Preliminary Evidence for Incomplete Neutralization of Coda Liquids in Puerto Rican Spanish

Miquel Simonet[1], Marcos Rohena-Madrazo[2] and Mercedes Paz[1]
University of Illinois at Urbana-Champaign[1], New York University[2]

1. Introduction

The term 'incomplete neutralization' refers to a phenomenon in which speakers produce small differences in the articulation of words which, based on phonological processes of positional neutralization, would be expected to be homophones (Warner, Jongman, Sereno & Kemps 2004). Interestingly, these phonemic differences are also perceptible; that is, listeners may make use of them in word-identification (e.g. Port & O'Dell 1985, Warner et al. 2004). The phonetic difference between sounds that are incompletely neutralized is much smaller than that between their corresponding phonemes in other positions, in which they are not supposed to neutralize, but is still significant. According to Warner et al. (2004:252) this "has implications for phonological theory because it runs contrary to the idea of categorical distinctions among segments." The most studied case of incomplete neutralization is final devoicing, which is a phenomenon in which /d/ and /t/ are (nearly) merged in word-final position. This occurs in languages such as German, Polish, Dutch, and Catalan (Charles-Luce 1985, 1993, Port & O'Dell 1985, Slowiaczek & Dinnsen 1985, Port & Crawford 1989, Warner et al. 2004, among many others). Incomplete neutralization effects may be attributable to differences in the words' underlying forms, their orthographic configuration, and/or the speech style in which they are produced (e.g. Port & Crawford 1989, Warner et al. 2004).

Puerto Rican Spanish is usually described in the literature as one of the Caribbean Spanish dialects that neutralizes the liquids /r/ and /l/ in post-nuclear position (e.g. /árma/ → [álma] 'weapon', /álma/ → [álma] 'soul'; Navarro Tomás 1948 [1974], López-Morales 1983, Lipski 1994, Ramos-Pellicia 2004). The two Spanish rhotic phonemes, the trill and the tap, are neutralized in syllable-coda position in all varieties. In most cases, they are neutralized as a tap, even though a trill is optional in emphatic speech for speakers of some dialects. Puerto Rican Spanish goes a step further in also neutralizing the contrast between the rhotic (a tap, in most dialects) and the lateral. Coda liquid neutralization, however, seems to affect /r/ more frequently than /l/, that is, 'lateralized' realizations of /r/ are more common than 'rhoticized' realizations of /l/ (Lipski 1994). Navarro Tomás (1974) observed considerable regional variation with regards to this phonological neutralization process. Lipski (1994), almost fifty years later, affirmed that the neutralization of liquids in coda position is now socially, rather than regionally, indexed. In a groundbreaking study, Navarro Tomás (1974) collected data from 43 speakers from different locations in Puerto Rico. He reported that neutralization of coda /r/ and /l/ to [r] was almost as common as neutralization to [l] (52.5% versus 41% respectively), so that both etymological phonemes could variably receive a rhotic or lateral pronunciation. More recently, López Morales (1983) examined the speech of a group of subjects from the San Juan metropolitan area. López Morales, based on his auditory-impressionistic analysis, described the use of four variants: (i) fricatives – including approximant realizations with no occlusion and different constriction degrees (45.6%); (ii) laterals – including a few 'mixed' sounds in which the lateral element is salient (34.6%); (iii) taps (14%); and (iv) elided or deleted liquids (5.6%). Regarding the lateralization of coda /r/, López Morales' results

* We wish to thank our informants/participants. We are indebted to Adam Baker, for writing the R procedure that made the Smoothing Spline ANOVA implementable, and to Eunice Díaz, for assistance at different stages during the preparation of the materials. For comments, we are grateful to José I. Hualde and Lisa Davidson. We appreciate the hard work of the editors and the reviewers, for their insights and corrections. All errors remain our own.

 Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology, ed. Laura Colantoni and Jeffrey Steele, 72-86. Somerville, MA: Cascadilla Proceedings Project.

showed a correlation with the socio-educational level of the speakers, since the percentage of lateralized rhotics found for the highest socio-economic strata was significantly lower than the percentages found for the lower levels. Other studies also reported a process of coda liquid neutralization involving the lateralization of /r/ (Vaquero 1996, Penny 2000). In a recent sociolinguistic investigation of Puerto Rican Spanish both in and outside Puerto Rico, Ramos-Pellicia (2004) described an almost categorical neutralization of coda /r/ to [l] in the speech of a number of rural speakers. In summary, the picture that arises from a literature review is unclear: some researchers describe multiple different allophones in similar proportions of use for both /r/ and /l/ in coda position, while others find lateralization of /r/ to be the most common process.

One reason for exercising caution in accepting categorical /r/ and /l/ neutralization in the Spanish of Puerto Rico comes from a recent finding by Paz (2005). Paz (2005) builds on the following comment by Navarro Tomás: "Otros [puertorriqueños], por otra parte, reducen la pronunciación de *r* y *l* a un sonido intermedio que no se deja clasificar bajo ninguno de ambos tipos" 'Other [Puerto Ricans], on the other hand, reduce the pronunciation of *r* and *l* to an intermediate sound that cannot be classified as either one' (Navarro Tomás 1974:76). In other words, Navarro Tomás perceived some liquids in post-nuclear position as being *both* rhotic *and* lateral, or something in between. Paz (2005) sought to describe the cross-dialectal differences in the perception of this 'intermediate' liquid sound by analyzing the speech of a single speaker, who read minimal pairs such as *alma* 'soul'-*arma* 'weapon', and *farsa* 'farce'-*falsa* 'false'. The speaker in Paz's study was instructed to read in an 'informal' style and did not produce /r/ with a tap realization, as is the norm in other dialects, but as a continuant. It was found that, although /r/ and /l/ received similar realizations, they were kept distinct. The acoustic difference between etymological laterals (*alma* 'soul') and etymological rhotics (*arma* 'weapon') seemed to lie in the frequency of F3, while F2 and F1 were not different. In words with etymological /r/, F3 presented a steeply falling trajectory and a lower value than F3 in words with etymological /l/. The formant structure of this 'intermediate' rhotic corresponds to that of a retroflex or bunched approximant (Delattre & Freeman 1968, Boyce & Espy-Wilson 1997). Since we do not have clear articulatory evidence to determine whether this sound has a retroflex or bunched realization, we will refer to this rhotic allophone with the terms 'rhotic approximant' and/or 'approximant (r)'. Examples of the productions of the speaker in Paz (2005) are shown in Figure 1. In *triunfa**r*** **'**triumph' and ***a**rma* 'weapon', we observe that the F3 of the coda liquid exhibits a falling trajectory, whereas in *triunfa**l*** **'**triumphal**'** and *a**l**ma* 'soul', it does not.

Most importantly, however, Paz (2005) tested her hypothesis that the contrast between approximant /r/ and /l/ in coda position would be difficult to perceive for speakers of dialects other than Puerto Rican Spanish. This prediction derived from the fact that the two phonemes are produced with a continuant realization in Puerto Rican Spanish, while in other dialects, coda /r/ is realized as a tap. The absence of a continuant or approximant rhotic in other dialects may prevent non Puerto Ricans from correctly identifying /r/, that is, /r/ may sound like /l/ to speakers of other dialects. According to this hypothesis, Puerto Rican listeners should experience no difficulty in identifying word pairs such as *arma* 'weapon' and *alma* 'soul' produced by a Puerto Rican speaker, while listeners of other origins should commit more identification errors. This is precisely what Paz (2005) found. A large number of Puerto Rican (n=30) and Argentinean (n=30) listeners were included in the experiment. Puerto Rican listeners identified the target words with much more accuracy (81% correct identification), than the Argentinean listeners (48% correct identification), whose performance was not different from chance. In other words, approximant /r/ was perceived as being different from /l/ only by the speakers that use and know the rhotic approximant /r/ allophone. Paz's (2005) main conclusion was that coda liquid neutralization in Puerto Rican Spanish is less common than what had previously been reported, and that its alleged high frequency of occurrence may have been due to the fact that most studies were based on transcriptions made by linguists who were native speakers of other Spanish dialects, who may have experienced some difficulty in perceiving this near-merger. Another way to interpret this finding would be to claim that the 'intermediate' sound described by Navarro Tomás was an approximant that may have involved some tongue tip retroflection or tongue body bunching.

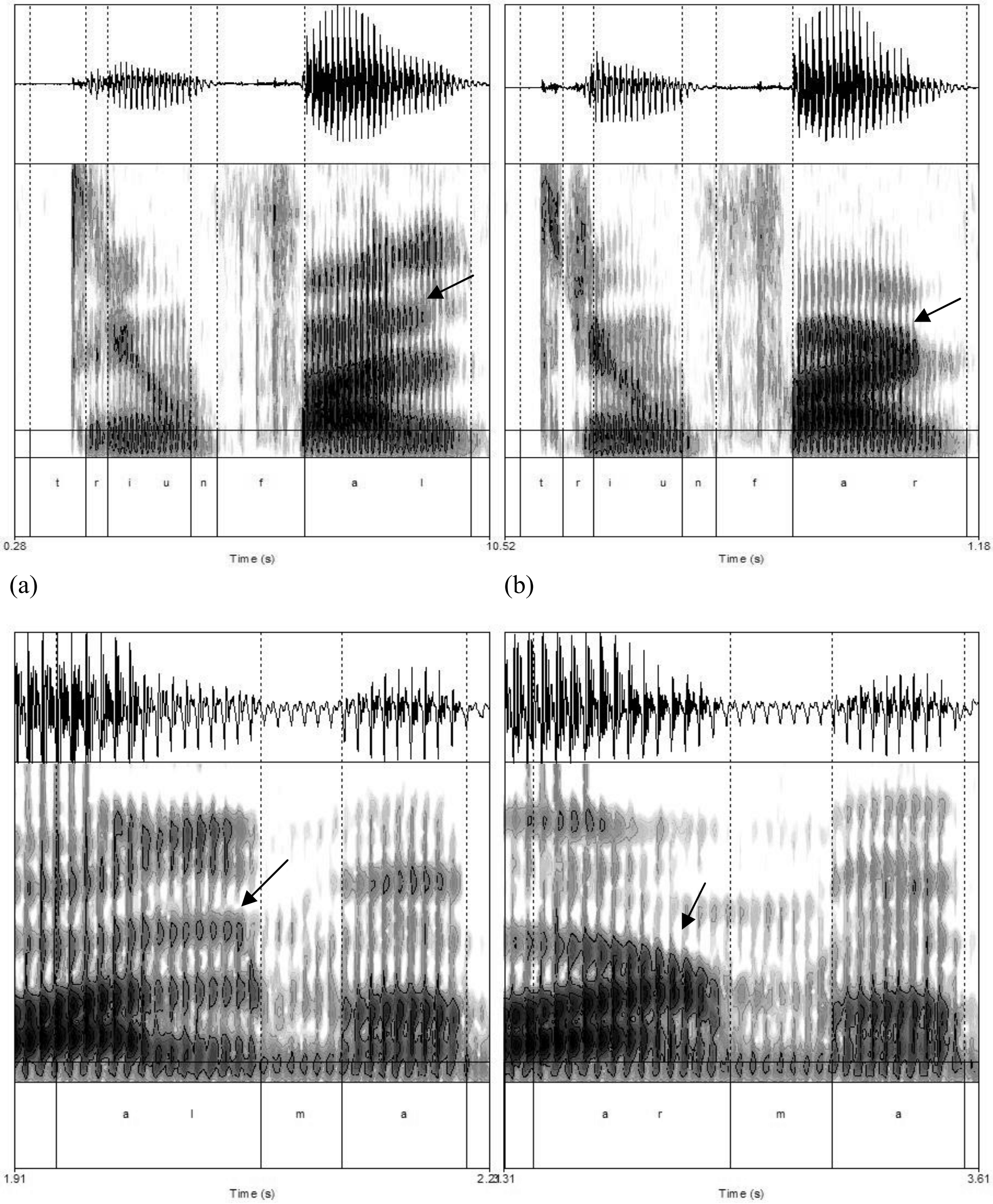


Figure 1. Waveforms and spectrograms of minimal pairs *triunfal* 'triumphal' (a), versus *triunfar* 'to triumph' (b); and *alma* 'soul' (c), versus *arma* 'weapon' (d). Compare F3 trajectories. (From Paz 2005)

In summary, both the enormous variability reported in the literature and Paz's (2005) production and perception findings highlight the fact that positional neutralization of liquids in Puerto Rican Spanish does not follow a categorical, all-or-nothing rule. Positional neutralization does not account for the complete picture. However, Paz's (2005) study is not without problems: (i) the production data is based on the speech of a single speaker who was reading minimal pairs in a highly artificial laboratory setting, and who was not naïve as to the purpose of the experiment; and (ii) the perception study was also based on the production of this single speaker. In order for Paz's (2005) claims to hold, they should be corroborated in further studies in which the speech of more subjects is analyzed.

1.1. Main goals of the present study

The main goal of the present paper is to test Paz's (2005) claims that /r/ and /l/ are phonetically different in Puerto Rican Spanish, even though they may be more similar to each other, and prone to neutralization, in this dialect of Spanish than in others. In other words, we aim to find out whether the two phonemes are completely or incompletely neutralized. More specifically, we are concerned with describing the 'approximant' realization of the rhotic. To do this, we investigate primarily the F1, F2 and F3 trajectories of tokens containing coda /r/, /l/ or no syllable-final liquid (e.g. ***ha**rta* 'fed up', ***al**ta* 'tall', ***at**a* 's/he ties'). This is a preliminary investigation in that only two correlates are studied: (i) duration of the vowel+liquid sequences (which we refer to as 'vocalic portion') and (ii) formant structure (F1, F2 and F3 trajectories). A number of other potential phonetic exponents could be explored, such as those related to the place and manner of articulation of the adjacent consonantal stops: burst duration, type of release and spectral composition of the burst, using center of gravity values (see Plug & Ogden 2003). In Section 2, we describe the experimental procedures including the characteristics of the subjects and the data collection methods. In Section 3, we present the results of our analyses of duration and formant frequencies, for which we implement a relatively new statistical technique for analyzing curves, the Smoothing Spline ANOVA (Gu 2002, Davidson 2006). In Section 4, we discuss the results and present our conclusions.

2. The experiment

2.1. Research questions and hypotheses

Our research questions are the following: Are Puerto Rican 'approximant' /r/ and /l/ acoustically different? If so, where do the acoustic differences between these two sounds lie? How systematic are these differences across speakers and tokens?

The operationalized hypotheses are as follows:

H_0: There are no acoustic differences between approximant /r/ and /l/ in post-nuclear position in Puerto Rican Spanish, i.e. they are completely merged or neutralized.

H_A: There are acoustic differences between approximant /r/ and /l/ in post-nuclear position in Puerto Rican Spanish, i.e. they are not completely neutralized.

Based on Paz's (2005) findings, we hypothesize the acoustic differences between coda approximant /r/ and /l/ to be concentrated in the F3 trajectories. More specifically, we hypothesize approximant coda [r] to present lower F3 values than [l] and/or steep descending trajectories.

2.2. Methods: participants and experimental procedures

In order to collect a sufficient number of tokens for comparison, four speakers were recorded while participating in two tasks: a sentence reading task and an interactive map task (Anderson et al. 1991). Our participants included two females (subjects 1F and 2F) and two males (subjects 3M and 4M) in their mid-twenties to early-thirties. Subjects 1F, 2F and 3M were recorded in Urbana-Champaign, IL and subject 4M was recorded in New York, NY. All four subjects were born and raised in Puerto Rico. At the time they were recorded, they had been living in the USA for 3 to 5 years. They were all students either at the University of Illinois at Urbana-Champaign (1F, 2F, 3M) or at New York University (4M). Three of the four participants are from the San Juan metropolitan area (1F, 3M and 4M) while the fourth is from Mayagüez (2F).

The Map Task (Anderson et al. 1991) consists of an interaction between an instruction giver and an instruction receiver. The instruction giver receives a fictitious map in which there appear several landmarks and a drawn route to be followed from start to finish, passing by all of the landmarks. The names of the landmarks contain instances of the variables under investigation. The instruction receiver is given the same fictitious map, with the same landmarks but without the route. The instruction giver is asked to guide the receiver from start to finish. In this way, both participants, especially the instruction giver, tend to pronounce the names of the landmarks several times. Presumably, attention paid to speech is diverted in map tasks due to the need to carry out a common goal. The participants were paired with another speaker of the dialect. The participant with the role of instruction giver was the one in the pair whose voice was recorded. Landmark names were controlled for the following potential conditioning factors:

(i) Preceding and following (trans-consonantal) vowel: /a/ or /i/. The trans-consonantal vowel was the same as the one that preceded the liquid (e.g. *c*[á]***rt***[a] 'letter', *c*[i]***rqu***[í]*to* 'small circus').

(ii) Lexical stress configuration. The syllable of the target coda liquid was either lexically stressed or unstressed (e.g. *c*[á]***rta*** 'letter', *c*[a]***rtab***[ó]*n* 'drawing triangle').

(iii) Following consonant: /p/, /t/ or /k/. Tokens such as ***car***[t]*a* 'letter', ***bar***[k]*a* 'boat' and ***ar***[p]*a* 'harp' were included.

Nonce words were used in order to control for phonetic context, in cases where there were lexical gaps in Spanish. Since our intention was to measure formant trajectories, and we anticipated that only small effects were to be found, controlling carefully for phonetic context was crucial: place of articulation of the following consonant, preceding and trans-consonantal vowels (due to potential trans-consonantal vowel-to-vowel coarticulation effects) and lexical stress configuration all affect the formant structure. Landmark names such as *Urbanización El M**apa***, *Parque La B**arca**za*, and *Teatro de M**alta*** were used. The participants were given three maps, each having 12 landmarks, for a total of 36 landmark names. The same maps, with different routes were re-used to obtain three repetitions of each token. Thus, approximately 108 tokens were recorded per speaker. It should be noted, however, that speakers uttered the target names different times, or did not pronounce some at all. Thus, on final count, we obtained between 90 and 120 tokens per participant.

In the reading task, the participants were asked to read aloud, as naturally as possible and at a self-selected rate, a sentence list in which the target words were included. Sentences were pseudo-randomized and presented to the speakers on a sheet of paper. The target words in the sentences included the same tokens that were used in the map tasks, with the exception of those with no coda liquids. That is, only target words with coda /r/ and coda /l/ were included here. Each speaker read 24 sentences three times, for a total of 72 sentences per participant. Since the same words were used as in the map tasks, the phonetic context was controlled for the same factors: preceding (and following) vowel, following consonant and lexical stress configuration. For example, sentences such as *Comer **alca**chofas es lo que más le gusta a Juan* 'Eating artichokes is what Juan likes best', or *En aquel **altar** se casaron mis papás hace más de veinte años* 'At that altar, my parents got married more than twenty years ago' were used as stimuli.

In total, 331 tokens with orthographic /r/, 298 tokens with orthographic /l/, and 152 tokens with no orthographic liquid in syllable-coda position were elicited.

The participants were recorded in a quiet room or in a sound treated booth using professional equipment: for 1F, 2F and 3M, a Marantz PMD660 digital solid-state recorder and a Shure SM10A head-worn dynamic microphone were used; for 4M, a Marantz PMD670 solid-state digital recorder and a Shure WH20XLR head-worn dynamic microphone. The speech signal was digitized at 44.1kHz (16-bit) and then transferred to a computer for further analysis. The sound files were then down-sampled at 22.05kHz using the Praat signal-processing package (Boersma & Weenink 2006).

3. Results

3.1. Auditory analysis

As a first step towards a description of the phonetic differences between orthographic /r/ and /l/ tokens, an auditory, impressionistic classification was conducted (for a similar methodology, see Plug & Ogden 2003). As discussed in Section 1, coda /l/ and /r/ can have several phonetic realizations in Puerto Rican Spanish ranging from a tap, to an approximant, to a lateral. Therefore, prior to any instrumental analysis, tokens were classified according to how they were *perceived* to be realized. The second author, who is a phonetically-trained native speaker of Puerto Rican Spanish, conducted the perceptual classification. The fact that a native Puerto Rican Spanish speaker performed the perceptual categorization is relevant due to Paz's (2005) findings that the /l/ and /r/ contrast in Puerto Rican Spanish is difficult to perceive for speakers of other dialects. Tokens were classified according to the following six categories:

(0) Deleted or elided liquid
(1) Tap or trill realization
(2) Inconclusive tap/approximant percept
(3) Approximant realization. This realization is similar to the one Paz (2005) described as a retroflex or bunched rhotic, and was mistakenly identified as /l/ by Argentinean listeners
(4) Inconclusive approximant/lateral percept
(5) Lateral realization: clear /l/ percept

Types (2) and (4) were added, as in Foulkes & Docherty (2000), in order to minimize the imposition of categories on the data. Of course, this does not eliminate subjectivity, but we believe it reduces it. All tokens, including those with etymological /l/ and those with no liquid in coda position (∅), were subjected to the auditory analysis.

	Perceived category					
Orthography	(0) deleted	(1) tap	(2) inconclusive	(3) approximant	(4) inconclusive	(5) lateral
∅	151	1	0	0	0	0
l	1	0	0	9	3	285
r	7	114	11	112	17	70

Table 1. The subjects' productions of coda liquids (/r/, /l/) and no-coda tokens (∅) as perceived by the second author, a phonetically-trained native speaker of Puerto Rican Spanish. Data pooled from all speakers.

Table 1 shows the results of the perceptual categorization of the data by the second author. The table relates orthographic coda liquids as presented in written form to the speakers with the perceived quality of the production of all the speakers, according to the second author. It can be observed that the produced tokens that had no liquid in coda position in their written form were perceived as such in 151 of the 152 cases. Tokens with an orthographic /l/ received a lateral pronunciation in 285 of the 298 cases, they were produced as rhotic approximant in only 9 cases, and with no liquid in post-nuclear position in only one case, all this according to the perceptual categorization conducted by the second author. Orthographic /r/ tokens received a more variable pronunciation: 114 were realized as a tap, 7 as having no coda, 70 as a completely neutralized lateral, and 112 as an approximant rhotic; again, as perceived by the second author. Inconclusive percepts (2) and (4) were selected only in a handful of cases, and thus will not be further discussed.

	Perceived category					
Orthography	(0) deleted	(1) tap	(2) inconclusive	(3) approximant	(4) inconclusive	(5) lateral
1F	0	15	1	47	4	1
2F	1	2	3	19	7	49
3M	0	56	3	28	3	3
4M	6	41	4	18	3	17

Table 2. The subjects' productions of tokens with orthographic coda <r> as perceived by the second author, a native speaker of Puerto Rican Spanish.

Table 2, which provides the distribution of the perceived realizations of coda /r/ for each speaker, reveals the following points: (i) speaker 1F's preferred realization has a rhotic approximant percept; (ii) speakers 3M and 4M prefer taps, which arose perhaps due to the artificialness of the tasks, since, for these two participants, rhotic approximants were much more common in an unscripted conversation and a shadowing task (which are not reported in this study, but cf. Rohena-Madrazo et al. 2006) than in the reading and the map tasks; (iii) complete lateralization or neutralization of /r/ into [l] is common for only one of the speakers. Notice that this speaker (2F) is the only one from Mayagüez, while the other three participants are from the San Juan metropolitan area. This finding may reveal a sub-dialectal difference.

3.2. Instrumental analysis

For the instrumental analysis, we divided the dataset in three groups, which were to be contrasted in the acoustic analysis:

(i) Tokens with no coda. (Orthographic <∅> perceived as token type “0”)
(ii) Tokens with an /l/ that was perceived as a lateral [l]. (Orthographic <l> perceived as token type “5”)
(iii) Tokens with an /r/ that was perceived as a rhotic approximant. (Orthographic <r> perceived as token type “3”)

As stated above, the main goal of the present paper is to describe the acoustic differences between coda laterals and approximant rhotics in Puerto Rican Spanish. Presumably, these are the two allophones that were found in Paz (2005) to be distinguishable by Puerto Rican but not by Argentinean listeners. Auditory classification was necessary in order to minimize the introduction of noise in the instrumental analysis. 70 tokens of coda /r/ were perceived as being produced as laterals, and thus completely neutralized. These tokens were excluded from the present analysis and are thus left for future investigations.

3.2.1. Duration of vocalic portion (vowel+liquid sequence)
3.2.1.1. Analytical procedures

Previous research on post-vocalic rhotics in Dutch (Plug & Ogden 2003) has shown that exponents of rhoticity are variable and may be distributed over a large region, that is, beyond the area where /r/ is thought to appear. For this reason, we followed Plug & Ogden (2003) and Nooteboom (1972) in not segmenting the signal into discrete vowel+liquid+stop chunks. Although identifying the region of laterals in /l/ tokens was not always difficult, identifying the boundary between vowels and approximants in approximant /r/ tokens would have been very difficult and arguably meaningless. All tokens were segmented and labeled as follows: (1) ONSET: The onset of the vowel preceding the target liquid; this boundary corresponds in the spectrogram to the emergence of F2 energy; (2) OFFSET: The offset of the vowel+liquid sequence or, in other words, the onset of the closure for the stop consonant following the liquid; this boundary corresponds in the waveform to a steep amplitude fall and in the spectrogram to the loss of F2 structure. The duration of vowel+liquid sequences (vocalic portions) was measured by subtracting the time value of OFFSET from that of ONSET.

3.2.1.2. Results

Table 3 (a-d) provides the means and standard deviations of the duration of vocalic portions as a function of coda type (/r/, /l/ and ∅), grouped in function of the preceding vowel and following consonant. Data from the four speakers is shown separately. Values were submitted to four different three-factor ANOVAs, one for each speaker, with coda type (/r/, /l/ and ∅), vowel (/a/, /i/), and consonant (/p/, /t/, and /k/) as main factors. The effects of stress were not explored. The ANOVAs revealed coda type to have a significant effect on the duration of the vocalic portions for the four participants: 1F ($F_{(2,182)}$=12.66, p<.001); 2F ($F_{(2,117)}$=10.65, p<.001); 3M ($F_{(2,119)}$=13.38, p<.001), and 4M ($F_{(2,114)}$=16.20, p<.001). Of the other factors, only vowel was significant and only for speaker 2F ($F_{(1,117)}$=10.10, p<.001). The effects of coda type were then submitted to post-hoc tests (Tukey HSD). Post-hoc tests revealed the following patterns: (i) for subject 1F, the duration of each coda type is different (/r/ ≠ /l/ ≠ ∅); (ii) for the two male participants 3M and 4M, /l/ vocalic portions were different from both /r/ and ∅ ones, but those for /r/ were not different from those for ∅ (i.e. /l/ ≠ /r/ = ∅); (iii) for the last speaker, 2F, only /l/ and ∅ were different, while /r/ tokens were not different from any other configuration.

	∅		r		l	
Sequence	Mean	SD	Mean	SD	Mean	SD
/a_p/	80	45	119	33	129	37
/i_p/	106	23	96	18	115	27
/a_t/	90	53	133	37	139	44
/i_t/	92	33	123	26	132	32
/a_k/	93	49	115	25	129	40
/i_k/	94	22	105	38	132	22
Total Mean	93	36	115	32	131	38

(a) Speaker 1F

	∅		r		l	
Sequence	Mean	SD	Mean	SD	Mean	SD
/a_p/	76	41	82	21	106	39
/i_p/	66	41	79	14	92	20
/a_t/	75	30			117	28
/i_t/	67	18			90	23
/a_k/	77	39	138		111	34
/i_k/	74	22	93	22	80	23
Total Mean	73	30	86	22	99	30

(b) Speaker 2F

	∅		r		l	
Sequence	Mean	SD	Mean	SD	Mean	SD
/a_p/	63	43	83	3	117	17
/i_p/	113	37	86	12	113	21
/a_t/	91	33	91	14	112	21
/i_t/	81	13	102	34	100	29
/a_k/	85	29	90	17	105	35
/i_k/	77	19	66	4	101	24
Total Mean	84	32	86	19	108	25

(c) Speaker 3M

	∅		r		l	
Sequence	Mean	SD	Mean	SD	Mean	SD
/a_p/	89	28	89	18	140	30
/i_p/	111	13	90	35	111	34
/a_t/	100	21			137	26
/i_t/	97	4	97	10	123	22
/a_k/	117	30			128	38
/i_k/	106	25	114	30	139	34
Total Mean	103	23	98	28	130	32

(d) Speaker 4M

Table 3. Duration of vowel+liquid intervals as a function of speaker, coda type (∅,r, l), preceding vowel (/a/ and /i/) and following consonant (/p/, /t/, /k/).

It seems that duration increases in the progression ∅ < rhotic < lateral, but the rhotic approximant and ∅ are more similar to each other than are the lateral and the rhotic. These results support the following interpretation: The production of /l/ may require a higher constriction degree, and thus may take more time to realize, than the production of approximant rhotics. Approximant rhotics may also involve a tongue raising gesture, but without constriction. The gesture may be lower in magnitude and thus there may not be a clear increase in duration vis-à-vis regular inter-consonantal vowels.

3.2.2. Formant structure differences
3.2.2.1. Analytical procedures

Formant trajectories were automatically extracted from the vocalic portion (vowel+liquid interval = OFFSET – ONSET) using a Praat script. Time was normalized by extracting formant values (F1, F2, and F3) from seven equidistant points within the vocalic portion: (p0= 25%; p1= 37.5%; p2= 50%; p3= 62.5%; p4= 75%; p5= 87.5%; p6= 100%). Formants were extracted using a Burg LPC-spectra formant estimation function. Praat's recommended settings were selected: for male speech, a maximum of 5 formants in the 0-5kHz region, with a rectangular window of .025 seconds and pre-emphasis from 50Hz; for female speech, a maximum of 5 formants in the 0-5.5kHz region, with a rectangular window of .025 seconds and pre-emphasis from 50Hz. Formant values in Hertz were converted to Bark units (Zwicker 1961). The Bark scale, which is a logarithmic psychoacoustic scale, was selected since it is believed to better reflect human perception. There were a total of 3 matrices for each token: one for each formant (F1, F2 and F3), with two columns (Time*Bark) and seven rows (7 time points) each. The matrices were then submitted to R for statistical computing and graphics.

Formant analyses were conducted in two steps. First, matrices were submitted to various Smoothing Spline ANOVAs (Gu 2002, Davidson 2006). Second, two static points in the formant structure were submitted to regular factorial ANOVAs, in order to confirm the results found in the Smoothing Spline ANOVAs using more traditional tests. The Smoothing Spline ANOVA (SSA) is a statistical technique that has been developed for the holistic statistical comparison of entire curves. Formant trajectories are usually studied by comparing frequency values at static time points, i.e. by performing ordered regular ANOVAs or t-tests on values extracted from static time points. However, formant curves are dynamic elements and thus a statistical analysis that takes dynamicity into account should better reflect their nature. The SSA has been recently introduced to speech research for the analysis of tongue curve shapes extracted from Ultrasound in Davidson (2006). More recently, Baker (2006) and Nycz & De Decker (2006) have advocated its use in formant explorations.

The SSA has two basic parts: the Smoothing Spline and the ANOVA. The smoothing spline is a type of natural cubic spline that connects points (knots), here, *Time * Bark* coordinates. The smoothing spline finds the best-fit curve for a series of knots. The smoothing spline is reached by comparing two terms: one that fits the data and one that penalizes the best-fit curve for not having an appropriate amount of smoothness. In this way, a smooth best-fit curve may be automatically determined. The ANOVA will then compare the smoothing parameters. The SSA has the form: $f = \mu + \beta x$ + main groups effects + smooth(x) + smooth(x; group). The main group effects correspond to the smoothing splines for each dataset (in our case, this would be the spline for vowel+rhotic and vowel+lateral). The smooth(x) is the best-fit spline for the aggregate data. The smooth(x;group) is the interaction term which is the spline representing the difference between a main effect and the general smooth(x) spline. The interaction term is the one to determine whether the smoothing splines (best-fit curves) representing each group in the dataset are significantly different or not. The SSA does not return an *F*-value nor a *p*-value. Significance is determined by comparing smooth(x) and smooth(x;group). If the two terms are of the same magnitude, then there is a significant difference at some point in the curves. One then needs to determine where the difference between the two curves lies along the x-axis. For this, 95%-Bayesian-confidence intervals may be calculated, and plotted along the mean best-fit curve. The point(s) at which the 95%-Bayesian-confidence intervals for the two (or more) lines do *not* overlap is (are) the point(s) where the two (or more) curves are different.

Best-fit curves were determined for each of two groups according to only two of the three levels of the main factor 'coda type', that is, the dependent variable was the curves themselves and the factor was coda type (vowel+/r/; vowel+/l/). 95%-confidence intervals were generated along the best-fit splines for each group. Results are plotted in Figure 2(a-h). Significance is reached if the 95%-Bayesian-confidence intervals do not overlap. The three formants (F1, F2 and F3) are plotted on the same figure, unless otherwise indicated. Results are presented for each of the four speakers separately.

Only paired comparisons are shown, since our goal was to see whether coda-/r/ tokens are significantly different from those for coda-/l/. Data were averaged across consonant and lexical stress configurations. Data were grouped according to the preceding vowel. In the figures, thick lines represent the best-fit spline, while the two thin lines surrounding them represent the confidence intervals.

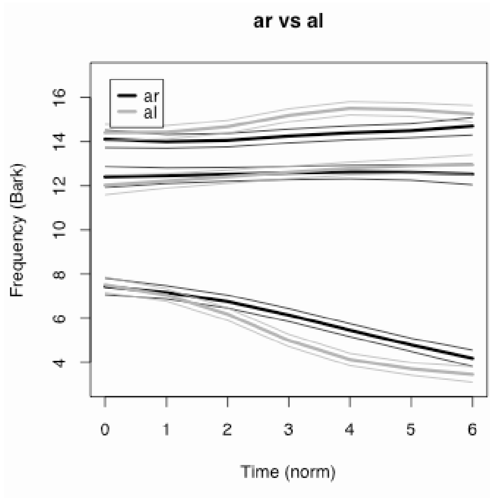
3.2.2.2. Results

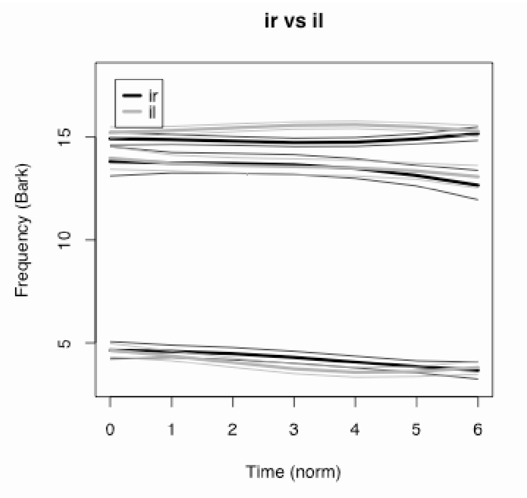
Figure 2(a-b) shows the SSA results for subject 1F for +/r/ vs. +/l/ vocalic portions, grouped by the preceding vowel: the left panel shows /a/-sequences, the right panel /i/-sequences. In the left panel it can be observed that the confidence intervals for the two F1 trajectories are significantly different: F1 is higher in +/r/ sequences than in +/l/ ones, and this is so from time point *p2* onwards; prior to this, the F1 trajectories are not different. Descending F1 trajectories for both types of sequences suggest that a tongue raising gesture is involved in both cases. Differences in the shape of the curves show that +/l/ sequences involve more raising than +/r/ sequences. In the right panel, F1 also presents a descending trajectory in both cases, but this time, the two lines are different only for a very brief period, and differences, although in the expected direction, are smaller. F3 values, which are a correlate of rhoticity (possibly retroflection or bunching), are also different in both panels: F3 is significantly lower from time point p2 to time point p5 in +/r/ cases than in +/l/ ones. Results for F1 and F3 are important in that they suggest that +/r/ and +/l/ sequences involve different tongue configurations. Notice also that all lines are equal at the beginning (time points p1 and/or p2, close to vowel midpoints) and the end (where consonantal locus plays a significant role), but differ in the middle (where the magnitude of the liquid gestures are at their greatest).

Figure 2(c-d) shows a similar situation for speaker 2F: there are differences in the F1 trajectories around the middle of the vocalic interval, and in the expected direction in the F3 trajectories, but not in those for F2. Thus, there are differences in tongue configuration, but they do not seem to affect tongue fronting, for which F2 is a correlate. They may affect tongue raising (F1) and tongue-tip retroflection or tongue-body bunching (F3). The right panel in Figure 2(c) does not show F1. This is due to the fact that, in order to better observe F3 differences, the range on the y-axis was reduced. Although not shown, F1 trajectories for speaker 2F in the lower panel are similar to those for speaker 1F.

Figure 2(e-f) shows the data for speaker 3M. For this subject the left panel reveals differences in F3 curves between +/r/ and +/l/ tokens in the expected direction but they are reached only at the vocalic portion's midpoint and seem to be very small. Most striking is the fact that /i/+liquid sequences are not different from each other in F3. There seem to be differences only in F2 and for only a very brief period. Recall that different consonantal configurations have been pooled here and thus this may be obscuring the small pattern that was revealed in the cases of the other speakers.

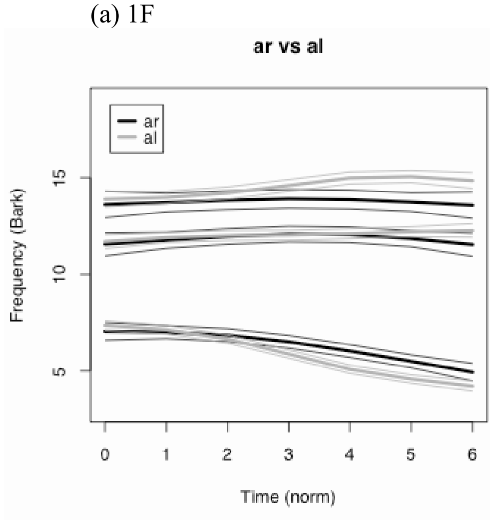
Figure 2(g-h) presents the trajectories for speaker 4M. For this participant, there seem to be no significant differences between the trajectories, but there seem to be trends toward the same direction. That is, there are no differences in F2 and F3 trajectories, but there seem to be small effects in F1, both in the context of /a/ (left panel) between time points p4 and p5, and in the context of /i/, not shown here.

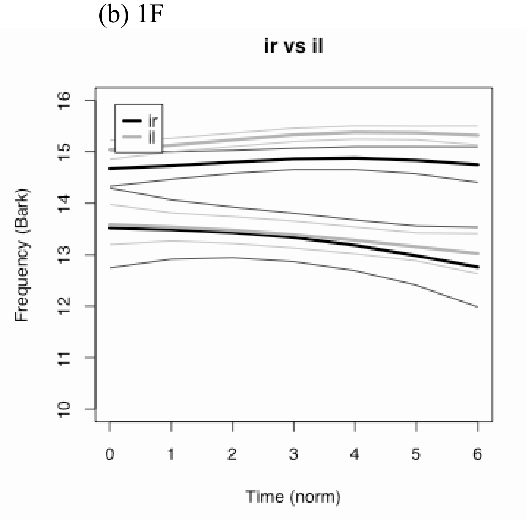




. (a) 1F

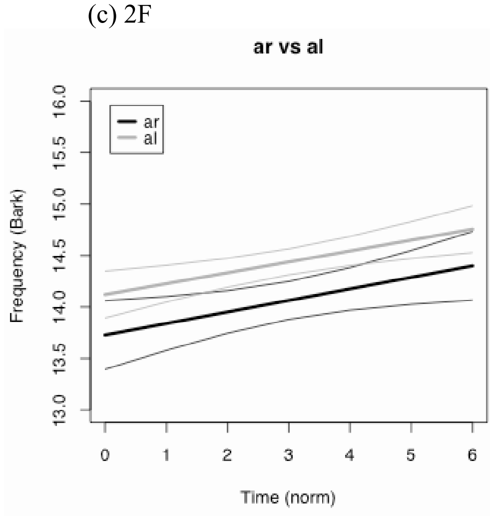
(b) 1F

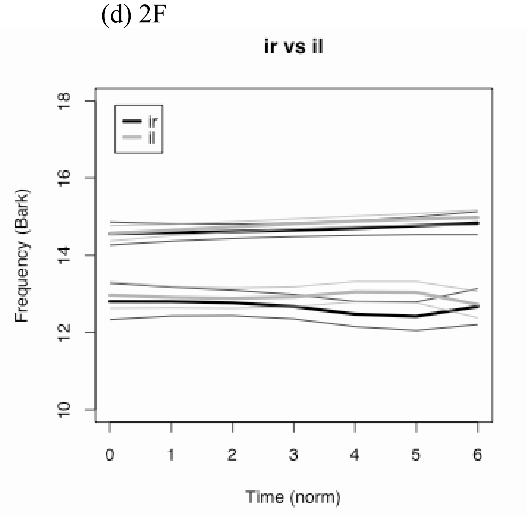




(c) 2F

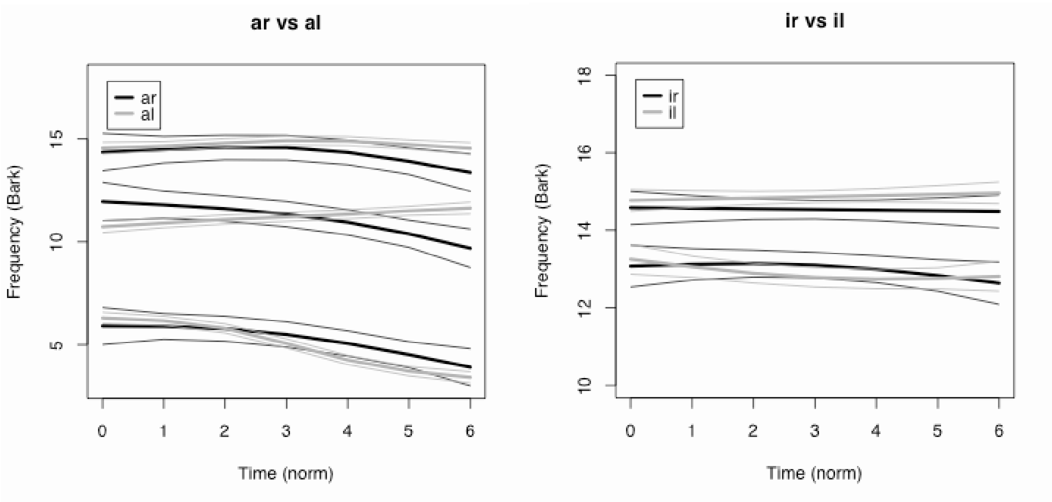
(d) 2F





(e) 3M

(f) 3M



(g) 4M (h) 4M

Figure 2. Formant trajectories (F1, F2, F3) of vowel+liquid sequences (/r/ vs. /l/) in Bark units, as calculated by the smoothing function in the Smoothing Spline ANOVA.[1]

3.2.2.3. ANOVAs on static time points

In order to assess whether these phonetic differences hold using a more traditional statistical method, we performed statistical comparisons of formant values extracted from two time points in the trajectories, with the use of regular factorial ANOVAs. The results of the SSAs are thought to be both reliable and interpretable. However, since they have been newly introduced to speech research, we thought it would be beneficial to provide a point of comparison with more traditional techniques. The statistical tests were conducted as follows. First, averages for each token were calculated of: (i) time points p0, p1 and p2 for each formant (thus we obtained average F1, F2 and F3 values for the region of the vowel's midpoint), and (ii) time points p3, p4 and p5 for each formant (F1, F2 and F3 values for the region corresponding to the coda liquids, in the case of vowel+liquid sequences). The first time region (p0, p1 and p2) will be referred to as Time 1 (T1), and the second time region (p3, p4 and p5) will be referred to as Time 2. The data from each speaker were submitted to 6 different one-way ANOVAs, with formant (F1, F2 or F3) values in T1 or T2 as dependent variables, and 'coda type' (/r/, /l/, ∅) as the main factor. The results are summarized in Table 4.

Results revealed that F1 values at T1 did not vary as a function of coda type (/r/, /l/, ∅), except marginally for speaker 4M, while F1 values at T2 did vary as a function of coda liquid for all speakers. ANOVAs that revealed significant effects were submitted to post-hoc comparisons using a Tukey HSD. Post-hoc tests on T2 data revealed the following significant differences: for speaker 1F, all three levels were different from each other at T2; for speaker 2F, laterals were different from ∅ tokens, as well as from rhotic tokens, while rhotics were not different from ∅ tokens; for speaker 3M, ∅ tokens were different from the other two levels, while rhotics and laterals were not different from each other; and for speaker 4M, laterals were different from the other two levels, while rhotics were not different

[1] In Figure 2, time is normalized by extracting formant values from seven equidistant points in vowel+liquid sequences. Smoothed, averaged formant trajectories are represented by the thicker lines, while thinner lines represent 95% confidence intervals. Significance is reached when confidence intervals of two different trajectories do not overlap. The different panels are organized as follows: (a-b) speaker 1F, (c-d) speaker 2F, (e-f) speaker 3M, and (g-h) speaker 4M. The panels on the left display data in /a/ contexts while the panels on the right display data in /i/ contexts. Panels (a), (b), (c), (e), (g) display F1-F3 trajectories, while panels (d), (f) and (h) display only F2 and F3, and panel (e) displays only F3. In the cases were only the upper formants are shown, this was done because differences were observable only by maximizing the y-scale.

from ∅ tokens. In other words, F1 values at T2 were different between +/r/ and +/l/ sequences for speaker 1F, 2F and 4M, but not for 3M. Anticipatory F1 coarticulation was found only for speaker 4M.

	Formant			
	F1 Time period		F3 Time period	
Speaker	T1	T2	T1	T2
1F	$_{2,143}$ = 1.4	$_{2,143}$ = 42.4 **	$_{2,143}$ = 14.7 **	$_{2,143}$ = 39.4 **
2F	$_{2,130}$ = 0.3	$_{2,130}$ = 11.4 **	$_{2,130}$ = 1.9	$_{2,130}$ = 12.2 **
3M	$_{2,134}$ = 0.3	$_{2,134}$ = 5 **	$_{2,134}$ = 2.4	$_{2,134}$ = 4.1 *
4M	$_{2,129}$ = 4.3 *	$_{2,129}$ = 18.5 **	$_{2,129}$ = 2	$_{2,129}$ = 4.7 *

Table 4. ANOVA results (df = F-value) of F1 and F3 formant trajectories at two time points (T1 and T2) as a function of speaker (1F, 2F, 3M, 4M). Significance is marked with asterisks: p<.05 = *; p<.001 = **

With regards to the second formant (F2), results revealed that liquid type does not seem to have any effect on its shape. Only one test revealed marginally significant results: F2 at T1 for speaker 2F ($F_{(2,129)}$=3.9, p<.05), but it is unclear why. The role of chance cannot be excluded here due to the large number of comparisons.

F3 values at T1 did not vary as a function of coda liquid type, except for speaker 1F, while at T2 it did vary significantly. In order to investigate whether rhotic tokens were different from lateral tokens, we submitted the T2 ANOVAs to post-hoc Tukey HSD tests. Tukey HSD tests revealed the following findings: at T2, for speaker 1F, each level was different from the others; for speaker 2F, laterals were different from the other two levels, while rhotics were not different from ∅; for speaker 3M, rhotics were different from laterals, while ∅ tokens were not different from any of the other two levels; and for speaker 4M, only rhotics were different from no coda tokens, but rhotics were not different from laterals. In other words, the results revealed that the type of post-nuclear liquid significantly affects F3 values: rhotics trigger lower F3 values than laterals do, at least for three speakers (1F, 2F and 3M). There was evidence for anticipatory F3 coarticulation only for speaker 1F.

In summary, the inspection of the trajectories generated by the smoothing spline part of the SS ANOVA, and the results of regular ANOVAs on T1 and T2 provided a very similar picture: differences between +/r/ and +/l/ tokens concentrate on F1 and F3. In general, F1 was found to be higher in +/r/ than in +/l/ sequences, which may indicate a tongue raising gesture of greater magnitude in +/l/ sequences than in the former. Regarding F3 trajectories, they were found to be lower in +/r/ than in +/l/ tokens. Low F3 values have been associated with rhoticity, possibly due to tongue retroflection or tongue body bunching (Plug and Ogden 2003, among others).

4. Discussion and conclusions

This article has shown that Puerto Rican Spanish speakers use a rhotic allophone in post-nuclear position that can be characterized as an approximant. Paz (2005) found that speakers of dialects other than Puerto Rican Spanish experience some difficulty in discriminating between rhotic approximants and laterals in coda position in the production of a Puerto Rican speaker. Thus, there seems to be evidence of incomplete neutralization of the two phonemes: while /r/ and /l/ are maximally different in, say, syllable-initial position, their acoustic difference is reduced, but not lost, in syllable-final position.

Using data from four speakers, two females and two males, elicited from sentence reading and map tasks, acoustic differences were found between tokens with approximant post-nuclear /r/ and tokens with a coda lateral. The acoustic exponents we analyzed were as follows: (i) duration of the vocalic interval (vowel+liquid sequence); (ii) F1, F2 and F3 dynamic trajectories from the 25% to the 100% duration points of the vocalic portion; and (iii) F1, F2 and F3 static values from two time points. Differences between coda approximant /r/ and /l/ were found for all of these correlates.

Table 5 presents a summary of the results showing where approximant rhotics and laterals were different for every given acoustic correlate, for each speaker separately. Differences between tokens in the context of /a/ and those in the context of /i/, found in the SSA analyses, are not included in this table's presentation: if differences were found in one vocalic context, we count the difference to hold.

Acoustic parameter	Speaker			
	1F	2F	3M	4M
Duration of vocalic portion	√		√	√
F1 (trajectory)	√	√		√
F2 (trajectory)				
F3 (trajectory)	√	√	√	
F1 (static T2)	√	√		√
F2 (static T2)				
F3 (static T2)	√	√	√	
Total √	5	4	3	3

Table 5. Summary of significantly different comparisons (+/r/ vs. +/l/ tokens) for each acoustic parameter, as a function of speaker: (i) duration of vocalic portion, (ii) observed dynamic trajectories of F1, F2 and F3, and (iii) static values of F1, F2 and F3 at T2, which corresponds to the portion where the liquid gesture reaches its peak.

Table 5 shows that coda approximant /r/ and /l/ are different for each speaker at least on some dimension: speaker 1F presents differences in all measured acoustic correlates; speaker 2F, only in those that have to do with formant structure, speaker 3M, in duration and F3 (both dynamic and static); and speaker 4M, in duration and F1 (both dynamic and static).

The acoustic evidence that we have gathered is compatible with the following articulatory interpretation: in +/r/ tokens, there may be an approximation of the tongue tip to the post-alveolar region, that is, a tongue body raising gesture, which causes F1 lowering, though less than in +/l/ tokens; and there may also be some degree of tongue tip retroflection, or tongue body bunching, which may be responsible for F3 lowering.

It is obvious also that the acoustic correlates of the Puerto Rican approximant /r/ are variable and only slightly different from those of /l/. On the one hand, there seem to be systematic acoustic differences between +/r/ and +/l/ tokens for at least three of the four speakers we recorded. Both +/r/ and +/l/ sequences involve a tongue body raising gesture, and are continuant sounds. Thus, they display similar acoustic configurations. However, our findings suggest that +/r/ and +/l/ tokens are not completely merged, at least for our three speakers from the San Juan metropolitan area. On the other hand, notice that complete neutralization of post-nuclear /r/ and /l/ seems to have also been found. This is the case for the speaker from Mayagüez, speaker 2F (see also Ramos-Pellicia 2004 for a recent report of complete neutralization in rural Puerto Rican Spanish).

Another important point we may raise is the following: previous research has shown that adult listeners find it difficult to discriminate between non-native speech sounds, especially when two non-native speech sounds can be perceptually assimilated to one native speech category (e.g. Best & Strange 1992). This seems to be what occurs with Japanese listeners' perception of American English /l/ and /r/ (e.g. Aoyama et al. 2004). Regarding the Puerto Rican approximant /r/, Paz showed that Argentinean listeners experienced difficulty in distinguishing this sound from /l/, while native Puerto Ricans did not. Argentinean Spanish does not have an approximant /r/ allophone. The phonemic tap is phonetically realized as a tap. Thus, /r/ and /l/ phonetic differences are preserved in all positions in Argentinean Spanish. This fact may have provoked the perceptual assimilation (or fusion) of the two categories by the Argentinean hearers in Paz's study. However, in the present paper, it has been shown that acoustic differences between coda approximant /r/ and /l/ are robust, at least for speakers from the San Juan metropolitan area. Thus, we may conclude that Puerto Rican listeners, who know/use this rhotic allophone, are trained to distinguish it from /l/ (Paz 2005).

The present findings question the alleged frequency of +/r/ and +/l/ neutralization in Puerto Rican Spanish. Coda liquid neutralization may have been over-reported in previous studies due to the fact that results have often been based on phonetic transcriptions conducted by speakers of non-Caribbean dialects who seem to have difficulty distinguishing between +/r/ and +/l/ sequences.

References

Anderson, Anne, Miles Bader, Ellen G. Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Cathy Sotillo, Henry Thompson & Regina Weinert (1991) The HCRC Map Task Corpus. *Language and Speech*, 34(4):351-366.

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada & Tsuneo Yamada (2004) Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32(2):233-250.

Baker, Adam (2006) Quantifying diphthongs: a statistical technique for distinguishing formant contours. Paper presented at NWAV35, November. Ohio State University, Columbus, OH.

Best, Catherine T. & Winnifred Strange (1992) Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20:305-330.

Boersma, Paul & David Weenink (2006) Praat: doing phonetics by computer. Digital signal-processing tool retrieved from http://www.fon.hum.uva.nl/praat/.

Boyce, Susan & Caroline Espy-Wilson (1997) Coarticulatory stability in American English /r/. *Journal of the Acoustical Society of America*, 101(6):3741-3753.

Charles-Luce, Jan (1985) Word-final devoicing in German: Effects of phonetic and sentential context. *Journal of the Acoustical Society of America*, 13:309-324.

---------- (1993) Effects of semantic context on voicing neutralization. *Phonetica*, 50:28-43.

Davidson, Lisa (2006) Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America*, 120(1):407-415.

Delattre, Pierre & Donald Freeman (1968) A dialect study of American r's by x-ray motion picture. *Linguistics*, 44:29-68.

Foulkes, Paul & Gerry Docherty (2000) Another chapter in the story or /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics*, 4(1):30-59.

Gu, Chong (2002) *Smoothing Spline ANOVA Models*. New York: Springer.

Lipski, John (1994) *Latin American Spanish*. New York: Longman. (Spanish translation: *El español de América*. [1996] Madrid: Cátedra.)

López Morales, Humberto (1983) Lateralización de /-r/ en el español de Puerto Rico. *Philologia Hispanensia in honorem Manuel Alvar*, 387-398. Madrid: Gredos.

Navarro Tomás, Tomás (1948) *El español en Puerto Rico*. San Juan: Editorial de la Universidad de Puerto Rico. (3rd edition, 1974)

Nooteboom, Sieb G. (1972) *Production and perception of vowel duration: a study of durational properties of vowels in Dutch*. Unpublished PhD Dissertation, Rijksuniversiteit Utrecht.

Nycz, Jennifer & Paul De Decker (2006) A new way of analyzing vowels: comparing formant contours using Smoothing Spline ANOVA. Paper presented at NWAV35, November. Ohio State University, Columbus, OH.

Paz, Mercedes (2005) Retroflection of post-nuclear /r/ in Puerto Rican Spanish. Paper presented at the 9th Hispanic Linguistics Symposium, November. Pennsylvania State University, University Park, PA.

Penny, Ralph (2000) *Variation and change in Spanish*. Cambridge: Cambridge University Press.

Plug, Leendert & Richard Ogden (2003) A parametric approach to the phonetics of post-vocalic /r/ in Dutch. *Phonetica*, 60:159-186.

Port, Robert & Penny Crawford (1989) Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17:257-282.

Port, Robert & Michael O'Dell (1985) Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13:455-471.

Ramos-Pellicia, Michelle F. (2004) *Language contact and dialect contact: cross-generational phonological variation in the Midwest of the United States*. Unpublished PhD Dissertation, Ohio State University.

Rohena-Madrazo, Marcos, Miquel Simonet & Mercedes Paz (2006) The vernacular in the laboratory. Paper presented at NWAV35, November. Ohio State University, Columbus, OH.

Slowiaczek, Louisa & Daniel A. Dinnsen (1985) On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13:325-341.

Vaquero, María (1996) *El español de América I: pronunciación*. Madrid: Arco Libros.

Warner, Natasha, Allard Jongman, Joan Sereno & Rachèl Kemps (2004) Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics*, 32:251-276.

Zwicker, Eberhard (1961) Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America*, 33:248.

Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology

edited by Laura Colantoni and Jeffrey Steele

Cascadilla Proceedings Project Somerville, MA 2008

Copyright information

ISBN 1-57473-424-9 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Simonet, Miquel, Marcos Rohena-Madrazo, and Mercedes Paz. 2008. Preliminary Evidence for Incomplete Neutralization of Coda Liquids in Puerto Rican Spanish. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, ed. Laura Colantoni and Jeffrey Steele, 72-86. Somerville, MA: Cascadilla Proceedings Project.

or:

Simonet, Miquel, Marcos Rohena-Madrazo, and Mercedes Paz. 2008. Preliminary Evidence for Incomplete Neutralization of Coda Liquids in Puerto Rican Spanish. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, ed. Laura Colantoni and Jeffrey Steele, 72-86. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1715.