 13–34.

Maxwell, E. M. and Weismer, G., 1982, The contribution of phonological acoustic and perceptual techniques to the characterization of a misarticulating child's voice contrast. *Applied Psycholinguistics*, **3**, 29–43.

Mota, C. de la, 1990, A study of [r] in spontaneous speech. *Proceedings of the XII International Congress of Phonetic Sciences*, **4**, 386–389.

Navarro-Tomàs, T., 1990, *Manual de Pronunciación Española* (Madrid: CSIC).

Nittrouer, S., Studdert-Kennedy, M. and Mcgowan, R. S., 1989, The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research*, **32**, 120–132.

Ohde, R. N., McCarver, M. E. and Sharf, D. J., 1989, Perceptual and accoustic characteristics of distorted /r/. *Journal of The Acoustical Society of America* (Suppl.), **86**, S48–S49.

Ohde, R N. and Sharf, D. J., 1988, Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children and /r/-misarticulating children. *Journal of Speech and Hearing Research*, **31**, 556–568.

Pentz, A., Gilbert, H. R. and Zawadski, P., 1979, Spectral properties of fricative consonants in children. *Journal of The Acoustical Society of America*, **66**, 1891–1893.

Powell, T. W., Elbert, M. and Dinnsen, D. A., 1991, Stimulability as a factor in the phonological generalization of misarticulating preschool children. *Journal of Speech and Hearing Research*, **34**, 1318–1328.

Powers, M., 1971, Functional disorders of articulation: Symptomatology and etiology. In L. Travis (Ed.) *Handbook of speech pathology and audiology* (New York: Appleton-Century-Crofts).

Quilis, A., 1981, *Fonética Acústica de la Lengua Española* (Madrid: Gredos).

Quilis, A. and Hernàndez, C., 1990, *Lingüística Española Aplicada a la Terapia del Lenguaje* (Madrid: Gredos).

Recasens, D., 1991, On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics*, **19**, 267–280.

Rimac, R. and Smith, B. L., 1984, Acoustic characteristics of flap productions by American English-speaking children and adults: implications concerning the development of speech motor control. *Journal of Phonetics*, **12**, 387–396.

Robb, M. and Saxman, J., 1990, Syllable durations of preword and early word vocalizations. *Journal of Speech and Hearing Research*, **33**, 583–593.

Serra i Raventos, M., 1984, Normas estadísticas de articulación para la población escolar de 3 a 7 años del área metropolitana de Barcelona. *Revista de Logopedia, Foniatría y Audiología*, **4**, 232–235.

Sharf, D. J. and Benson, P. J., 1982, Identification of synthesized /r-w/ continua for adult and child speakers. *Journal of The Acoustical Society of America*, **71**, 1008–1015.

Sharf, D. J. and Ohde, R. N., 1984, Effect of formant frequency onset variation on the differentiation of synthesized /w/ and /r/ sounds. *Journal of Speech and Hearing Research*, **27**, 475–479.

Shuster, L. I., 1996, Linear predictive coding parameter manipulation/synthesis of incorrectly produced /r/. *Journal of Speech and Hearing Research*, **39**, 827–832.

Smith, B., 1978, Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics*, **6**, 37–67.

Smith, B. L., 1992, Relationships between duration and temporal variability in children's speech. *Journal of The Acoustical Society of America*, **91**, 2165–2174.

Smith, B. L., 1994, Effects of experimental manipulations and intrinsic contrasts on relationships between duration and temporal variability in children's speech. *Journal of Phonetics*, **22**, 155–175.

Smith, B. L., Sugarman, M. D. and Long, S. H., 1983, Experimental manipulation of speaking rate for studying temporal variability in children's speech. *Journal of The Acoustical Society of America*, **74**, 744–749.

Strange, W. and Broen, P. A., 1981, The relationship between perception and production of /w/, /r/, and /l/. *Journal of Experimental Child Psychology*, **31**, 81–102.

Surwillo, W. W., 1977, Developmental changes in the speed of information processing. *Journal of Psychology*, **8**, 468–482.

Weismer, G. and Elbert, M.., 1982, Temporal characteristics of 'functionally' misarticulated /s/ in 4- to 6-year old children. *Journal of Speech and Hearing Research*, **25**, 275–287.