UNIVERSITAT DE BARCELONA

Estudios de Fonética Experimental
Journal of Experimental Phonetics

LF Laboratori de Fonètica

Acoustic analysis of rhotics in coda position in five- to seven-year-old children

Jordi Cicres[a], Sílvia Llach[b]

[a] Universitat de Girona (Spain), jordi.cicres@udg.edu
[b] Universitat de Girona (Spain), silvia.llach@udg.edu

ARTICLE INFO

Article history
Received: 03/07/2020
Accepted: 08/03/2021

Keywords
Rhotics
Coda position
Children's speech
Acoustics
Catalan

ABSTRACT

This article studies the acoustic production of central Catalan rhotics in a group of 90 boys and girls between the ages of five and seven, in medial and final coda. An acoustic description is provided based on an analysis of the number and type of components of both the closing and opening phases. Based on combinations of different components, up to 78 different rhotic types were detected. The results show that most rhotic sounds have one or two components, and that rhotics with five or more components are almost always at the end of a word. There is variability with respect to acoustic properties, with a higher presence of vocalic components in the medial coda and more fricative or occlusive components in the final coda.

1. Introduction

1.1. Acquisition processes and rhotics

In general, acoustic research has shown that, compared to adult speech, children's speech presents a higher pitch, higher formant frequencies, longer segment durations, and more spectral variability (Eguchi and Hirsh, 1969; Kent, 1976; Kent and Forner, 1980; Smith, 1978, 1992; Smith, Kenney and Hussain, 1995; Hillenbrand et al., 1995). In some cases, these differences continue well into adolescence. For example, Lee, Potamianos and Marayanan (1999) acoustically analysed some American English segments and short sentences in children and adolescents aged between five and 17 years old and found that some of the parameters analysed, such as relative amplitude and duration, converged with adult values at age 12, but other variables, such as fundamental frequency and formant frequency variability, did not converge until the participants were 14 or 15 years of age. With respect to prosodic variables, Aoyama, Akbari and Flege (2016) found that children's speech (in children aged seven to 14 years) was slower and higher pitched when absolute values were compared, but when they analysed the values proportionally (i.e. by comparing semitones instead of hertz, or based on rhythm), the productions were similar.

Detailed descriptions of the acoustic characteristics of, firstly, parameters relating to the frequency, amplitude and duration of individual sounds and, secondly, prosodic parameters (intonation, rhythm, etc.) are critical to understanding the chronology and development of the organs involved in children's verbal production and motor control. A thorough description should therefore make it possible to reproduce a precise vocal tract developing model at different ages of verbal development. On a practical level, this model would prove useful in various applied linguistics fields. For example, in forensic linguistics, it could be helpful in voice comparison with child speakers or when determining an unknown speaker's age and sex from a recording; in the field of automatic speech recognition, such models are essential if the goal is to achieve levels of success comparable to those of adult speech; and in the field of speech synthesis, it could serve to synthesize the voices of children of different ages with greater accuracy. However, few acoustic studies have focused on this age group, at least in relation to Spanish and Catalan. Notable studies relating to English have been conducted by Lee, Potamianos and Narayanan (1999), McGrowan, Nittrouer and Manning (2004), and Klein et al. (2013), who all focused on this chronological stage but did not compare the results with those of adult speech. Only Dalston (1975) compared the spectral structure and temporal variables of rhotics in word-initial position in adults and children; in this case, however, the children were three and four years old.

This lack of acoustic studies is due to the fact that most focus on the phonological aspects of sound system acquisition. These studies provide data to suggest that initial acquisition of the phonological system occurs at four years of age (Ingram, 1976; Ingram et al., 1980; Oller, 1980; Stark, 1980; Grunwell, 1981; Haelsig and Madison, 1986; Vihman, 1996; Roberts, Burchinal and Footo, 1990, among others). This can be partly explained by the significant anatomical changes that occur in the first few years of life, which, in terms of phonation, affect primarily the morphology of the tongue, lips and palate, thereby transforming the human resonating chambers (Caplan, 1993). This is also the case with Catalan: Bosch (1987), De Ribot (1992) and Llach (2007) carried out naming or repetition tests to analyse how phonological errors are distributed between the ages of three and seven and how these evolve during this period of development. In doing so, they were able to establish the order in which individual sounds are acquired. For example, all three authors agreed that the consonant sounds acquired first (and which therefore present fewest errors) are nasals, followed by glides, occlusives and approximants. By contrast, the sounds that present most errors are laterals, followed by trills and fricatives (Llach, 2007: 142). Bosch (1987: 24) found that, in Catalan, the trill sound is acquired by more than 90% of children at age seven (palatal laterals, voiced prepalatal affricates, voiced fricatives and /s/ present similar or lower percentages of correct production among children at age seven; other sounds are produced without errors before that age). Bosch (2004: 54) also showed, this time in relation to Spanish, that the trill is the second-most late-acquired sound (behind only the palatal lateral); fewer than 50% of three-year olds produce the sound correctly; by five years old, 70% can pronounce it; at age six, 80% produce it correctly; and, by age seven, 90% are able to produce the sound correctly. Gómez-Fernández (2004) conducted a study on children from Seville between the ages of 12 and 36 months and also found that rhotic consonants rarely appear before the age of two. In a comparative study on children with Spanish or German as L1, Kehoe (2018) found that the alveolar rhotic is acquired later in Spanish than in German (a language in which the rhotic is an uvular approximant). Other studies relating to Spanish have described this trend and have detailed the simplification processes that occur in rhotic sounds (Melgar de González, 1976; Miras Martínez, 1992; León and Prinz, 1996; Goldstein and Cintrón, 2001; Vivar, 2009). With respect to English, a study by Shriberg (1993) found that children present articulatory problems with the approximant rhotic sound up to the age of eight, difficulties that may still be present as residual errors in adolescent speakers (Preston and Edwards, 2007). Another series of studies (Stemberger and Bernhardt, 2018; Másdóttir, 2018; Lundeborg Hammarström, 2018; Ramalho and Freitas, 2018; Pérez et al., 2018; Ignatova et al., 2018; Ozbič et al., 2018; Tar, 2018; Bernhardt and

Stemberger, 2018) have described the characteristics of word-initial rhotic clusters in typical and protracted phonological development in seven languages with taps or trills: Icelandic, Swedish, Portuguese, Spanish, Bulgarian, Slovenian and Hungarian. These studies revealed that, during the acquisition of these languages, rhotics usually appear later than /l/ (except in Portuguese) and that the most common process in the early years is elision of the rhotic sound. In later development, the most common process is replacement with other, non-nasal sonorants that share the place and manner of articulation with the substituted segment. Another phenomenon observed in all seven languages is vowel epenthesis.

The delayed acquisition of rhotic sounds is related to articulatory difficulties. An ultrasound study carried out by Boyce, Hamilton and Rivera-Campos (2016) with speakers of English, Malayalam, French, Persian and Spanish provided data to suggest that an articulatory characteristic shared by all rhotic sounds is a secondary articulation involving a constrictive tongue root gesture towards the back pharyngeal wall, and that maintenance of this double articulation (i.e. the blade or dorsum on the one hand and the root on the other) constitutes an articulatory difficulty that is observed in the acquisition stages of several languages. This phonetic phenomenon unites these sounds, which have traditionally been regarded as a natural class based more on phonological evidence than on phonetic evidence (Ladefoged, 2005). According to this study, therefore, this phonetic quirk, which had already been described by Lindau (1985) and Catford (1986), unifies the variability of rhotics as a natural class of sounds, and also helps explain the articulatory difficulties observed in such sounds during typical development in many languages.

1.2. The rhotics of Catalan

Catalan has two rhotic phonemes: /ɾ/ and /r/. The former, called a tap, involves brief contact of the tip of the tongue against the alveolar ridge; the latter, a trill, is the result of a series of rapid contacts. The two rhotics are apicoalveolar, although they present differences with respect to the point of articulation:

> While the dorsum is not directly involved in production of taps, production of trills involves some lowering of the predorsum and retraction of the postdorsum and, therefore, a constrained dorsal configuration that is resistant to coarticulatory effects. In syllable-final position, these tongue configuration characteristics may help distinguish between the two types of rhotic consonant to a greater extent than the number of apicoalveolar contacts. (Recasens, 2014: 216, our translation).

From an acoustic point of view, accurate trill pronunciations present F1 values of between 350 and 600 Hz, F2 values of between 1000 Hz and 1600 Hz and F3 values of between 2000 and 2600 Hz; and in the case of taps, F1 values of between 200 and 600 Hz, F2 values of between 1000 and 1600 Hz and F3 values of between 2100 and 2600 Hz (Recasens, 1986). In terms of sonority, they can be both voiced and voiceless, with complete or partial sonority, depending on the context (Recasens, 1996: 325).

This work followed the classification proposed by Blecua (2001) and Blecua and Cicres (2019) used to analyse trills in Spanish. It consisted of determining the number of "components" of the rhotic (in other words, the number of closing and opening phases), as well as the main acoustic traits. Based on the number of components, therefore, four different rhotic types can be described: trills, which consist of three or more components, with alternating closing and opening phases (Figure 1); those formed by two components, i.e. a closing phase followed by an opening phase (Figure 2); those formed by a single component, which corresponds to a closing phase (Figure 3); and elisions. Blecua and Cicres (2019) analysed data on adult spontaneous speech in relation to the Spanish spoken in Madrid and Salamanca in implosive position and found that more than half the realizations were formed by two components (53.3%), while the second-largest group (30.6%) corresponded to single-component rhotics; cases of elisions (9.4%) and trills (6.7%) were clearly in the minority and, in the case of the latter, were found mainly in prepausal position. The literature also shows that the rate of speech affects rhotic realizations (Recasens, 2014) and that, more specifically, the more spontaneous the speech, the more likely it is to contain simpler rhotics (i.e. those with fewer components); for example, a study by Blecua (2001), which was based on a corpus of paragraph readings, found fewer elisions (0.3%) and more trills (13.5%) than a study by Blecua and Cicres (2019), which was based on a corpus of spontaneous speech.

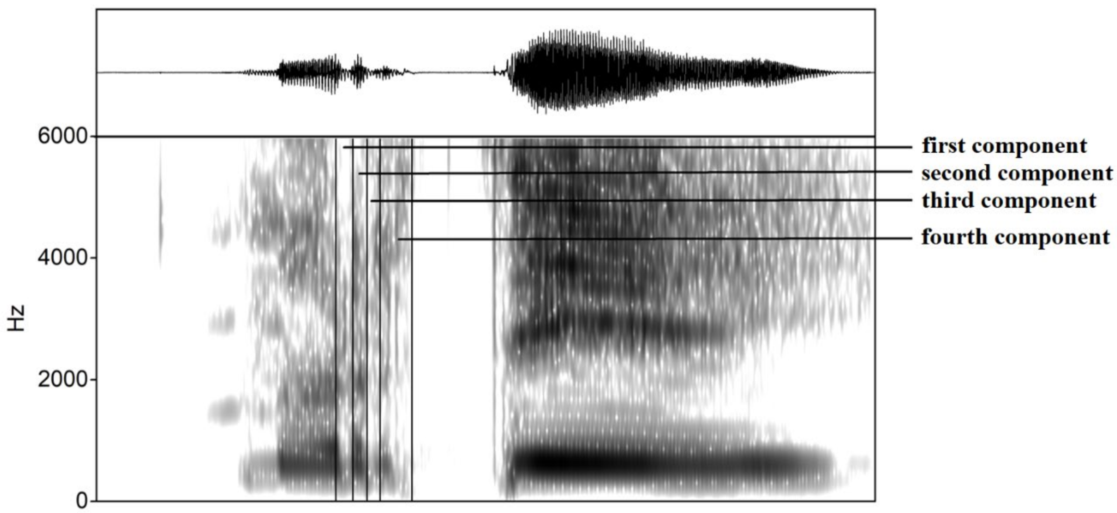

Figure 1. Oscillogram and spectrogram corresponding to the word *martell* ("hammer") produced by a six-year-old boy. The rhotic sound, a trill, is made up of four components.

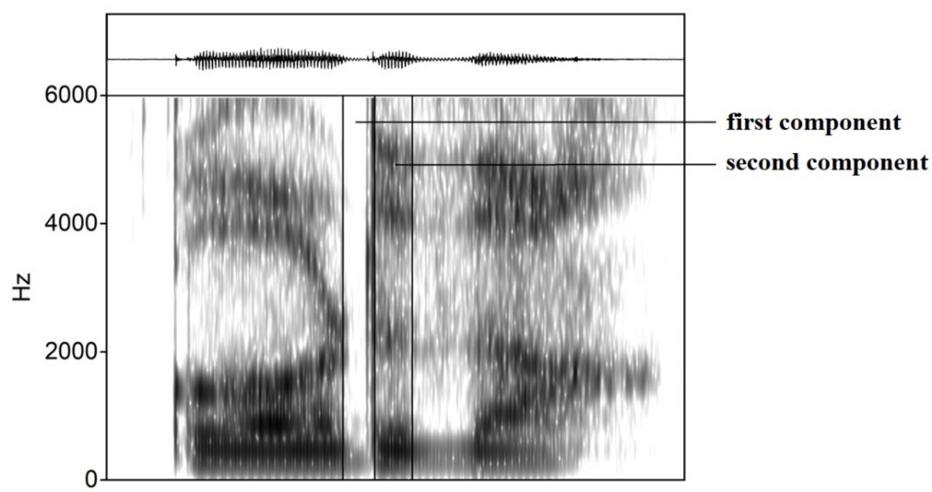

Figure 2. Oscillogram and spectrogram corresponding to the word *corda* ("string") produced by a five-year-old boy. The rhotic sound is made up of two components.

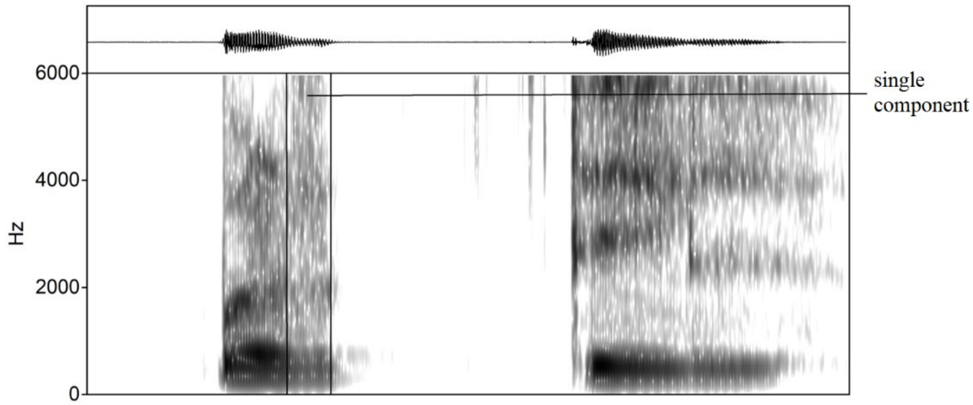

Figure 3. Oscillogram and spectrogram corresponding to the word *martell* ("hammer") produced by a six-year-old boy. The rhotic sound consists of a single component.

With respect to the acoustic properties of components, Blecua and Cicres (2019) described four types: occlusive, fricative, approximant and vocalic (see examples in figures 4-11, section 3). In this regard, the authors considered the vocalic element (also known as a "svarabhakti" vowel) to form part of the rhotic sound as an opening phase. The combination of different numbers of closing and opening phases with the different acoustic traits presented by the components results in a high number of possible realizations. With respect to Spanish rhotics in implosive position, with adult spontaneous speech, Blecua and Cicres (2019: 32) identified around 20 possibilities. The most common were realizations consisting of two components (an initial occlusive or approximant phase, followed by a vocalic element; or an initial occlusive component followed by a fricative phase) and those consisting of one component (approximant, occlusive or fricative).[1]

As we have seen, there is wide variation in the production of rhotic sounds, especially in preconsonantal and prepausal position. This variation is attributed to stylistic, dialectal and individual (or idiolectal) factors, in addition to the phonetic context in which the rhotic segment occurs.

First, with respect to stylistic factors, trills have been observed as having more components in controlled speech (which is usually slower and more careful) than in spontaneous speech (see, for example, Recasens, 2014, in relation to Catalan, and Martínez Celdrán, 1984, and Blecua and Cicres, 2019, in relation to Spanish).

Second, with regard to dialectal variation, the preconsonantal rhotic generally has a single brief contact in Mallorcan and Valencian (shorter in Valencian than in Mallorcan). In the case of central Catalan, it has one contact, and less frequently, two contacts that last longer than in other dialects. The type of realization, however, also depends on the characteristics of the following consonant. Other articulatory and acoustic differences occur in dialects, which generally point to more open articulation in Valencian and more retraction and constriction in central Catalan. Therefore, the differences lie not only in the number of contacts, but also in the overall articulatory gesture. With respect to the prepausal rhotic, this has a single

[1] Regarding rhotics with a single fricative component, Cicres and Blecua (2015) carried out an acoustic comparison of the sound (centre of gravity, standard deviation, skewness and kurtosis) with the sound of fricatives /s/ in prepausal position. They observed significant differences among all variables: the fricative consonants /s/ had higher mean values for centre of gravity, standard deviation and duration than those of the fricative components of rhotics; by contrast, skewness and kurtosis presented lower values.

contact and a long vocalic explosion before a pause in Valencian. In central Catalan, realization of the prepausal rhotic can be considered longer and stronger than that of Valencian, because it can be articulated with one or more contacts, it can be voiced or voiceless, it is primarily posterior (in terms of both alveolar constriction and dorsopalatal contact), has a low F2, and presents no vocalic explosion. In the Mallorcan dialect, there is phonological opposition between the two rhotics through differences in duration. There are also differences in the manner of articulation and formant frequency; on the other hand, both rhotics tend to be voiceless.

Another feature of the central dialect is the frequent insertion of epenthetic vowels, or postconsonantal epitheses that reinforce the prepausal rhotic (Bernat Baltrons, 1991; Fernández Planas, 1993). Roussillonese and the Catalan spoken in Sóller are also worth mentioning. With respect to northern Catalonia, Recasens (1996: 325) observed that interference from the French language caused taps and trills to be replaced by postdorsal-uvular trills, fricatives or approximants ([ʀ] o [ʁ]); in relation to the Catalan spoken in Sóller, Llompart (2013) described the "French r" of these Mallorcan people, similar to the uvular fricative found in French, and attributed it to the return of numerous migrants who had settled in the south of France in the second half of the 19th century and early 20th century.

Finally, few studies have examined speaker-related variation; we know only of a study by Blecua, Cicres and Gil (2014), who detected significant differences in prepausal rhotics, in terms of both the number of components and the type of first and second components. However, this was based on adult speakers of Spanish.

The main objective of this study is therefore to carry out an acoustic analysis of rhotic sounds in the Catalan spoken in central Catalonia in coda position, with respect to children's speech, in accordance with the methodology applied by Blecua (2001) to Spanish, i.e. by analysing the number and type of components. Three independent variables were considered: position in the word (word-final or word-medial coda position), sex and age.

2. Methodology

2.1. Corpus

This study was based on a sample from a corpus of words used for a naming test (used in Llach, 2007). The participants were 90 children aged between five and seven, divided into three groups of 30 (15 boys and 15 girls) according to their chronological age (from five years to five years and 11 months; from six years to six years and 11 months; and from seven years to seven years and 11 months). All the children lived in Hostalric or Arbúcies (in the *comarca* of La Selva) and fulfilled two requirements: first, that Catalan was their home language (that is, they spoke it with their parents); and, second, that they presented typical overall development. Therefore, children who presented disabilities related to the organs involved in speech production or reception (orofacial structures and hearing) or language difficulties (whether or not they were receiving speech therapy) were excluded from the study.

The evaluator then read part of a sentence, which the participant had to complete by saying the target word only, with the aid of an image. For example, to encourage the participant to say the word *martell* ("hammer"), the evaluator held up an image of this tool and said "To knock a nail into a wall, we use a...". The participant had to complete the sentence by saying the target word only. This ensured that the words were pronounced in isolation and that no coarticulatory effects occurred. A single sample of each word was obtained.

Recording was done with a unidirectional Shure 515SD microphone connected to an external sound card, with a sampling frequency of 22 kHz and 16-

bit resolution. The test was administered in a quiet but non-soundproofed room.

The rhotic phonemes under study were those that appear in the words *mar* (“sea”), *cor* (“heart”), *corda* (“string”) and *martell* (“hammer”). In all four words, the rhotic phoneme is in implosive position; in two cases, it is in prepausal position, and in the other two, it is followed by a coronal consonant (such that the point of articulation is similar in both cases). It has been shown that the difference in the approximant and occlusive manners of articulation of posterior consonants does not cause significant changes to the properties of the anterior rhotic sound, at least with respect to central peninsular Spanish, since both manners favour two-component rhotics (Blecua and Cicres, 2019: 35–36). This same study also found that the sonority of the posterior consonant does not have any significant effect on the rhotic. The timbre of the anterior vowel was also controlled; as per Gibson (2015), there are similarities in the articulatory configuration of trills and low and mid-low vowels. Thus, the three vowels preceding the rhotic ([a], [ə] and [ɔ]) should marginally favour trills (i.e. rhotics with three or more components).

The corpus under analysis consisted of 360 words, distributed equally according to the age and sex of the participants (Table 1). The original corpus contained more recordings per age group, but only the samples in which the participants actually said the word were chosen (until all 360 words that made up the final corpus were included); therefore, samples were discarded whenever the children did not say the word (because they did not know it or said a synonym instead) or introduced variations.

	5	**6**	**7**
/ˈkɔr/	30	30	30
/ˈkɔr.ðə/	30	30	30
/ˈmar/	30	30	30
/mər.ˈteʎ/	30	30	30
Total	120	120	120

Table 1. Number of words in the corpus, according to age.

2.2. Method

To carry out an acoustic description of the various components, the methodology proposed by Blecua and Cicres was followed (2019: 24–25), based on analysis of the oscillogram and spectrogram. Version 6.1.07 of the Praat program (Boersma and Weenink, 2019) was used for the acoustic analysis.

Table 2 shows the criteria used to determine the characteristics of the various components. It should be noted that both vocalic and approximant components have the same combination of acoustic characteristics, so they were labelled depending on their position in a closing phase (approximant) or opening phase (vocalic), in accordance with the criteria used by Blecua and Cicres (2019: 25).

		Defined formants	Noise at high frequencies	Occlusion (with or without a release burst)	High number of zero-crossings[2]
Component type	Vocalic	Yes	No	No	No
	Approximant	Yes	No	No	No
	Fricative	No	Yes	No	Yes
	Occlusive	No	No	Yes	No

Table 2. Acoustic characteristics used to determine component types.

Once the number and characteristics of each component had been identified, the data were analysed by means of contingency tables. The data in these tables were arranged in rows and columns (to represent the different study variables: number of components, component types, age and sex of the participants and position of the rhotic sound in the word), such that each field contained the number of cases of each combination of variables (for example, component types according to the sex of the participants, or the number of components according to the position of the rhotic in the word). In addition, these tables also presented the number of cases that would be expected if the crossed variables were independent. Pearson's chi-squared test was used to analyse the discrepancy between the observed and expected frequencies. Finally, the contingency tables reflected the corrected typified residuals; in other words, the difference between the observed and expected frequency in each field of the contingency table, but corrected so that they had zero mean and standard deviation 1. In line with Pardo and Ruiz (2002: 250), a 95% confidence level was used, so that corrected typified residual values greater than 1.96 or lower than -1.96 indicated significant differences.

3. Results and discussion

3.1. Number of components

Most rhotic realizations in syllable coda position were made up of one (19.2%) or two components (45.8%). Trill realizations (consisting of three or more components, i.e. with at least two closing phases separated by an opening phase) made up the remaining 35%. More than half this group corresponded to realizations with three or four components (11.4% and 8.6%, respectively). The remaining realizations (15%) had five or more components. Examples of realizations with as many as 14 components were observed, although the number was negligible (Table 3). Finally, no cases of elision were detected.

The position of the rhotic in the word (syllable coda in word-initial or word-final position) had a significant effect on the number of components (Pearson's chi-square = 60.625, significance level <0.001). Thus, there were almost twice as many cases of rhotics with two components in word-medial position than in word-final position (60.6% and 31.1%, respectively). Similarly, rhotics with five or more components were almost always found in word-final position (Table 3).

[2] A zero-crossing is the point at which a continuous signal (which can be observed in an oscillogram) has a value of zero; in other words, when the signal changes value from positive to negative, or vice versa. The number of zero-crossings is associated with the noise level in the acoustic signal (thus, it is a variable used to describe fricatives and other sounds).

	Number of components											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**14**
/r#/ (N)	42	56	17	16	16	10	9	9	1	2	1	1
/r#/ (%)	23.3	31.1	9.4	8.9	8.9	5.6	5	5	0.6	1.1	0.6	0.6
/r.C/ (N)	26	109	25	15	1	3	1	0	0	0	0	0
/r.C/ (%)	14.4	60.6	13.9	8.3	0.6	1.7	0.6	0	0	0	0	0

Table 3. Number and percentage of realizations according to position in the word and number of components.

	Number of components											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**14**
/r#/	**2.3**	**-5.6**	-1.5	0.2	**3.7**	**2.0**	**2.6**	**3.0**	1.0	1.4	1.0	1.0
/r.C/	**-2.3**	**5.6**	1.5	-0.2	**-3.7**	**-2.0**	**-2.6**	**-3.0**	-1.0	-1.4	-1.0	-1.0

Table 4. Contingency table on the number of components depending on the position of the rhotic in the word. The fields reflect the corrected typified residuals; statistically significant values are highlighted in bold.

The contingency table that analyses the relationship between the number of components and the position of the rhotic sound in the word (Table 4) shows, first, that there were significantly fewer cases of two-component realizations in word-final position than expected (with respect to those in word-medial position) and, second, that there was a higher rate of both single-component rhotics and trill rhotics (with between five and eight components); due to the low number of total realizations, the differences in rhotics with more than eight components were non-significant, although they appeared only in word-final position.

This apparent contradiction observed in prepausal position (i.e. that there were significantly more single-component rhotics, but also more trills with more than five components) can be explained by the fact that, although the intensity and general articulatory tension in this position generally decrease with respect to other positions, the type of test administered caused some participants to pronounce words in an emphatic manner. In addition, it is important to consider speaker variation (see, for example, Blecua, Cicres and Gil, 2014) caused by the individual articulatory preferences of speakers who favour different types of realizations.

Finally, in terms of the number of components, the statistical tests did not show significant differences, even with respect to the participants' sex (Pearson's chi-square = 11.512, significance level = 0.401) and age (Pearson's chi-square = 31.025, significance level = 0.096).

The results obtained in this study on children's speech were broadly in line with the findings of other studies based on corpora of non-spontaneous adult speech, with respect to the number of components. For example, with regard to central Catalan, Recasens (2014) found that the most common realizations in preconsonantal position had one contact (and, therefore, one or two components, depending on whether an opening phase was observed) and, to a lesser extent, two contacts; by contrast, trills with two or more contacts increased in word-final position. Similarly, a study by Blecua (2001), based on a corpus of paragraph readings, also found that the most common realizations had one or two components (one contact) and also failed to detect elisions. However, the results were slightly different when compared to studies on spontaneous speech. For example, a study by Blecua and Cicres (2019), which focused on Spanish, observed elisions (9.4%) and fewer trills than this study on children's speech. It is clear, therefore, that the degree of spontaneity has an effect on the number of components; the more spontaneous the speech, the more likely it is to present simpler realizations (with fewer components) and even elisions. However, with corpora of readings or word naming, elisions are not observed and more cases of trills (with three or more components) are detected.

3.2. Component types

The results broken down by component type are presented below. To simplify the presentation of the results, we chose to limit the analysis to configurations that represented more than 1% of cases (Table 5). Although 78 different rhotic types were identified in the entire corpus (in view of the various combinations of occlusive, fricative, approximant and vocalic components), only 14 combinations exceeded the 1% threshold. Together, these accounted for 75% of the rhotic realizations. Figures 4 to 11 are examples of the rhotics that appeared most frequently in the corpus (more than 10 realizations).

Configurations (components)	Frequency	Percentage
Occlusive–vocalic element (OV)	56	15.6
Approximant–vocalic element (AV)	52	14.4
Fricative (F)	37	10.3
Approximant–fricative (AF)	28	7.8
Approximant (A)	27	7.5
Occlusive–fricative (OF)	21	5.8
Approximant–vocalic element–approximant (AVA)	12	3.3
Approximant–vocalic element–fricative (AVF)	7	1.9
Occlusive–vocalic element–approximant (OVA)	6	1.7
Approximant–fricative–approximant–fricative (AFAF)	5	1.4
Approximant–vocalic element–approximant–vocalic element (AVAV)	5	1.4
Fricative–vocalic element (FV)	5	1.4
Approximant–occlusive–fricative (AOF)	4	1.1
Occlusive–fricative–approximant (OFA)	4	1.1
Total	**269**	**75**

Table 5. Rhotic configuration types according to the number and type of components with a percentage of occurrence in the corpus greater than 1%.

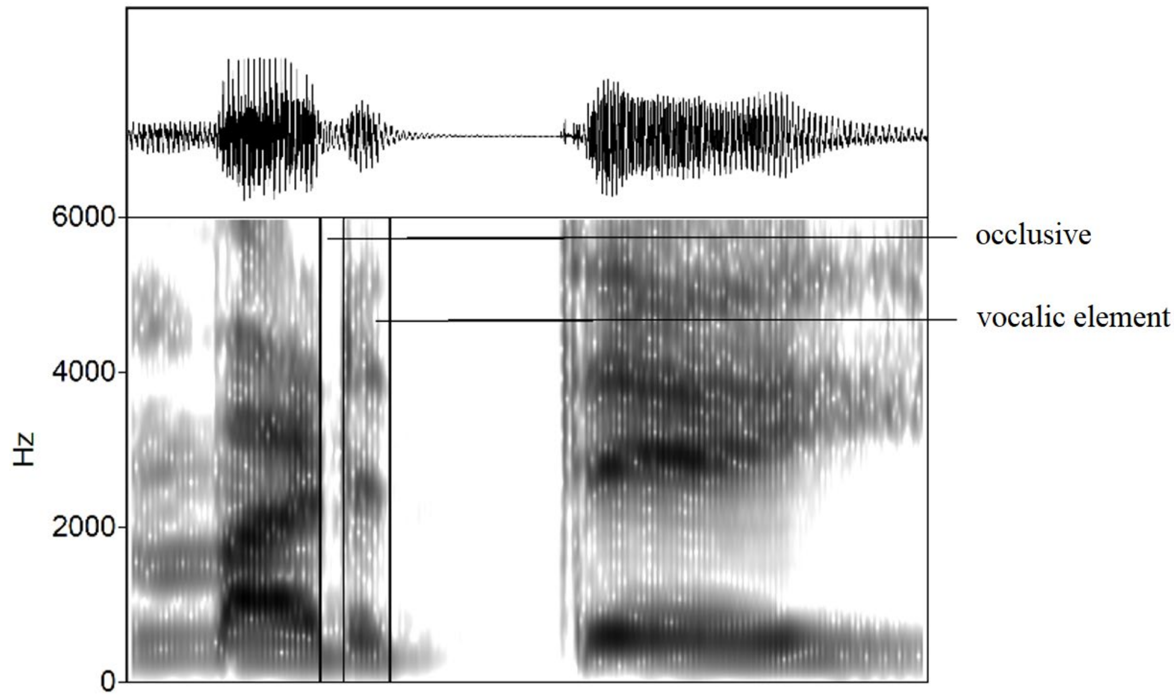

Figure 4. Oscillogram and spectrogram corresponding to the word *martell* ("hammer") produced by a six-year-old boy. The rhotic sound is made up of two components (occlusive and vocalic).

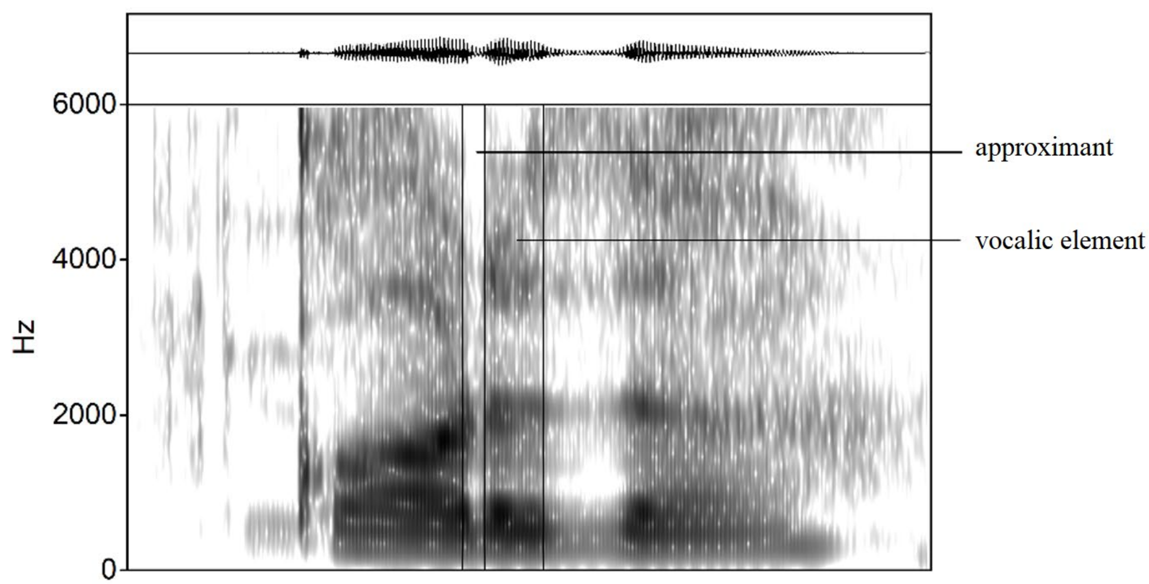

Figure 5. Oscillogram and spectrogram corresponding to the word *corda* ("string") produced by a five-year-old boy. The rhotic sound is made up of two components (approximant and vocalic).

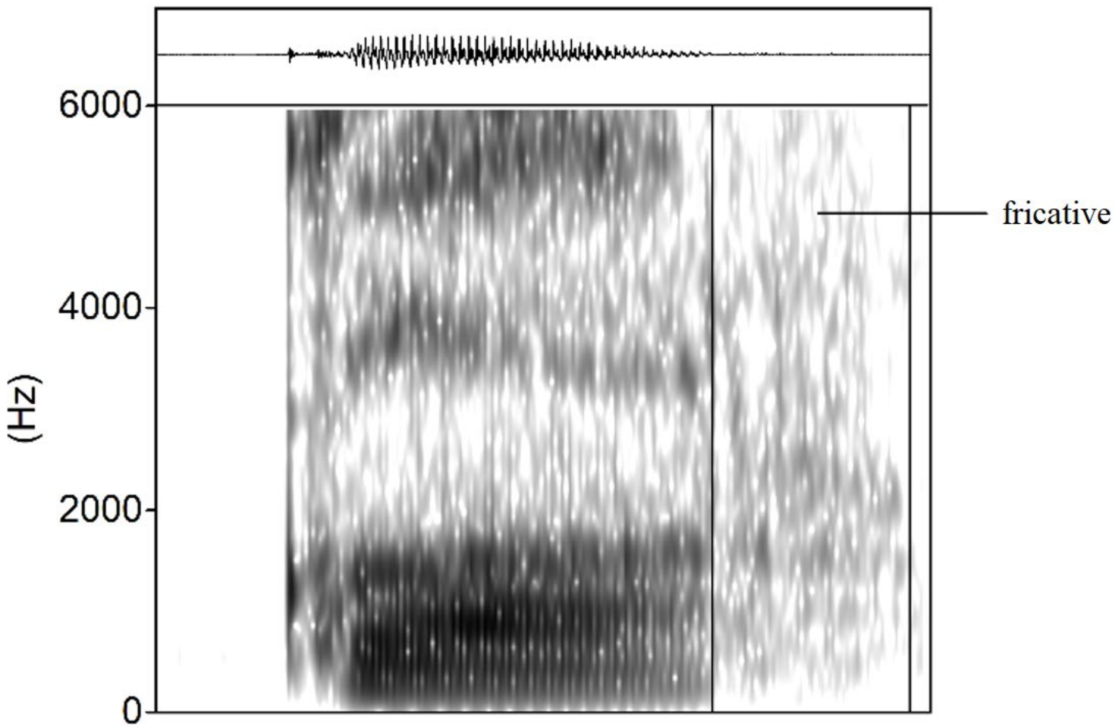

Figure 6. Oscillogram and spectrogram corresponding to the word *cor* (“heart”) produced by a five-year-old boy. The rhotic sound consists of a single fricative component.

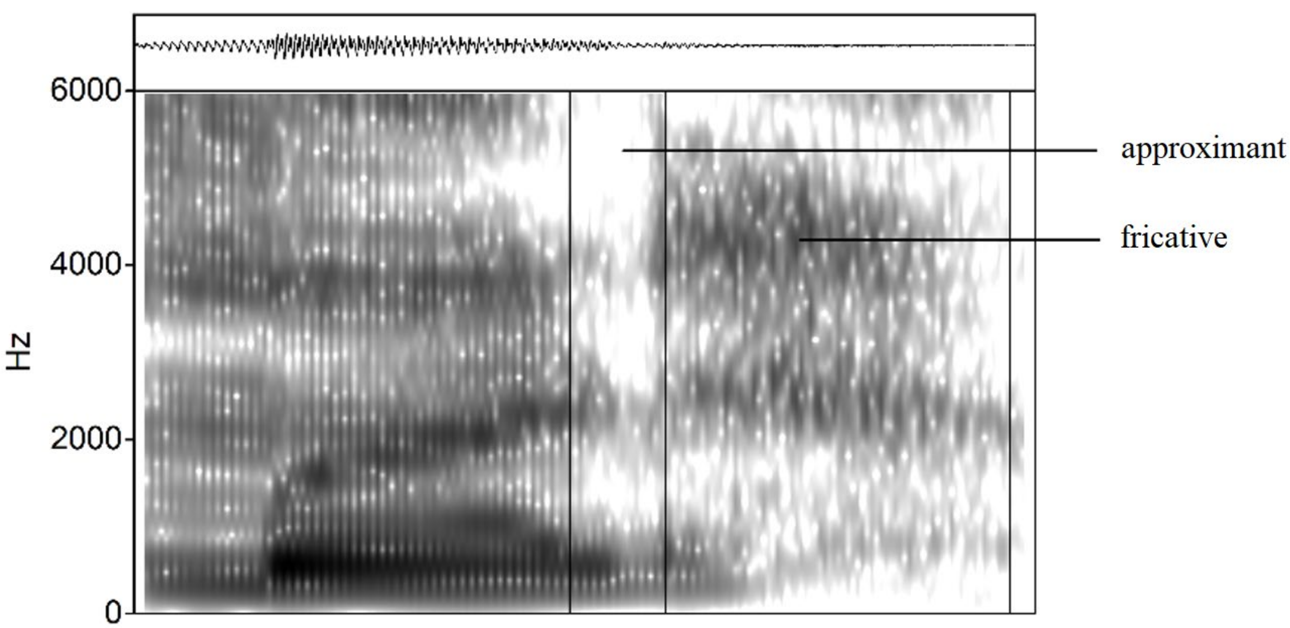

Figure 7. Oscillogram and spectrogram corresponding to the word *mar* (“sea”) produced by a five-year-old boy. The rhotic sound is made up of two components (approximant and fricative).

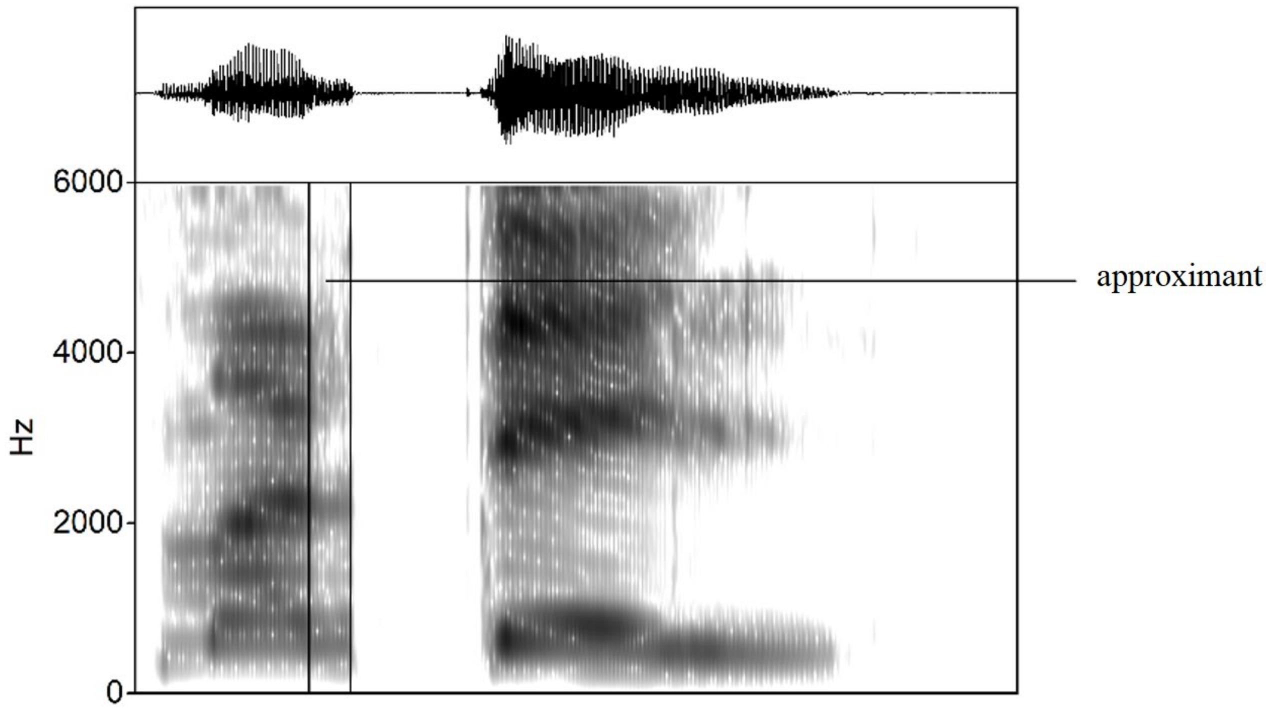

Figure 8. Oscillogram and spectrogram corresponding to the word *martell* (“hammer”) produced by a five-year-old boy. The rhotic sound consists of a single approximant component.

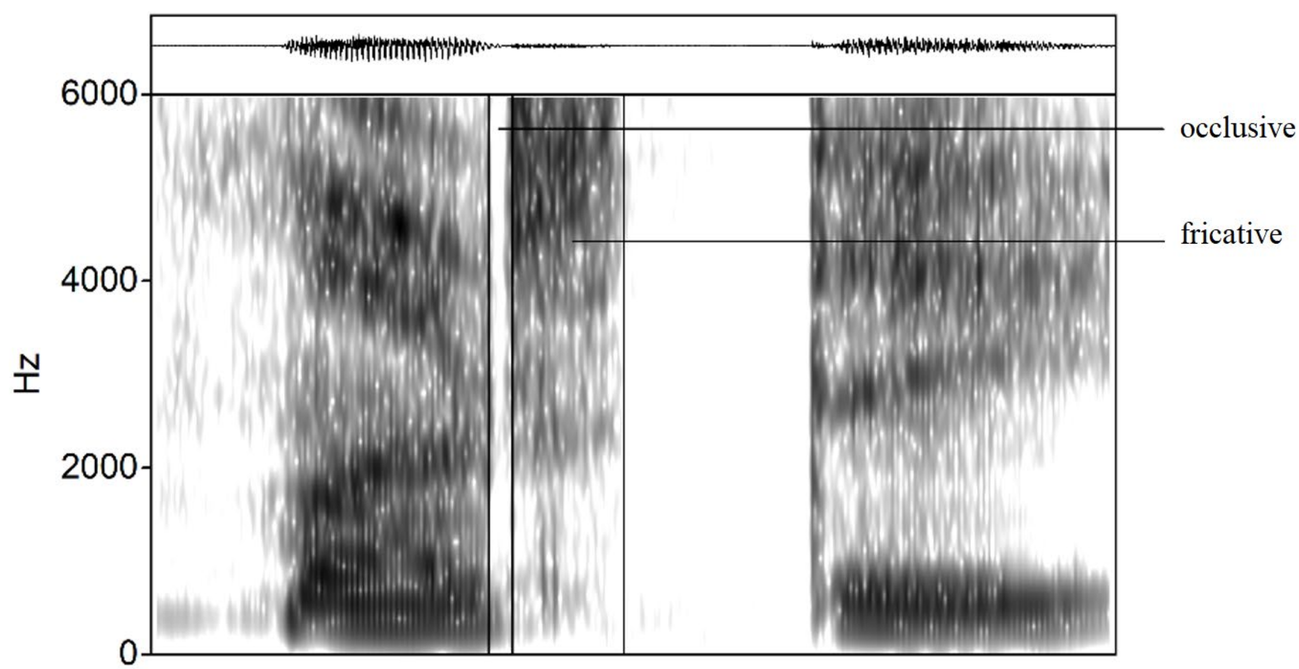

Figure 9. Oscillogram and spectrogram corresponding to the word *martell* (“hammer”) produced by a five-year-old boy. The rhotic sound is made up of two components (occlusive and fricative).

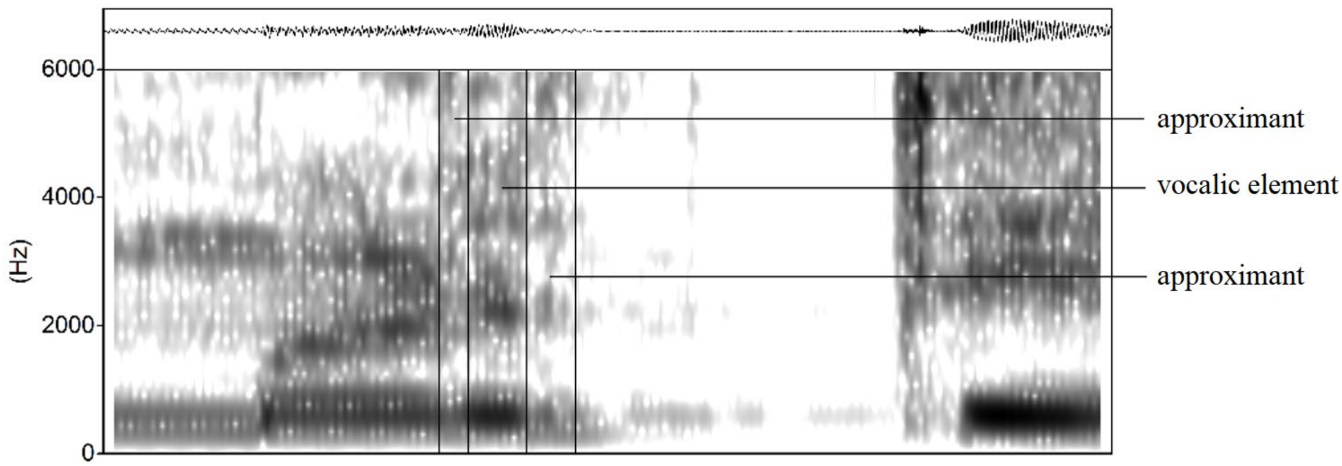

Figure 10. Oscillogram and spectrogram corresponding to the word *martell* ("hammer") produced by a five-year-old boy. The rhotic sound is made up of three components (approximant, vocalic and approximant).

Nonetheless, it is important to note that other, more complex combinations of components were observed, albeit to a negligible degree. As an example, Figure 11 illustrates an eight-component rhotic. If we focus only on the most common rhotic types (as shown in Table 5), it is clear that the position of the rhotic in the word (prepausal or implosive before a heterosyllabic consonant) influences the combination of components (Pearson's chi-square = 97.933; significance level <0.001), as illustrated in the graph (Figure 12).

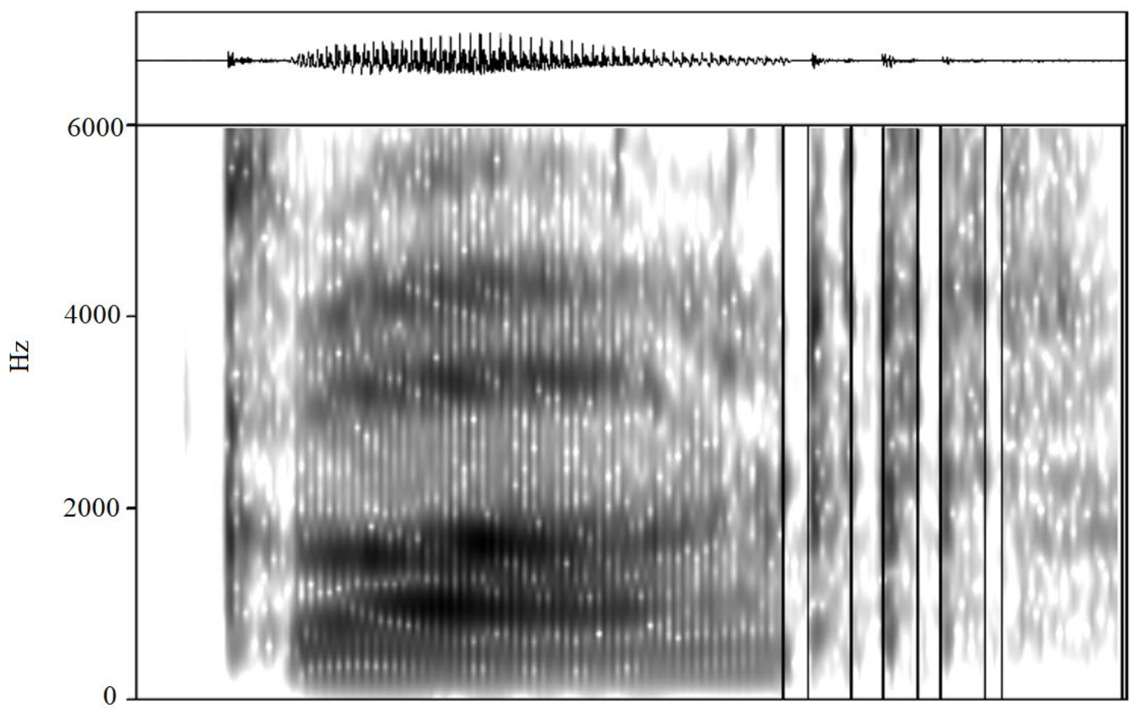

Figure 11. Oscillogram and spectrogram corresponding to the word *cor* ("heart") produced by a six-year-old boy. The rhotic sound is made up of eight components that alternate between occlusive and fricative segments.

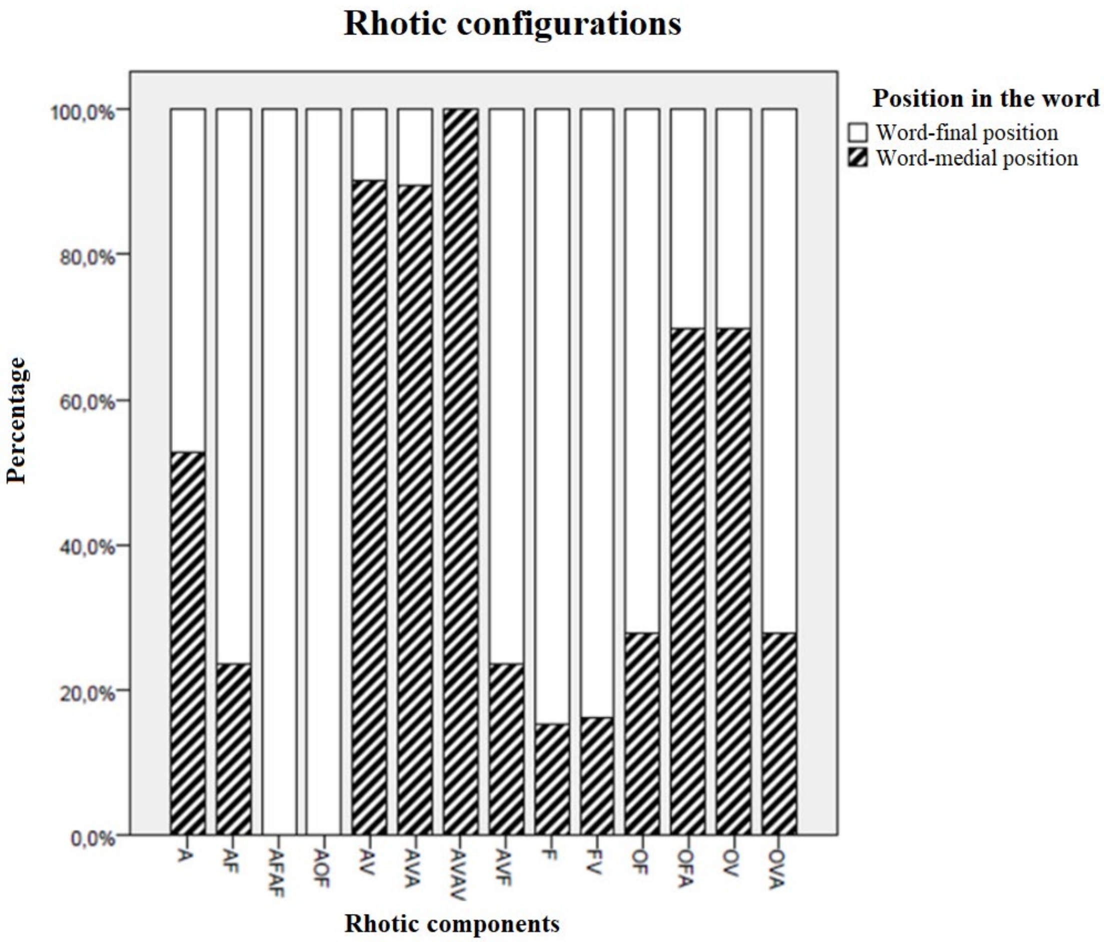

Figure 12. Distribution of rhotic configurations according to word-medial or word-final position.

Six rhotic types were found in word-medial position in more than half of cases. These were the following combinations: occlusive–vocalic element; approximant–vocalic element; approximant; approximant–vocalic element–approximant; approximant–vocalic element– approximant–vocalic element; and occlusive–fricative–approximant. The implosive position before a heterosyllabic consonant, therefore, usually (though not categorically) favours components with a higher degree of sonority (vocalic elements and approximants); by contrast, rhotic combinations with fricative or occlusive components (with a lower degree of sonority) are more common in prepausal position.

It should be noted, however, that differences in the distribution of each rhotic type were not always significant. Table 6 shows the contingency table with the corrected typified residuals of the various rhotic types in relation to prepausal or word-medial position. It shows that significant differences occurred in just nine cases. A positive or negative sign indicates a significantly higher or lower presence than would be expected if the contrasted variables were independent.

	A	AF	AFAF	AOF	AV	AVA	AVAV
Syllable coda in word-final position	-0.1	**3.1**	**2.6**	**2.3**	**-5.8**	**-2.5**	**-2**
Syllable coda in word-medial position	0.1	**-3.1**	**-2.6**	**-2.3**	**5.8**	**2.5**	**2**

	AVF	F	FV	OF	OFA	OV	OVA
Syllable coda in word-final position	1.5	**4.9**	1.7	**2.2**	-0.8	**-3.2**	1.1
Syllable coda in word-medial position	-1.5	**-4.9**	-1.7	**-2.2**	0.8	**3.2**	-1.1

Table 6. Contingency table on the structure of rhotic components according to the position of the rhotic in the word. The fields reflect the corrected typified residuals; statistically significant values are highlighted in bold.

This table reveals that only the component combinations approximant–vocalic; approximant–vocalic–approximant; approximant–vocalic–approximant–vocalic; and occlusive–vocalic presented more cases than expected when the syllable coda was in word-medial position. A feature shared by all these combinations was a vocalic component in second position (typically an opening phase). This vocalic component, which has a formant structure but is shorter and less intense than a vowel, is known as a "svarabhakti" vowel (it is also referred to an "intrusive" or "epenthetic" vowel in the literature). The frequency of formants is similar to that of adjacent vowels, but because it does not correspond to a phoneme, it is not perceived by speakers. In fact, the combination of an obstructive phase (either a complete occlusion or a fricative or approximant element) and an opening phase (this svarabhakti element) is considered to be the most common canonical realization (Blecua, 2001; Hualde, 2005; Martínez Celdrán and Fernández Planas, 2007). As shown in Table 5, these canonical realizations were also the most common in a controlled corpus of children's speech, although many other component combinations appeared.

On the other hand, rhotics with a single fricative component or a combination of approximant–fricative, occlusive–fricative, approximant–fricative–approximant–fricative or approximant–occlusive–fricative components presented significantly more realizations in prepausal position, at the end of a word. None of these cases presented a vocalic element (or svarabhakti vowel), but lower-sonority segments (mainly fricatives and occlusives) occurred more frequently. In fact, if the rhotic consists of only one component in word-final position, it is more likely to be fricative. In this regard, Recasens (2014) observed fricative rhotics in word-final position in central Catalan. This result is also similar to the findings of Blecua and Cicres (2019: 38) in relation to Spanish and adult spontaneous speech.

Thus, with respect to opening phases (that is, phases corresponding to the components that appear in positions two, four, six, etc.), fricative realizations are most common before a pause, while vocalic realizations predominate in preconsonantal position (again, the data are consistent with the findings of Blecua and Cicres, 2019: 42–43). However, these fricative rhotics present significant acoustic differences with respect to fricative phoneme realizations (see note 1).

By contrast, the statistical tests showed that neither the sex variable (Pearson's chi-square = 20.355; significance level = 0.087) nor the age variable (Pearson's chi-square = 17.911; significance level = 0.879) significantly influenced the occurrence of certain configurations over others.

4. Conclusions

This study, which focused on children's speech after initial acquisition of the phonological system (from age five), demonstrated considerable diversity in the production of rhotic sounds, in both preconsonantal and prepausal positions. This diversity was evident in terms of both the number of components (in the case of slow, controlled speech, no cases of elision were observed, but rhotics consisting of between one and 14 components were detected) and combinations of components. However, the majority of the most common combinations (occlusive–vocalic element; approximant–vocalic element; fricative; approximant–fricative; approximant; occlusive–fricative; approximant–vocalic element–approximant; occlusive–vocalic element–approximant, etc.) have also been documented in other studies on adult speech. Nonetheless, some combinations (such as approximant–fricative–approximant–fricative and approximant–occlusive–fricative) have been described here for the first time. In addition, other combinations that were much more complex but occurred to a much lesser extent were observed.

In general terms, however, some recurring patterns were detected. First, we observed how preconsonantal position (in medial coda) favours the appearance of svarabhakti vowels, while rhotics with a vocalic element appear significantly less often in prepausal position. Second, the degree of sonority of components is also influenced by the position of the rhotic in the phonic sequence: occlusive and fricative components (with a low degree of sonority) are more common in word-final position, while vocalic elements and approximants (components with a higher degree of sonority) are more common in word-medial position.

Finally, despite the fact that the experiment involved a highly controlled speech model (since it consisted of a picture-naming test, in which participants said only the target words), the data indicate that rhotic productions in children between the ages of five and seven are already comparable to those of adult speakers, although the children presented much greater variability, probably as a result of the type of test administered and their immature motor control. Therefore, more studies are needed to carry out a more direct comparison, either with corpora of spontaneous children's speech or with corpora of controlled adult speech, under similar conditions and with a similar number of participants.

References

Aoyama, K., Akbari, C. and Flege, J. E. (2016). Prosodic Characteristics of American English in School-Age Children. Dins *Speech Prosody 2016*, Boston, 572–576.

Bernat Baltrons, F. (1991). Estudi sobre la pronunciació d'erra final a Blanes, *Anuari de filologia XIV, secció C, Llengua i literatura catalanes*, *2*, 23–44.

Bernhardt, B. M. and Stemberger, J. P. (2018). Tap and trill clusters in typical and protracted phonological development: Conclusion, *Clinical Linguistics and Phonetics*, *32*(5–6), 563–575.

Blecua, B. (2001). *Las vibrantes del español: manifestaciones acústicas y procesos fonéticos* (PhD dissertation). Universitat Autònoma de Barcelona, Bellaterra.

Blecua, B. and Cicres, J. (2019). Rhotic variation in Spanish codas: Acoustic analysis and effects of context in spontaneous speech. In M. Gibson and J. Gil (Eds.). *Romance Phonetics and Phonology* (pp. 21–47). Oxford University Press.

Blecua, B., Cicres, J. and Gil, J. (2014). Variación en las róticas del español y su implicación en la identificación del locutor, *Revista de Filología Románica*, *31*, 13–35.

Boersma, P. and Weenink, D. (2019). *Praat v.6.1.07*, www.praat.org [2/12/2019].

Bosch, L. (1984). El desarrollo fonológico infantil: una prueba para su evaluación. In M. Siguán (Dir.). *Estudios sobre psicología del lenguaje infantil* (pp. 33–58). Ediciones Pirámide.

Bosch, L. (1987). *Avaluació del desenvolupament fonològic en nens catalanoparlants de 3 a 7 anys*. Publicacions de l'ICE.

Bosch, L. (2004). *Evaluación fonológica del habla infantil*. Masson.

Boyce, S. E., Hamilton, S. M. and Rivera-Campos, A. (2016). Acquiring rhoticity across languages: An ultrasound study of differentiating tongue movements, *Clinical Linguistics and Phonetics*, *30*(3–5), 174–201.

Caplan, D. (1993). *Language: Structure, processing, and disorders*. MIT Press.

Catford, J. C. (1986). Comment on "variability in feature specifications". In J. S. Perkell and D. H. Klatt (Eds.). *Invariance and variability in speech processing* (pp. 478–479). Lawrence Erlbaum Associates.

Cicres, J. and Blecua, B. (2015). Caracterización acústica de las róticas fricativas prepausales en español peninsular, *Loquens*, *2*(1), e019.

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*(2), 462–467.

De Ribot, M. D. (1992). *Problemàtica de l'adquisició del sistema fonemàtic de la llengua catalana a les comarques de Girona*. Publicacions de l'Abadia de Montserrat.

Eguchi, S. and Hirsh, I. J. (1969). Development of speech sounds in children, *Acta Oto-laryngologica. Supplementum*, *257*, 1–51.

Fernández Planas, A. M. (1993). L'epítesi postconsonàntica de mots amb vibrant final a Blanes, *Anuari de filologia XIV, secció C, Llengua i literatura catalanes*, *4*, 73–102.

Gibson, M. (2015). A stochastic approach to rhotic variation in Spanish codas, *Loquens*, *2*(1), e015.

Goldstein, B. A. and Cintrón, P. (2001). An investigation of phonological skills in Puerto-Rican Spanish-speaking 2-year-olds, *Clinical Linguistics and Phonetics*, *15*, 343–361.

Gómez-Fernández, D. (2004). La adquisición de las líquidas en los niños de Sevilla y su provincia: Desde los doce a los treinta y seis meses, *Revista de Filología y su Didáctica*, *27*, 125–185.

Grunwell, P. (1981). The development of phonology: A descriptive profile, *First Language*, *8*, 161–191.

Haelsig, P. C. and Madison, C. L. (1986). A Study of Phonological Processes Exhibited by 3-, 4-, and 5-Year-Old Children, *Language, Speech, and Hearing Services in Schools*, *17*, 107–114.

Hillenbrand, J., Getty, L. A., Clark, M. J. and Wheeler, K. (1995). Acoustic characteristics of American English vowels, *The Journal of the Acoustical Society of America*, *97*, 3099–3111.

Hualde, J. I. (2005). *The Sounds of Spanish.* Cambridge University Press.

Ignatova, D., Bernhardt, B. M., Marinova-Todd, S. and Stemberger, J. P. (2018). Acquisition of word-initial trill clusters in children with typical versus protracted phonological development: Bulgarian, *Clinical Linguistics and Phonetics*, *32*(5–6), 506–522.

Ingram, D. (1976). *Phonological Disability in Children*. Elsevier.

Ingram, D., Christesen, L., Veach, S. and Webster, B. (1980). The acquisition of word initial fricatives and affricates in English by children between 2 and 6 years. In G. Yeni-

Komshian, J. Kavanagh and C. Ferguson (Eds.). *Child Phonology: Production,* volum I (pp. 169–192). Academic Press.

Kehoe, M. (2018). The development of rhotics: A comparison of monolingual and bilingual children, *Bilingualism: Language and Cognition*, *21*(4), 710–731.

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic study, *Journal of Speech and Hearing Research*, *19*(3), 421–445.

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

Klein, H. B., McAllister Byun, T., Davidson, L. and Grigos, M. I. (2013). A Multidimensional Investigation of Children's /r/ Productions: Perceptual, Ultrasound, and Acoustic Measures, *American Journal of Speech-Language 540 Pathology*, *22*, 540–553.

Ladefoged, P. (2005). *A course in phonetics.* Harcourt (1st ed. 1975).

Lee, S., Potamianos, A. and Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters, *The Journal of the Acoustical Society of America*, *105*(3), 1455–1468.

Lindau, M. (1985). The story of /r/. In V. A. Fromkin (Ed.). *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 157–168). Academic Press.

Llach, S. (2007). *Fonaments fonètics de l'adquisició de la fonologia de les consonants del català* (PhD dissertation). Universitat Autònoma de Barcelona, Bellaterra.

Lleó, C. and Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment, *Journal of Child Language*, *23*, 31–56.

Llompart, M. (2013). La «r francesa» en Sóller (Mallorca) y su relación con la adquisición de L1 y L2 en hablantes bilingües catalanodominantes, *Estudios de Fonética Experimental*, *XXII*, 193–232.

Lundeborg Hammarström, I. (2018). Word-initial /r/-clusters in Swedish-speaking children with typical versus protracted phonological development, *Clinical Linguistics and Phonetics*, *32*(5–6), 446–458.

Martínez Celdrán, E. (1984). *Fonética*. Teide.

Martínez Celdrán, E. and Fernández Planas, A. M. (2007). *Manual de fonética española*. Ariel.

Másdóttir, T. (2018). Word-initial /r/-clusters in Icelandic-speaking children with protracted versus typical phonological development, *Clinical Linguistics and Phonetics*, *32*(5–6), 424–445.

McGowan, R. S., Nittrouer, S. and Manning, C. J. (2004). Development of [r] in young, midwestern, American children, *Journal of the Acoustical Society of America*, *115*(2), 871–884.

Melgar de González, M. (1976). *Cómo detectar al niño con problemas de habla*. Editorial Trillas.

Miras Martínez, F. (1992). *El desarrollo fonoarticulatorio del habla infantil*. Instituto de Estudios Almerienses.

Oller, D. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. Kavanagh and C. Ferguson (Eds.). *Child Phonology: Production,* volum I (pp. 93–110). Academic Press.

Ozbič, M., Kogošek, D., Bernhardt, B. M., Stemberger, J. P., Muznik, M. and Brce, J. N. (2018). Word- initial rhotics in slovenian 4-year-olds with typical versus protracted phonological development, *Clinical Linguistics and Phonetics*, *32*(5–6), 523–543.

Pardo, A. and Ruiz, M. Á. (2002). *SPSS 11. Guía para el análisis de datos*. McGraw Hill.

Perez, D., Vivar, P., Bernhardt, B. M., Mendoza, E., Ávila, C., Carballo, G. and Vergara, P. (2018). Word-initial rhotic clusters in

Spanish-speaking preschoolers in Chile and Granada, Spain, *Clinical Linguistics and Phonetics*, *32*(5–6), 481–505.

Preston, J. L. and Edwards, M. L. (2007). Phonological processing skills of adolescents with residual speech sound errors, *Language, Speech, and Hearing Services in Schools*, *38*(4), 297–308.

Ramalho, A. M. and Freitas, M. J. (2018). Word-initial rhotic clusters in typically developing children: European Portuguese, *Clinical Linguistics and Phonetics*, *32*(5–6), 459–480.

Recasens, D. (1986). *Estudis de fonètica experimental del català oriental central.* Publicacions de l'Abadia de Montserrat.

Recasens, D. (1996). *Fonètica descriptiva del català*. Institut d'Estudis Catalans (1st ed. 1991).

Recasens, D. (2014). *Fonètica i fonologia experimentals del català: vocals i consonants*. Institut d'Estudis Catalans.

Roberts, J., Burchinal, M. and Footo, M. (1990). Phonological process decline from 2½ to 8 years, *Communication Disorders*, *23*, 205–217.

Shriberg, L. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders, *Journal of Speech and Hearing Research*, *36*, 105–140.

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, *The Journal of the Acoustical Society of America*, *91*, 2165–2174.

Smith, B., Kenney, M. K. and Hussain, S. (1995). A longitudinal investigation of duration and temporal variability in children's speech production, *The Journal of the Acoustical Society of America*, *99*, 2344–2349.

Stark, R. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh and C. Ferguson (Eds.). *Child Phonology: Production,* Volume I (pp. 73–90), Academic Press.

Stemberger, J. P. and Bernhardt, B. (2018). Tap and trill clusters in typical and protracted phonological development: Challenging segments in complex phonological environments. Introduction to the special issue, *Clinical Linguistics and Phonetics*, *32*(5–6), 411–423.

Tar, É. (2018). Word-initial tap-trill clusters: Hungarian, *Clinical Linguistics and Phonetics*, *32*(5–6), 544–562.

Vihman, M. (1996). *Phonological development. The origins of language in the child.* Blackwell.

Vivar, P. (2009). Evaluación de grupos consonánticos de ataque complejo en un grupo de niños de la ciudad de Concepción con prueba articulatoria CEFI, *Onomázein*, *20*, 33–44.