A REAL-TIME CASE STUDY OF RHOTIC ACQUISITION IN SOUTHERN BRITISH ENGLISH

Rachael-Anne Knight[†], Christina Villafaña Dalcher[†], and Mark J. Jones[‡]

†City University, London, ‡University of Cambridge
knight@city.ac.uk, cvd1@city.ac.uk, mjj13@cam.ac.uk

ABSTRACT

Development in the production of /r/ is attested in a speaker of Standard Southern British English (SSBE) between the ages of 3;8 and 3;11. The progression towards adult-like apical approximant /r/ is manifested along multiple dimensions and primarily involves a gradual raising of F2 and lowering of F3. In addition to changes in absolute formant frequencies, however, we find that this speaker's development involves reduction in variation of her output, elimination of apparent [w] substitutions concomitant with increased labiodental realisations, and decrease of F3-F2 distance. This latter acoustic cue is worth exploration, as F3 may remain stable between perceptibly different outputs. Moreover, the data show that development of /r/ may include a shift in the relative salience of F3 by an increase in that formant's amplitude, again compensating for non-lowering of F3 to canonically /r/-like frequencies.

Keywords: Rhotics, Acquisition, Child /r/-production

1. INTRODUCTION

1.1. The acquisition of /r/

Research into phonological acquisition suggests that adult-like English approximant /r/ does not emerge until around the age of 4;5 (after most other sounds) and remains highly variable before being mastered at around 6;0 [7], [15]. There has, however, been relatively little research focussed on the acquisition of /r/, and much of that which has been conducted is of limited use: it is out-of-date, considers only American children, focusses on absolute formant measurements, or lacks acoustic analysis altogether, e.g. [2], [14]. An additional problem in previous studies is the tendency to classify any mispronunciation of /r/ as [w]-like. Some children may indeed substitute [w] for /r/, though instrumental confirmation of this is lacking [10], but it is also likely that child speakers of Standard Southern British English (SSBE) actually substitute a labiodental (rather than a labial-velar) approximant, a common occurrence on the developmental pathway towards apical /r/ [9], [12].

1.2. Acoustics of English /r/ and other approximants

The standard description of /r/ in SSBE is a voiced postalveolar approximant, where the tongue tip is in wide approximation to the region of the palate behind the alveolar ridge. However, studies of American English /r/ show that speakers use many different articulations, on a continuum from a bunched to a retroflex tongue shape [16]. Despite the variety of articulations which may be used to produce an approximant /r/, e.g., [3], [5], [16], the acoustic properties are largely consistent across the articulatory continuum. The most striking acoustic feature for /r/ is a low third formant (F3), which is in close proximity to the second formant's frequency (F2) [6]. Acoustic characteristics of the labial-velar approximant /w/, on the other hand, include a lower frequency and lower amplitude F2 in conjunction with a raised F3 [13]. In SSBE, there is a further adult approximant variant of /r/ where F2 and F3 occur in closer approximation than for /w/, but which does not involve the more extreme lowering of F3 and raising of F2 attested in apical /r/ productions [8]. Such alternative realisations – 'labiodental' /r/s – are relevant because they are often regarded as infantilisms, a label whose origins may lie in the fact that developing speakers move gradually from [w]-like articulations of /r/ to more adult-like articulations, producing a labiodental variant along the way.

1.3. The present study

The speaker exhibits a progression towards the production of adult-like apical /r/ in a very short time span, and at a very young age. Acoustic analysis of /r/ realizations at two points in time show that her development is manifested along several dimensions and involves not a categorical change, but a gradual shift from [w]-like to [ɹ]-like outputs. The overall strategy of the speaker involves a consistent raising of F2 and lowering of F3 over time, resulting in incorrect /r/ productions that are not merely [w] substitutions, as suggested in previous studies [7], [11]. We also find that there is more to her development than simple approximation of absolute /r/-like formant targets.

2. METHODS

The data presented is part of a larger study of the production and perception of rhotics by seven child speakers of SSBE. The subject under discussion was tested at two age points, 3;8 and 3;11, and was singled out for further analysis due to the noticeable shift in her productions of /r/. Preliminary analysis at the second testing phase (out of three), shows that other subjects exhibit some, but not all, of the same developmental patterns as the speaker discussed here.

At each phase, 32 tokens with /r/ in various contexts (Cr clusters, singleton onset, and intervocalic) were elicited via a picture-naming task and recorded on a Marantz digital recorder using a unidirectional condenser microphone (one token in Phase 2 was eliminated due to background noise). The subject was administered the Diagnostic Evaluation of Articulation and Phonology (DEAP, [4]), before Phase 1 to ensure that her articulatory abilities were on target for her age. At each testing phase, three tokens with word-initial /w/ were also collected as a control.

Acoustic analysis was carried out in Praat [1]. To maximise the consistency of formant measurements, we created spectra of the ±5ms window surrounding the F2 minima during each production of the target. Spectra were LPC-smoothed using 12 poles, and enabled accurate measurement of formant frequencies and amplitudes. Tokens were labeled impressionistically as [ɹ], [w], or [ʊ] by the second author with a blind check by the first author.

3. ANALYSIS

We analyse a number of different aspects of the data between testing phases: auditory impression, absolute frequencies of F2 and F3, variability of output, F3-F2 distance, and formant amplitudes.

1.4. Impressionistic analysis

Of 32 /r/ tokens in Phase 1 of testing, 17 were classified as [w], 8 as [ʊ], and 7 as [ɹ]. At Phase 2, three months later, no [w]s were heard, labiodental variants increased to 12, and adult-like apical realisations were counted at 19. These impressions are corroborated by mean F3-F2 distance for each output in the testing phases, illustrated in Table 1.

Table 1: Formant frequencies at Phases 1 and 2

	F3-F2, Phase 1	*F3-F2, Phase 2*
/w/ → [w]	2987	2533
/r/ → [w]	2714	n/a
/r/ → [ʊ]	1951	1461
/r/ → [ɹ]	970	1004

1.5. Formant values for /r/ (absolute values)

Table 2 shows a significant difference between each of the first three formant frequencies at each stage of testing ($p<.001$ for each formant).

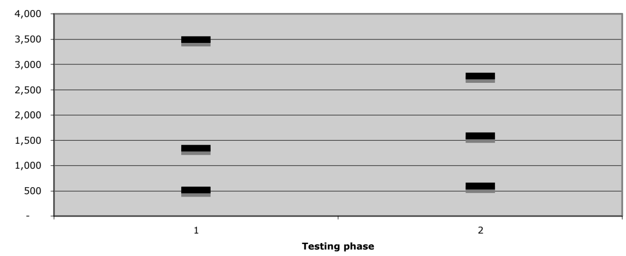
Table 2: Comparison of F values at Phases 1 and 2

Variable	*M*	*SD*	*t*	*df*	*p*
F1			-3.63[a]	47.4[a]	.001
Phase 1	510	106			
Phase 2	587	56			
F2			-3.37	61.0	.001
Phase 1	1342	297			
Phase 2	1585	274			
F3			6.71[a]	36.7[a]	.000
Phase 1	3483	579			
Phase 2	2765	173			

[a] *t* and *df* were adjusted because variances were not equal.

Inspection of the means at each phase indicates that F1 and F2 increased over time while F3 decreased. This rapid change is also illustrated in Figure 1.

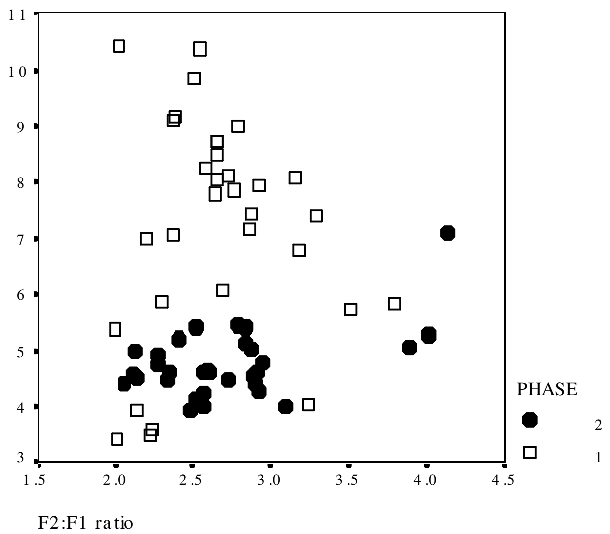
Figure 1: Comparison of F values at Phases 1 and 2



1.6. Variability

To illustrate the variability of productions we examine formant frequency ratios [2] at the two testing phases. Expressing formant values of F2 and F3 relative to F1 (that is F2/F1 and F3/F1), we derive the representation in Figure 2.

Figure 2: F3:F1 ratio vs. F2:F1 ratio at Phases 1 and 2



We see that in the short time span between testing phases, the variability in the speaker's output is significantly reduced, as was suggested by the impressionistic analysis. An added benefit of this analysis is that it allows comparison between the two testing times where variation in the speaker's fundamental frequency is possible.

1.7. F3-F2 distance

A logical sequela to concomitant F2 raising and F3 lowering is that the distance between these two formants is reduced. A comparison of the F3-F2 distance at the two testing phases shows a significant difference over time (p<.000).

Table 2: Comparison of F3-F2 at Phases 1 and 2

Variable	*M*	*SD*	*t*	*df*	*p*
F3-F2			6.72[a]	40.0[a]	.000
Phase 1	2142	754			
Phase 2	1181	287			

[a] *t* and *df* were adjusted because variances were not equal.

However, our speaker does not always raise F2 and lower F3 concurrently, or achieve expected formant targets for correct /r/ as reported in previous studies [2]. Perceptibly different tokens exist where one of these formants occurs at the same frequency, but where the other formant is lowered or raised to reduce F3-F2 distance. Thus the speaker approximates /r/-like output without necessarily producing /r/-like formants.

To illustrate this finding, two tokens of "rock" are compared, one involving an 'incorrect' /r/ realization with formant frequencies of F1=480; F2=960; F3=2580; the other produced 'correctly' (F1=570; F2=1660; F3=2630). LPC-smoothed 10 ms spectral slices are below and sound files incorrect_r.wav and correct_r.wav are attached.

Figure 3: Spectrum of 'incorrect' /r/; heard as [ʋ]

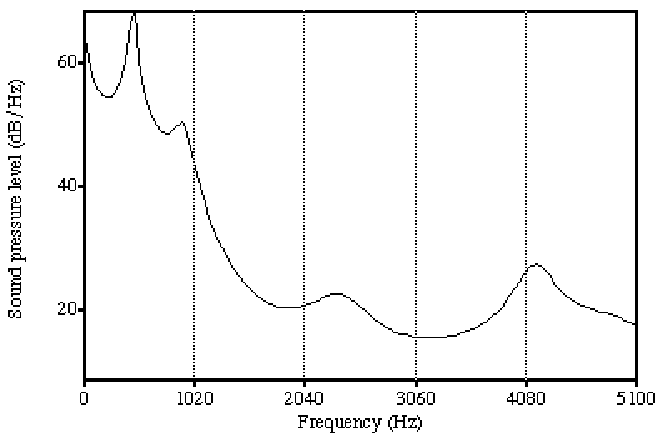
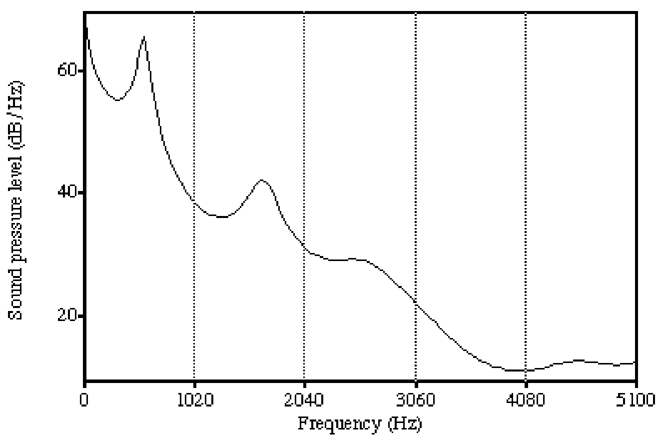


Figure 4: Spectrum of 'correct' /r/; heard as [ɹ]



Although the 'correct' /r/ production in Figure 4 exhibits none of the expected F3-lowering, we note two striking differences between its spectrum and that of the 'incorrect' /r/. First, the F3-F2 distance is considerably narrower in the 'correct' production (970 Hz) than in the 'incorrect' realisation (1620 Hz). Second, higher formant amplitudes differ considerably between the two /r/ outputs. In the labiodental version in Figure 3, F3 is of a lower amplitude than F4. In the 'correct' version, the expected amplitude pattern is seen with F3 exhibiting greater amplitude than F4.

1.8. Formant amplitudes

Examination of formant amplitudes for our subject's /r/-targets reveals significant differences between testing phases for higher formants as shown in Table 3.

Table 3: Comparison of F amplitude at Phases 1 and 2

Variable	*M*	*SD*	*t*	*df*	*p*
F1			.315	61	.754
Phase 1	60.16	6.43			
Phase 2	59.68	5.59			
F2			3.14	61	.003
Phase 1	48.25	7.58			
Phase 2	42.65	6.55			
F3			-3.59	61	.001
Phase 1	24.47	10.36			
Phase 2	33.23	8.94			

F3 amplitude increases considerably from Phase 1 to Phase 2 for this speaker. Such an increase is not likely an artefact of recording environment or speaking style: if it were, we would expect an overall increase in formant amplitudes. Instead, F1 amplitude remains stable across the testing phases while F2 amplitude actually decreases.

4. DISCUSSION

We have analysed various aspects of a child's /r/ production in the latter stages of her fourth year. While there is a general tendency towards raising F2 and lowering F3 to /r/-like targets, we find that this speaker's development is characterised by features other than absolute formant values.

We saw that the subject's /r/-realisations are highly variable, especially in Phase 1, and not simply a matter of [w] substitution. However, a drastic reduction in variation occurred in the 3-month time period between tests, and [w] substitutions ceased to surface. We also saw that even when canonical formant targets (e.g., low F3) for /r/ are not achieved, the speaker incorporates other strategies to produce an auditorially 'correct' /r/, such as exploitation of the F3-F2 distance by means of raising F2, and increasing the amplitude of F3 relative to surrounding formants.

We hypothesise that the greater amplitude of F3 in target-like /r/ outputs, in combination with a narrower F3-F2 gap, result in impressionistically correct /r/ even when there is a failure to lower F3. If this is supported by further testing, we must assume F3 lowering on its own is only one ingredient of 'correct' /r/, and measurement of absolute formant frequencies must be considered in combination with other acoustic properties.

5. CONCLUSIONS

This speaker exhibited rapid rhotic development over three months towards the end of her fourth year. Impressionistically her outputs become less variable and many more adult-like /r/ tokens are produced. Acoustic analysis, however, indicates that this subject is not only lowering F3, but employing a variety of compensatory strategies in order to produce /r/. For example F2 may be raised and the relative amplitude of F3 increased. The results suggest that, at least for some speakers, absolute formant values must be considered in combination with these additional factors.

6. REFERENCES

[1] Boersma, P., Weenink, D. 2006. Praat: doing phonetics by computer (Version 4.4.12) [Computer program]. http://www.praat.org/.

[2] Dalston, R.M. 1975. Acoustic characteristics of English /w, r, l/ spoken correctly by children and adults. *J. Acoust. Soc. Am.* 57, 462-469.

[3] Delattre, P., Freeman, D.C. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics* 44, 29-68.

[4] Dodd, B., Hua, Z., Crosbie, S., Holm, A., Ozanne, A. 2002. Diagnostic Evaluation of Articulation and Phonology. London: Psychological Corporation.

[5] Espy-Wilson, C.Y., Boyce, S.E., Jackson, M., Narayanan, S., Alwan, A. 2000. Acoustic modelling of American English /r/. *J. Acoust. Soc. Am.* 108, 343-356.

[6] Guenther, F.H., Espy-Wilson, C.Y., Boyce, S.E., Matthies, M.L., Zandipour, M., Perkell, J.S. 1999. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *J. Acoust. Soc. Am.* 105, 2854-2865.

[7] Grunwell, P. 1987. *Clinical Phonology*. London: Croom Helm.

[8] Jones, M.J. 2005. An acoustic study of labiodental /r/ in British English. *Unpublished manuscript, University of Cambridge.*

[9] Kerswill, P. 1996. Children, adolescents, and language change. *Language Variation and Change* 8, 177-202.

[10] Kuehn, D., Tomblin, J.A. 1977. A cineradiographic investigation of children's w/r substitutions. *J. Speech and Hearing Disorders* 42, 462-473.

[11] Menyuk, P., Anderson S. 1969. Children's identification and reproduction of /w/, /r/ and /l/. *Journal of Speech and Hearing Research* 12, 39-52.

[12] Shriberg, L.D., Kent, R.D.1982. *Clinical Phonetics*. New York: John Wiley and Sons.

[13] Stevens, K.N. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

[14] Strange, W.A., Broen, P.A. 1980. Perception and production of approximant consonants by 3-year-olds: a first study. In Yeni-Komshian, G.H., Kavanagh, J.F., Ferguson, C.A. (eds), *Child Phonology*, vol. 2. New York: Academic Press, 117-154.

[15] Vihman, M.M. 1996. *Phonological Development: the origins of language in the child.* Oxford: Blackwell.

[16] Westbury, J.R., Hashi, M., Lindstrom, M.J. 1998. Data collection techniques in the USF speech production and perception laboratory, and their reliability. *Third Workshop on Ultrasound Applications in Linguistic Research*. University of Arizona.