COPiL CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **11**, Article 3: 057–085, 2018 | ISSN 2050-5949

American English Rhotic Rhymes: Phonemic rhotacized schwa or underlying /V+ɹ/?

Connor McCabe
University of Cambridge

1 Introduction

The underlying phonological status of English syllabic consonants has been the subject of considerable research and debate in the literature over the past century. There is strong support for apparent syllabicity in sonorants /l m n/ being a surface realization of underlying /əC/, as discussed in section 2. In rhotic dialects of English, putatively 'syllabic' /ɹ/ occurs in a wider range of positions than its other sonorant counterparts, namely in both stressed and unstressed positions. By contrast, /l m n/ surface syllabically only in unstressed positions. Syllabic /ɹ/ is commonly transcribed as <ɚ>, although <ɹ̩> can also be found.[1] The wider distribution and distinct phonetic status of [ɚ] makes it tempting to suggest that American English (among other rhotic varieties) has developed a phonemic rhotic vowel distinct from consonantal /ɹ/. If, indeed, [ɚ] is able to occur in both stressed and unstressed positions and is consistently realized as a stable rhotacized schwa (or [ɝ] in stressed syllables), why should it not merit analysis as a vocalic phoneme in its own right, distinct from consonantal /ɹ/?

The organization of the remainder of this dissertation is as follows. In section 2, I present a cursory discussion of English syllabic consonants insofar as relevant to discussion of /ɹ/. Section 3 comprises a review of English rhotics, including their articulation and acoustic features and the importance of these factors in the analysis of [ɚ]'s phonological status. Section 4 contains a survey of phonological evidence for and against [ɚ]'s classification as a distinct phoneme, including distributional considerations, possible minimal pairs, and evidence from fluent backward talkers. In section 5, I present and discuss original acoustic evidence against phonemic /ɚ/ drawn from recordings of adult native speakers of American English. This is followed by a discussion and conclusion in section 6.

[1] The distinction between <ɹ̩> and <ɚ> is dismissed as notational variation by some authors, e.g. Akamatsu (2013), paralleling discussion of other approximant/vowel pairs, e.g. /i j/ or /u w/. Among other possible contrasts (cf. Espy-Wilson (2004) on formant structure of /ɹ/ in different positions), this distinction is often based on syllabic vs. syllable-marginal status (Laver 1994: 147-148). As such, for the purposes of syllabic /ɹ/'s candidacy for independent phoneme status, it seems more reasonable to posit /ɚ/ than /ɹ̩/. In addition to being more notationally distinct from <ɹ>, <ɚ> emphasizes the phone's vocalic (and syllabic) qualities

2 Relevant background on syllabic consonants in English

Before proceeding to discussion of putatively syllabic /ɹ/ in American English, it is worth discussing some background information on English syllabic consonants in general. In the most canonical sense, syllabic consonants (or, rather, consonants with the potential to obtain syllabicity in certain phonetic contexts and syllable positions) comprise the sonorants /ɹ l m n/ (Akamatsu 2013). Some authors also note phonetically "syllabic" cases of, e.g., /f s/ (Bonilla 2003), but these cases, more restricted in distribution than syllabic sonorants, are comfortably classed as surface phenomena and are therefore of only marginal relevance to the topic at hand.

The surfacing of syllabic consonants, although frequent in speech of all registers with the exception of particularly exaggerated citation forms, is generally conceived of as being restricted to unstressed syllables. These are most commonly word-final, but this is not a requirement (e.g. 'Italy' ['ɪtl̩i]. The fact alone that these phones do not surface uniformly but rather alternate with schwa plus consonant sequences suggests that their syllabicity is not present in the underlying representation. The latter analysis (i.e. /l̩ ɹ̩ m̩ n̩/) is unpalatable in terms of systemic economy, especially given that the difference between [l̩ ɹ̩ m̩ n̩] and [əl əɹ əm ən] is rarely lexically contrastive (Wells 1995). In those pairs which do allow a possible contrast in meaning to be ostensibly realized via the distinction between [C] and [C̩], this is never the only recourse for lexical distinction. Indeed, although not in free variation, [əC] and [C̩] act as covariants of one another (Akamatsu 2013). The schwa-containing string (e.g. 'lightening' as 'ligh[tə]ning' rather than 'light[n̩]ing' in contrast with 'lightning' 'light.[n]ing') is generally perceived as hyperarticulatory and/or characteristic of child or non-native speech, but is nevertheless viable.

The alternation between [əC] and [C̩] presents the possibility of two broad analysis types. Either surface [C̩] is underlying /əC/, produced in interaction with vowel syncope, discrete phonological rules, rate of speech, etc. or it is the realization of normal (non-syllabic) /C/ resyllabified in accordance with English's phonotactics (Wells 1995). Wells notes that the latter approach, supported by Jones (1967), struggles to capture the possible difference in syllabifications in pairs such as SSBE 'Hungary' ['hʌn.gɻ̩.i] and 'hungry' ['hʌn.gɹi] to both of which a Jones-style approach would presumably ascribe the phonological representation /'hʌng.ɹi/. As such, the family of analyses which derive surface [C̩] from underlying /əC/ seem most able to provide a range of empirical coverage while maintaining valuable systemic economy.

It is worth emphasizing that the above is neither an entirely uncontroversial conclusion nor an exhaustive survey of the literature on the topic. What is relevant is that for the purposes of the present discussion of [ɝ ɚ] and its abstract phonological counterpart(s), I favor /əC/ as the stronger candidate for the underlying representation of English's syllabic consonants in general.

As noted above, syllabic consonants are canonically cited as occurring in unstressed position, particularly by authors studying non-rhotic varieties of English (namely RP/SSBE, cf. Wells (1995), Bonilla (2003), Akamatsu (2013), de Jesus Arboleda & Monroy (2015) inter alia). This puts American English /ɹ/ in an interesting theoretical position, as its putatively "syllabic" variant frequently occurs in stressed

position in both mono- and polysyllabic words, in which positions it is frequently transcribed as <ɜˠ> in contrast with unstressed <əˠ>.[2] If the restriction of surface syllabic consonants to unstressed position (being the output of phonological rules and phonetic processes acting on a schwa+consonant string) is taken as definitional, then it is possible to logically eliminate <ɻ̩> as a possible representation of the vocoid encountered in 'her' and 'servant'. The question is then not whether phonemic /ɻ̩/ exists in English, since this has been eliminated along with its other sonorant counterparts. Rather, it is a matter of what underlying representation to ascribe to the rhotacized vowel ([ɜˠ] or [əˠ], as the case may be) attested in these rhotic rhymes.

Wells (2011) speculates that unstressed [əˠ] may be reconciled with the underlying /əC/ approach while stressed [ɜˠ] should be analysed as a phoneme in its own right corresponding to the British /ɜ/ (i.e. the NURSE vowel). In such an approach, it would seem that <ɻ̩> and <əˠ> should be more or less notational variations in unstressed positions for any level of transcription above the phonemic, while <ɜˠ> alone would be appropriate for stressed positions. It is my contention based on both phonological and phonetic evidence (discussed in sections 4 and 5, respectively) that both stressed and unstressed cases of "syllabic" /ɹ/ can be reconciled with underlying /V+ɹ/, regardless of whether [ɜ] is taken to be phonemic.

To summarize thus far, to posit underlying syllabic consonants for English is both uneconomical and unnecessary, as supported by research primarily on non-rhotic varieties of British English focusing on /l m n/. It is logical to extend this prohibition to cases of putatively syllabic /ɹ/. However, the realization of syllabic /ɹ/ as a (rhotacized) vocoid, unique among other syllabic consonants, and the availability of such a realization in both stressed and unstressed position leaves open the possibility of attributing a phonemic rhotacized schwa —as distinct from either /ɻ̩/ or /əɹ/ —to rhotic varieties such as American English. Below, I present a discussion of the various factors which must be taken into account when studying the articulation, acoustics, and phonology of rhotics followed by evidence for the presence of an underlying (non-rhotic) vowel in rhotic rhymes of all positions.

3 Background on rhotics

3.1 Articulatory variation

From an articulatory perspective, /ɹ/ is unique by virtue of its characteristic formant structure being able to be achieved by a variety of articulatory strategies. Details of the acoustic characteristics of /ɹ/ are discussed below in section 3.2. Although these

[2] Official IPA convention now includes [ɜ] as an open-mid central vowel distinct in height from central [ə] (International Phonetic Association 1993), although it previously recognized [ə] as valid only when it was necessary to denote a central vowel distinct from schwa (International Phonetic Association 1989). Earlier usage norms have persisted to some extent in the preference of authors such as Kuecker, Lockenvitz & Muller (2013) to distinguish the two (or the rhotacized counterpart of each) on the basis of stress. Gimson (1962: 116), for example, suggests that [ə] may be an unstressed allophone of phonemic /ɜː/. In this paper, I distinguish between [ɜ] and [ə] on the conventional basis of stressed vs. unstressed position, setting aside both the issues of the phonemic/allophonic status of either and of the height difference between the two.

strategies differ from one another considerably, the acoustic/auditory result of each is virtually indistinguishable from the others. However, it is logical to assume that the different tongue shapes and degree of pharyngeal constriction exhibited by different articulation strategies, while equivalent in isolation, behave differently from one another in interaction with syllabic position and phonetic context as a result of physiological constraints on articulation. This line of phonotactic consideration is of particular importance in explaining both the extent to which underlying /ə/ and following /ɹ/ merge with one another and the range of inter- and intraspeaker variation in the discreteness of the /əɹ/ → [ɚ] realization.

In their 1968 cineradiographic study of English /ɹ/ conducted using X-ray technology, Delattre & Freeman identified eight distinct strategies for /ɹ/-production across British and American English. I have extracted and presented the authors‘ original illustration of this eight-way typology in Figure 1. Types 1 and 8 are identified as characteristic of British English (Type 1 being ostensibly non-rhotic, and Type 8 lacking a pharyngeal constriction), while Types 2-7 are all exhibited by American English speakers (Type 2 being attributed to speakers of non-rhotic American varieties). Although Types 2-5 are cited as postvocalic strategies (drawing no distinction between syllabic and postvocalic /ɹ/ while acknowledging the surfacing of /əɹ/ as [ɚ] in, e.g., ‘curb’) and Types 6 and 7 as prevocalic ones, the authors observe that there is wide inter- and intraspeaker variation. Thus, alongside those speakers who consistently use a single articulation type and those who consistently contrast a single prevocalic type and a single postvocalic type are those who use a typically pre- or postvocalic strategy in all positions and those who exhibit variation among multiple articulation types.

Delattre & Freeman‘s (1968) work represents a substantial departure from the simpler distinction traditionally drawn between “bunched" and “retroflex" /ɹ/. It is worth noting that the authors do not abandon the bunched/retroflex distinction (see, e.g., Types 6 and 7) so much as acknowledge that evidence from imaging more comprehensive than that available to earlier researchers suggests the situation of /ɹ/ articulation to be more complex than a simple binary distinction.

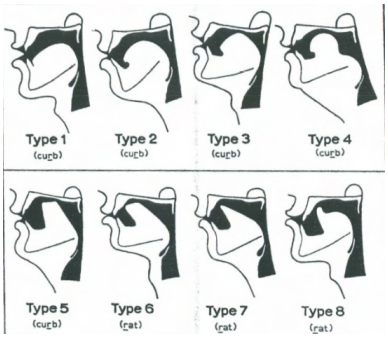

Figure 1 Types of ɹ articulation, taken from Delattre & Freeman (1968)

Variation in /ɹ/-articulation strategy has been more recently investigated for both American and British varieties of English, frequently with reference to Delattre & Freeman‘s (1968) descriptive classifications. Using the binary retroflexion/bunching

distinction paired with possible intraspeaker variation in strategy between onsets and codas, Scobbie, Lawson, Nakai, Cleland & Stuart-Smith (2015) present a broad four-way typology of /ɹ/ production, primarily focusing on approximant-utilizing Scottish English varieties. With specific reference to American English, the authors note that speakers who consistently use a single strategy overwhelmingly use a bunched articulation type, while those speakers who use retroflexion tend to exhibit positional variation in strategy. This is consistent with Mielke, Baker & Archangeli's (2010) division of (rhotic) American English speakers into three categories: those who consistently use a single articulation strategy, those who exhibit free variation between strategies, and those who exhibit constrained variation between strategies. The latter type of variation "can have very complex conditioning environments'specific to individual speakers (and their phonetic needs)" (Mielke et al. 2010). For example, speakers with constrained strategy variation appear to favor retroflexion prevocalically.

In addition to being sensitive to factors such as phonetic context and rate of speech, different /ɹ/ strategies logically interact with connected speech processes in different ways. Mielke et al. (2010) investigate the interaction between different /ɹ/ articulation strategies and /s/-retraction in initial /sCɹ-/ as in 'street', 'strand', 'spray', etc., finding that "different /ɹ/ allophones cause different degrees of coarticulation on [s]". Although all subjects produced variants of the bunched /ɹ/ articulation type, preventing comparison of the two broad categories of articulation strategy, degree of /s/-retraction (i.e. coarticulation between the /ɹ/ and /s/ targets) was observed to be more pronounced for speakers whose /ɹ/ target was least antagonistic to their /s/ target. The gradience of effect on coarticulation within the "bunched" /ɹ/ category is consistent with Delattre & Freeman (1968) enrichment of the /ɹ/ classification scheme.

Of more direct relevance to the present issue of syllabic /ɹ/'s phonological status is Lawson, Scobbie & Stuart-Smith's (2013) finding that bunched /ɹ/ articulation "exerts a stronger global coarticulatory force over preceding vowel tongue configurations than tongue-front raised [i.e. apical] /r/ does." If this is so, then it strengthens the case for analysing [ɚ] as a surface phenomenon resulting from coarticulation between /ɹ/ and /ə/, the extent of which coarticulation is at least partially a function of /ɹ/ articulation strategy. Cases of particularly strong (i.e. acoustically and perceptually stable) "syllabic" productions, as exhibited by several of the subjects discussed in section 5, could feasibly be the result of certain articulation strategies having a more pronounced global coarticulatory effect than others. Similarly, variability in degree of /əɹ/ → [ɚ] discreteness (or, conversely, evident degree of distinct [ə] articulation) may be the result of intraspeaker variability in /ɹ/ articulation strategy subject to complex environmental conditioning.

It is also worth noting Lawson et al.'s (2013) observation from ultrasound evidence that the most 'extreme' rhotic articulation ("the maximum of the /r/ gesture") is often reached after voicing has stopped. Although the data discussed in section 5 is purely acoustic, and therefore blind to the direct articulatory data which Lawson et al. were able to observe, it is important to bear in mind that speakers may articulate a

rhotic target which, although distinct from the (potentially rhotacized) vowel, does not have an acoustically salient peak.

To summarize, speakers achieve perceptual /ɹ/ through a variety of articulatory strategies which are virtually indistinguishable from one another in terms of formant structure. Some speakers consistently use one articulation type, while others exhibit either free or constrained variation between strategies. Crucially, although largely indistinct from one another when observed in isolation, different /ɹ/ articulation strategies exhibit different coarticulatory effects in phonetic context. Observations of certain articulation strategies exerting strong coarticulatory effects on neighboring segments (including schwa) may be useful in justifying surface [ɚ]'s analysis as /əɹ/, especially when analyzing speakers who produce this phone without any trace (acoustic or perceptual) of distinct vowel and rhotic segments.

3.2 Acoustic characteristics

The acoustic feature most commonly cited as characteristic of /ɹ/ is a lowering of the third formant (F3) relative to adjacent phones (Laver 1994). This generalization is tacitly relied upon in much of the available literature on rhotics (e.g. Lawson et al. (2013), and F3 decline has been used as a metric in studies involving rhotic duration (Lockenvitz, Kuecker & Ball 2015, Kuecker et al. 2013). There is, however, evidence to suggest that a decline in raw F3 value is not itself the primary cue of a perceptual /ɹ/ (distinct from other approximants such as /w/). Specifically, it seems that the attainment of a minimal distance between the second and third formants plays an important role in cueing perceptual rhoticity. Below, I discuss evidence for using the relationship between F3 and F2 (herein referred to as min.F3-F2) as the primary acoustic marker of rhoticity. This approach to rhotic definition is of interest to the present study as it provides a more nuanced view of rhoticity in syllabic /ɹ/ contexts, allowing for distinct vocalic and rhotic targets —which may be obscured by the broader F3-decline approach —to be teased apart in spectrograms.

The close proximity between F3 and F2 is not a recent observation in the available literature on rhotics (Delattre & Freeman 1968, Dalston 1975). In their primarily descriptive approach to the classification of rhotic articulation types, Delattre & Freeman (1968) observe closeness between F3 and F2 for all /ɹ/ articulation strategies (recall the 8-way classification presented above) with the exception of two; "Types 1 and 8 ... show a wider separation between the second and third formant." This is interesting in light of the fact that Delattre & Freeman's /ɹ/ articulation Types 1 and 8 are attested only in non-rhotic varieties of British English, consistent with the authors' characterization of F3-F2 narrowing being the result of the creation of a third resonance cavity in the vocal tract via what they term "palato-velar constriction". In updated terminology, this "palato-velar constriction" is equivalent to the (posterior) tongue root gesture which accompanies primary (anterior) tongue bunching or retroflexion ((Gick, Bacsfalvi, Berhardt, Oh, Stolar & Wilson 2007, Klein, Byun, Davidson & Grigos 2013). This pharyngeal constriction, along with correlated F3-F2 narrowing, is identified in (American) articulation Types 2-7 and is notably absent from Type 8 (used by British speakers in syllable onsets). The relationship

between complex articulation and the attainment of [ɹ]'s characteristic acoustics (including the narrowing of F3-F2) has also been related to the later mastery of min.F3-F2 by comparison to simple F3 decline by children (McGowan, Nitrouen & Manning 2004). Furthermore, acoustic information may be utilized in different ways by different speakers. This applies both to different raw acoustic cues (McGowan et al. 2004) and to, e.g., temporal vs. (more strictly) spectral information (Dalston 1975).

In their study of possible change in weighting of acoustic cues in perception of approximants in British English speakers, na Dalcher, Knight & Jones (2008) stress the importance of F2 in the attainment of target /ɹ/. The authors highlight the contrast between /ɹ/ and /w/, noting that the former characteristically exhibits low F3 and high F2, while the latter exhibits essentially the opposite (high F3 and low F2). The evidence presented by their perceptual tests run on both American and British English speakers suggests that low (i.e. canonically [ɹ]-like) F3, although perhaps the most immediately salient feature of [ɹ] in spectral analysis, does not reliably predict a listener identification of /ɹ/ if F2 is uncharacteristically low. Instead, the authors conclude that both F3 and F2 height play a role in cueing perceptual rhoticity. The relative importance of F2 height is implicated in the identification by naive American listeners of labiodental [ʋ] (which has [ɹ]-like F2 paired with relatively high F3) as a distorted rhotic rather than as a variant of /w/ in a forced identification task. The authors note the potentially widening acoustic and articulatory gap between the /ɹ/ variants presented by British and American English.

In a similar vein, Klein, Grigos, Byun & Davidson (2012) presented naive listeners with gradient series of natural and manipulated auditory stimuli in order to observe acoustic correlates for judgements of /ɹ/ vs. /w/. Although raw F3 height was found to be the best predictor of positive rhotic identification in a statistical sense, [ɹ]-like F2 in isolation was frequently sufficient to cue identification of a distorted rhotic. The contrast between onset and coda /ɹ/ is also commented on. The authors note that although listeners are willing to accept a wider range of F3-F2 distances in onset /ɹ/, interpersonal judgements are more consistent for codas (including both stressed and unstressed syllabic contexts). Both of these effects are suggested to be related to the durational asymmetry between the two positions which is paralleled in other sonorants (Klein et al. 2012).

In short, evidence from perceptual and acoustic studies suggests decline in raw F3 value is not necessarily the most accurate indicator of attainment of a rhotic target, especially in light of data involving [ɹ]-like F2 height cueing perceptual /ɹ/ judgements even in the absence of a canonical F3 value. Close proximity (and in many cases an essential merger) between the second and third formants, salient enough in relatively early descriptive observations of /ɹ/ articulation and acoustics to merit comment (Delattre & Freeman 1968, Dalston 1975), both demonstrably cues a perceptual rhotic (to the exclusion of, e.g., /w/) and is readily identifiable in spectrograms. Lack of significant F3-F2 proximity, according to Delattre & Freeman (1968), only occurs in the absence of pharyngeal constriction. Since such constriction appears to be standard in rhotic varieties of American English, and since there is evidence suggesting that F2 height and F3-F2 distance play an important role in

cueing perceptual rhoticity, it is reasonable to include min.F3-F2 as one of /ɹ/'s identifying acoustic characteristics.

4 Phonological evidence for [ɝ ɚ] as /V+ɹ/

Before proceeding to the main discussion of acoustic evidence for a schwa-inclusive underlying representation of syllabic /ɹ/, it is worth briefly discussing abstract phonological evidence for such an analysis. In the sections below, I discuss two points of interest: the rarity of either [ɝ] or [ɚ] behaving contrastively with [ɹ] and the apparent compositionality of rhotacized schwa for backward talkers. Note that for brevity's sake, unless otherwise specified I will use 'rhotacized schwa' as an umbrella term encompassing both stressed [ɝ] and unstressed [ɚ].

4.1 Minimal pairs contrasting rhotacized schwa and /ɹ/

If it is the case that [ɝ], [ɚ] or both are phonemes in American English, then it is reasonable to expect them to create minimal pairs with the rhotic consonant [ɹ] (or, rather, /ɹ/ since its phonemic status is not in question). It is evident that rhotacized schwa can minimally contrast with other non-rhotacized vowels, as in 'gully' ['gʌli] vs. 'girly' ['gɝli], 'Ellie' ['ɛli] vs. 'early' ['ɝli], or 'fury' ['fjɝ(ɹ)i] vs. 'Führer' ['fjɝ(ɹ)ɚ] (the identity between the vowels in the initial syllables of the latter pair being subject to variation, as some speakers may have [ʊ] in one or both words). This, however, is not particularly revealing as these contrasts can in principle be subsumed under either a /ɜɹ əɹ/ approach or a /ɝ ɚ/ approach. The syllabic /ɹ/ elicitation task discussed in section 5, although primarily acoustic, also provided a few points of more abstract interest, particularly with regards to minimal contrasts. Below, I discuss phonological considerations for both stressed and unstressed rhotacized schwa.

Stressed [ɝ] has a definitionally limited distribution, being restricted to the nucleus and coda of stressed syllables. This makes the creation of true minimal pairs in opposition with [ɹ] somewhat challenging. The ideal minimal pair would contain the sequence [ɝV] vs. [ɹV]. Close matches for this template include 'stirrup' ['stɝ(ɹ)ʌp] vs. 'strut' ['stɹʌt]. The production of such a contrast is certainly conceivable, and other more tenuous pairs (such as 'early' ['ɝᵊli] and 'really' ['ɹɪli] in the sufficiently reduced speech of a hypothetical American English speaker) come to mind. There is, however, a broader issue of syllabification which comes into play with near-contrasts of the 'stirrup'-'strut' sort which I have highlighted by the inclusion of an optional [ɹ] in my transcription of the former. Leaving aside orthographic bias, it is difficult to say whether the word medial [ɝ] of, e.g., 'stirrup' is truly restricted to the nucleus of the first syllable and is followed by a vowel-initial second syllable or is ambisyllabically involved in the onset of the (rhotic-initial) second syllable. It is intuitively appealing to think that the difference between ['stɝ.ɹʌp] and ['stɝ.ʌp] corresponds to the contrast between 'stirrup' and 'stir up'. This unfortunately does little to weigh in on the underlying representation of surface [ɝ], and only further complicates the identification of potential minimal pairs. I return to the issue of

rhotic and rhotacized vowel ambisyllabicity and its implications for the underlying representation of syllabic /ɹ/ below. For present purposes of discussion, simply note that the identification of minimal pairs which rely on a contrast between [ɹ] and [ɝ] is no easy task.

The possibility of minimal pairs contrasting [ɹ] with unstressed [ɚ] is raised by lexical items containing either phone immediately following a full vowel. Lexical items containing /Vɹ/ are easy to come by, as in 'bar', 'hair', and 'fear'. Cases of [ɚ] preceded by a vowel rather than a consonant are more difficult to find. Individual speakers' realizations of 'drawer' and 'horror' as ['dɹɔ.ɚ] and ['hɔ.ɚ] come to mind, and it would be interesting to see if either is able to be used contrastively. Potential contrasts between, e.g., a ['dɹɔ.ɚ] 'person who draws' and a ['dɹɔɹ] 'part of a desk or dresser?, or between the non-homographic 'horror' ['hɔ.ɚ] and 'whore' ['hɔɹ] come to mind. However, although potential contrasts of this sort appear to present minimal pairs between [ɚ] and [ɹ], the evidence which they would provide for phonemic /ɚ/ is at best weak. Consider in particular the fact that while a phonemic /ɚ/ approach is unable to economically explain the significant number of speakers who can be predicted to not produce a contrast along these lines, a /əɹ/ → [ɚ] approach is more easily reconciled with speakers of both the non-contrastive and (apparently) contrastive [ɚ] type. As this is closely related to acoustic evidence presented by cases of 'double' syllabic /ɹ/, see section 5 and discussion in section 6 for further evaluation.

In brief, minimal pairs contrasting either [ɝ] or [ɚ] with [ɹ] are difficult to identify, and, once identified, the evidence which they provide in favor of phonemic rhotacized schwa is not particularly strong. The scant evidence for true contrastiveness between rhotacized schwa and /ɹ/ is reminiscent of Akamatsu's (2013) identification of covariation between [əC] and [C̩] ([ɚ] in this case). Rather, an analysis of underlying /Vɹ/ remains the more economical approach.

4.2 Evidence from fluent backward talkers

A final piece of abstract phonological evidence of relevance to the underlying representation assigned to surface rhotacized schwa comes from the productions of fluent backward talkers. In their 1985 paper "The Phonological and Metaphonological Representations of Speech: Evidence from Fluent Backward Talkers", Cowan et al. describe the speech of 13 adult native speakers of American English who possess the ability to produce backwards strings of speech "[including] long words quickly and without rehearsal." The speech of such backward talkers is of particular interest for discussion of the underlying phonological forms of speech due to the overwhelming tendency to manipulate discrete phonemes, largely independent of acoustics (the affricate /t͡ʃ/, for example, being reversed as [tʃ] rather than [ʃt]) and orthography. This is not unilaterally true of Cowan et al.'s subjects —one subject manipulates the ordering of syllables rather than phonemes in his reversals, while heavy orthographic bias is reported by the authors to be common among backward talkers.

With specific regards to the underlying representation of surface syllabic /ɹ/, two predictions are possible for the productions of fluent backward talkers. If it is the case that rhotacized schwa (be it stressed [ɝ], unstressed [ɚ], or both) is a phonological primitive, then it should be preserved in phonemic reversals such as those produced by Cowan et al.'s (1985) subjects. Lexical items such as 'turn' [tɝn] and 'ladder' [lærɚ] (both elicited in Cowan et al. would have predicted reversals approximating [nɝt] and [ɚdæl]. On the other hand, if rhotacized schwa is a surface phenomenon, this phonemic compositionality should be evident in reversal tasks, yielding 'turn' [nɹɜt] and 'ladder' [ɹədæl]. Results favor an analysis corresponding to the latter prediction. Across stressed and unstressed tokens, rhotacized schwa is preserved in reversal in only 7 of 44 tokens; 6 of these 7 tokens were produced by a single subject. The latter subject is also the only one of the 13 individuals studied to frequently (though not exclusively) reverse syllable order rather than phoneme order. As such, it is not only unsurprising that rhotacized schwa should be preserved in this subject's reversals, but this phenomenon also sheds no light on the subsyllabic composition of the speaker's rhotacized schwas. The remaining majority of subjects and individual tokens consistently exhibit compositional reversals of the [ɹɜ]/[ɹə]-type. This points to a schwa-inclusive underlying representation of surface rhotacized schwa. Furthermore, the fact that subjects' compositional reversals may violate English phonotactics (e.g. initial [nɹ-] in reversals of 'turn') undermines potential arguments of [ɚ]'s reversal as [ɹə] being a matter of phonotactic consideration.

To conclude this section, phonological evidence for the underlying representation of surface rhotacized schwa can be difficult to source. Identification of potential minimal pairs is hampered by a matter of lexical availability, limiting (though not eliminating) the usefulness of a classic test for phoneme-status (Jones 1967: 15, inter alia). However, limited data suggesting potential contrasts distinguished by rhotacized schwa vs. [ɹ] (cf. the 'horror'-'whore' distinction extrapolated from certain subjects' production of the former in the acoustic study discussed below) is economically explicable within a schwa-inclusive approach to its underlying status, while the same cannot be said of phonemic rhotacized schwa approaches to the more frequently attested lack of such contrasts. Evidence from Cowan et al.'s (1985) study of fluent backward talkers is even stronger, the reversal of both stressed and unstressed rhotacized schwa as [ɹV] rather than as [ɚ ɝ] pointing to a compositional (i.e. biphonemic) underlying representation. Below, I present the structure and results of an acoustic study undertaken with the aim of revealing the phonological representation underlying surface rhotacized schwa.

5 Acoustic phonetic evidence for an underlying /ɜ ə/ in [ɝ ɚ]

While the phonological evidence discussed above is of somewhat mixed strength, even relatively strong evidence (albeit limited in quantity) for underlying /Vɹ/ in surface rhotacized schwa such as the production of fluent backward talkers can, with considerable cost to theoretical economy, be reconciled with a phonemic /ɝ ɚ/ approach. With this in mind, in the present section, I present and discuss the results

of an acoustic study of syllabic /ɹ/ undertaken with the aim of identifying evidence of distinct vowel and rhotic targets in the formant structure of perceptually stable [ɝ ɚ].

Previous preliminary studies which have investigated the formant structure of syllabic [ɝ ɚ], namely Lockenvitz et al. (2015) and Kuecker et al. (2013), have commented on the asymmetry exhibited between stressed and unstressed syllabic /ɹ/. Specifically, they note that duration of rhoticity is frequently coterminous with the rest of the vocoid sequence in stressed contexts (i.e. [ɝ]), while onset of rhoticity is slightly delayed in unstressed contexts (i.e. [ɚ]). If this is the case, it hints in favor of Wells' (2011) suggestion of phonemic /ɝ/ (i.e. the NURSE vowel) available in stressed positions contrasting with phonologically/phonetically derived [ɚ] in unstressed positions. In other words, 'bird' should have an underlying representation of /bɝd/ while 'father' should have /'fɑðəɹ/.

Crucially, however, the above conclusion relies on durational measurements taken using the decline in raw F3 height as the primary marker of rhoticity —"the rhotic portion [being] deemed to start (word finally) at the point where F3 [starts] to lower, and [being] deemed to finish (word initially) at the point where F3 [reaches] its highest point before becoming level" Lockenvitz et al. (2015). Would such generalizations about duration of rhoticity remain valid if, based on the evidence summarized in section 3, a more restrictive definition of rhoticity based on the attainment of close proximity between F3 and F2 (min.F3-F2) were to be adopted? If so, then there might indeed be a case for a rhotic vowel phoneme in American English. If, as I hypothesized at the outset of the present study, a min.F3-F2-based approach to rhoticity shows a consistent asymmetry between /Vɹ/ duration and rhotic duration (that is, if there is a consistent delay in onset of rhoticity), then the acoustic evidence would suggest underlying /Vɹ/ for syllabic /ɹ/ in all positions (both stressed and unstressed).

Below, I present an acoustic study undertaken on the syllabic /ɹ/ productions of 12 native adult speakers of American English. The study was designed with the aim of identifying acoustic evidence of distinct vowel and rhotic targets in perceptually stable syllabic /ɹ/ tokens, if such distinction were indeed found to occur.

5.1 Structure of the study

In order to study the formant structure of syllabic /ɹ/, 12 adult native speakers of rhotic American English were recruited from the American undergraduate and graduate student bodies of the University of Cambridge. Subjects came from eight states (Connecticut, New Jersey, New York, Texas, North Carolina, California, Ohio, and Maryland), and none had been living in the UK for longer than five years. No subject exhibited a particularly strong (or readily identifiable) regional accent. Before participation, subjects were asked if they had received any speech therapy during childhood. Of the 12 subjects, three had received speech therapy (subjects 01, 10 and 12); only one of these three (01) had specifically been treated for /ɹ/ production. 01's /ɹ/ production at the time of recording was indistinguishable from that of the other subjects.

A word list comprising 62 items selected to elicit syllabic /ɹ/ (both stressed and unstressed), singleton onset /ɹ/, postvocalic /ɹ/, and syllabic /l/ was prepared. At a maximum, 34 of these 62 tokens elicited syllabic /ɹ/. At a minimum, syllabic /ɹ/ was present in 30 of the 62. This variation is the result of interpersonal variability in vowel syncope ('favorite' as ['feɪvɚ(ɹ)ɪt] vs. ['feɪv.ɹɪt]), vowel reduction to schwa ('horror' as ['hɔɹɚ] vs. ['hɔɹɔɹ]), and lexical representation ('lure' as [lɝ] vs. [luɹ] vs. [lɔɹ]). Judgement of syllabic /ɹ/ status was carried out perceptually for potentially ambiguous tokens. The inclusion of tokens of prevocalic /ɹ/ allowed for comparison of relative duration of rhotic segments, while inclusion of postvocalic /ɹ/ allowed for comparison of rhoticity relative to the rest of the /Vɹ/ sequence in cases of uncontested /V+ɹ/ vs. in syllabic contexts surfacing as rhotacized schwa. Finally, the inclusion of a small number of syllabic /l/ tokens (7/62) allowed for comparison of the degree of [ə] production in sequences of syllabic /ɹ/ vs. syllabic /l/.

Subjects were recorded using a Zoom H4n audio recorder, and the results were analyzed using Praat (Boersma & D.Weenink 2016). There were three separate production tasks presented: the list of 62 lexical items read in isolation (word list), the same 62 items each embedded in a short sentence (sentential context), and a subset of the larger list of 62 items along with a small number of incidental tokens of syllabic /ɹ/ in the context of a short passage specifically written for this task (extended context). The materials presented to the subjects are included below in Appendix A.

After recording was completed, the resulting audio files were analyzed and annotated using Praat (Boersma & D.Weenink 2016). For /Vɹ/ and syllabic /ɹ/, the durations of both the full vowel+rhotic/rhotic vowel sequence and the rhotic itself were measured. The difficult task of identifying onset of rhoticity was done using a metric of min.F3-F2 attainment (as opposed to decline in raw F3 value). In other words, measurement of postvocalic rhotic duration began at the point at which a stable, minimal proximity between F3 and F2 was attained. This minimal relationship ideally occurred in a steady state, although cases of a stable relationship during transition were accepted in the absence of an eventual steady state. In many cases, min.F3-F2 was observable as the second and third formants merging for all practical purposes (see Figure 2 below). As discussed above, there is strong evidence for the importance of min.F3-F2 in defining a canonical /ɹ/, especially in postvocalic position (McGowan et al. 2004). Identification of rhotic onset using min.F3-F2, while satisfyingly less permissive than a raw F3-decline approach (see Figure 2 below), is still a substantially challenging and often uncooperative task.

In Figures 2 and 3 below, I provide an illustration of ideal cases of min.F3-F2-defined rhotic duration for tokens of stressed 'her' (putatively syllabic /ɹ/) and 'hear' (non-syllabic). Perceptual judgements were relied on in spectrally challenging cases, and in some contexts were seen to reinforce the more restrictively defined min.F3-F2 rhoticity. Perception of vowel rhoticity before the attainment of min.F3-F2 was common, contributing to perceptual stability of rhoticity even in the absence of acoustic stability (as in Figure 2). This was considered analogous to, e.g., heavy vowel nasalization before the achievement of a target nasal consonant (Hasegawa-Johnson, Chen, Cole, Borys, Kim, Cohen, Zhang, Choi, Kim, Yoon & Chavarria 2005,

Chen, Slifka & Stevens 2007). In this sense, 'her' /hɜɹ/ → [hɜ˞(ɹ)] parallels 'tan' /tæn/ → [tã(n)]. Vowel nasalization without the articulation of an accompanying nasal consonant is incidentally attested in the speech of at least one subject (subject 09), the same individual who exhibited the strongest vowel rhoticization in all contexts. Figure 4 is a case of spectrally ambiguous /ɹ/, which has been perceptually verified to be in the onset of the following syllable rather than in the nucleus or coda of the previous syllable (issues of ambisyllabicity and their implications are returned to below). However, since syllabic /ɹ/ is frequently perceptually stable, aural judgements were often of limited use.

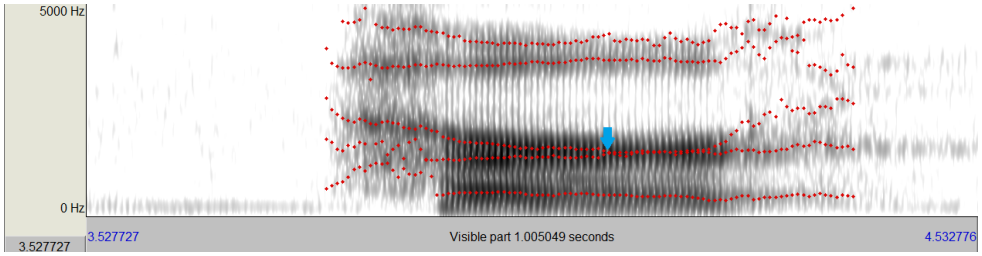

Figure 2 'her' with beginning of stable min.F3-F2 marked, note that in a raw F3-decline approach in the style of Lockenvitz et al. (2015), rhoticity in this token would be considered to begin at the decline in raw F3 immediately following [h]

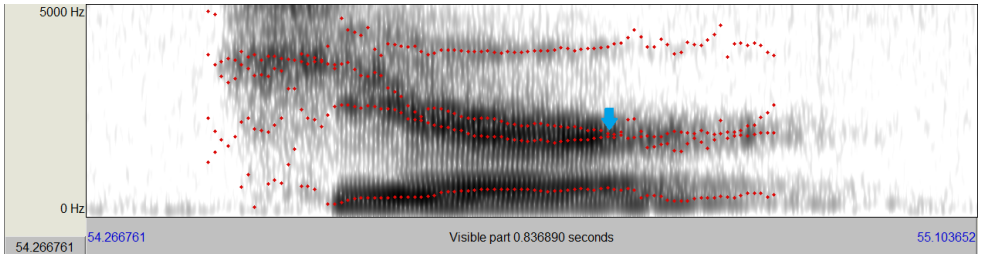

Figure 3 'hear' with beginning of min.F3-F2 marked, note the merger between F2 and F3 paralleling that of Figure 2

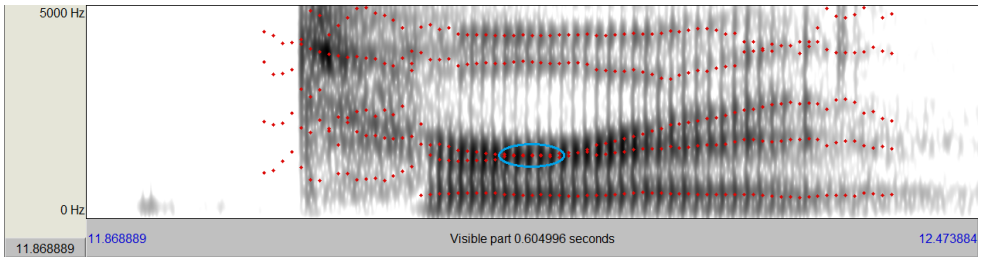

Figure 4 'turret' with ambiguous /ɹ/ perceptually verified to be in the onset of the second syllable

Unsurprisingly, cases of /Vɹ/ and rhotic duration being coterminous in syllabic tokens were common, even predominating in some speakers. The importance of

identification of immediate postconsonantal rhoticity in all speakers (to widely varying extents) in a variety of consonantal contexts using the min.F3-F2 metric should not be understated. This demonstrates the method's ability to distinguish between postconsonantal formant transitions and a significant lag in onset of rhoticity (which I suggest to be evidence of an underlying vowel). In Figures 5 and 6 below, I provide spectrograms distinguishing inevitable formant transition from a lag in attainment of rhoticity.

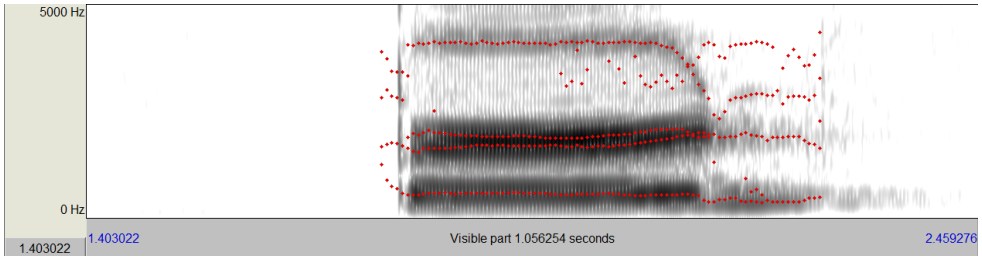

Figure 5 'bird' with stable postconsonantal min.F3-F2 rhoticity alongside minor formant transitions out of preceding [b]

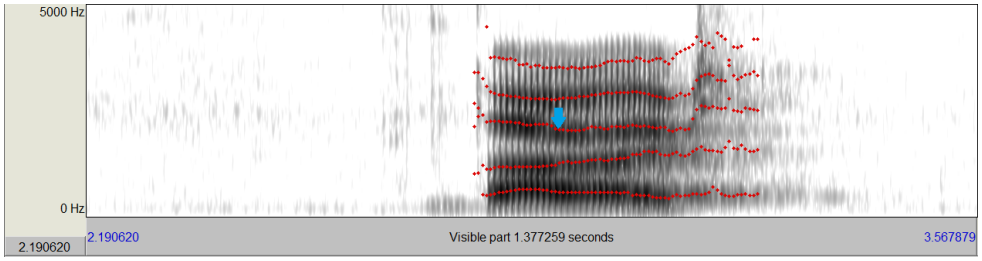

Figure 6 'bird' with significant lag in onset of rhoticity, beginning of min.F3-F2 (distinct from the beginning of a stable but more distant relationship between the two) marked. Note that the attainment of min.F3-F2 is contributed to by both a decline in F3 and a rise in F2 relative to the previous distance between the two formants, rather than by one or the other.

Syllabic /l/ was measured analogously, with both /əl/ and /l/ durations taken and compared. Prevocalic /ɹ/ was measured only for its own duration. For all token-types except the prevocalic, liquid duration was subtracted from the full vowel+liquid duration. The difference between the two was also converted into a percentage, allowing for comparison of different sequences? relative percentage of rhoticity.

The results of this process of recording and analysis are presented below. This is followed by a broader discussion in section 6.

5.2 Results

The results of the acoustic study show a large amount of both inter- and intraspeaker variation in duration of postvocalic rhotics, attested in both syllabic /ɹ/ and in sequences of uncontested /Vɹ/ (e.g. 'bar'). On an intraspeaker basis, variation

is evident as a function of phonetic context (e.g. 'her' vs. 'fur') and production context (word list, sentential context and extended context). I hypothesize that wide interspeaker variation can primarily be accounted for as a function of /ɹ/ articulation strategy interacting with phonetic context. Crucial to my later argumentation is the fact that rhotic duration coterminous with the vocoid sequence as a whole (e.g. everything following [h] in 'her') is the exception rather than the rule. Only 2/12 subjects produce this type of full rhoticity on a consistent basis, and while others do produce full rhoticity to varying extents, this is produced alongside a more common delay in onset of rhoticity. Below, I have separated discussion of the data collected into results from production of uncontroversial postvocalic /ɹ/, production of stressed syllabic /ɹ/, production of unstressed syllabic /ɹ/, production of syllabic /l/, and production of 'double' syllabic /ɹ/. It is worth emphasizing that the measurements collected suggest that stressed and unstressed syllabic /ɹ/ pattern more together than apart —the evidence does not support a rhotic NURSE phoneme /ɝ/ contrasting with an unstressed /əɹ/ which surfaces as [ɚ]. Rather, the relatively consistent delay in onset of rhoticity (even in light of frequent variation in relative size of the prerhotic vowel) is consistent with an underlying sequence of vowel plus /ɹ/ in both stressed and unstressed contexts.

5.2.1 Uncontroversial /V+ɹ/

In order to draw meaningful conclusions about the status of syllabic /ɹ/, it is important to first understand how /ɹ/ behaves in uncontroversial postvocalic position, that is, when following a full vowel as in 'hear', 'bear', 'far', etc. This allows for more direct consideration of whether syllabic /ɹ/ behaves like a vowel + /ɹ/ sequence. The chart below in Figure 7 shows average duration of rhoticity values in uncontroversial /Vɹ/ for all subjects across the three production contexts. The rhotic portion most frequently makes up less than 60% of the overall /Vɹ/ sequence, although this is not without exception. Subject 09, the most strongly rhotic of the 12, exhibits rhoticity as high as 68% in individual tokens of this type. Also evident in Figure 7 is the intercontextual variation exhibited by all subjects to varying degrees. Effects of production context on degree of rhoticity are somewhat sporadic. For 6 of 12 subjects, word list (WL) context shows the highest average rhotic duration, compared to sentential context (SC) for 2 (02 and 09) and extended context (EC) for the remaining 4 subjects (04, 07, 10 and 11).

The considerable spread in rhotic duration in uncontroversial /Vɹ/ sequences is important in considering the most appropriate phonological representation to assign to perceptually stable syllabic /ɹ/, as the evidence collected demonstrates that in no context does postvocalic /ɹ/ necessarily represent 50% or less of the /Vɹ/ sequence. Across WL contexts for all 12 subjects, for example, there is an average range across individual tokens of approximately 32 percentage points (pp.) in duration of rhoticity. The widest range for individual WL tokens (exhibited by subject 11) is approximately 53pp. while the narrowest is subject 10's 12pp. As noted above, no one context (WL, SC, or EC) consistently elicited more pronounced vowel/rhotic coarticulation than the others.

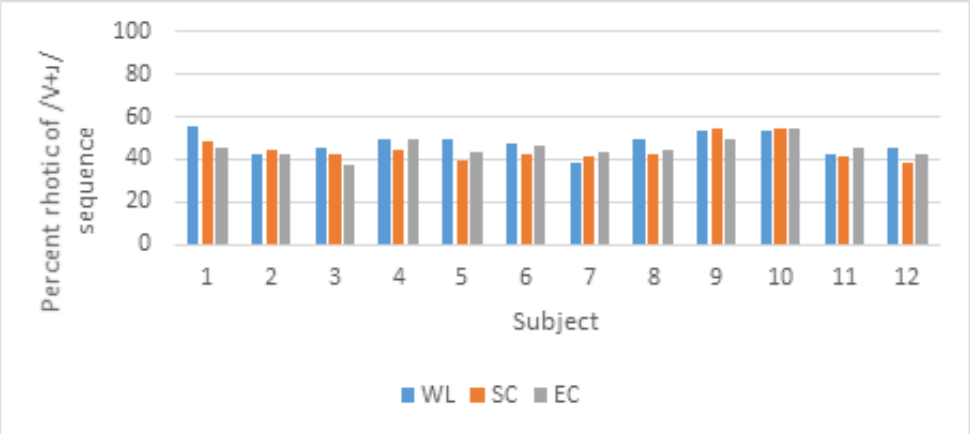

Figure 7 Average duration of rhoticity (measured as percentage of /V+ɹ/ sequence) for uncontroversial /Vɹ/

The postvocalic tokens being considered follow five different vowels: /i: eɪ ɑ ɔ u: ʊ/, so it is unsurprising that rhoticity should encroach on 'pure' vowel duration to varying extents. However, it is worth noting that interaction between different vowel qualities and rhotic duration is not consistent across speakers. The /i:/ context of tokens 'peer', 'fear', 'hear', and 'year', for example, may produce both the highest and lowest rhotic durations, depending on the subject in question. This points to the role played by interspeaker variation in /ɹ/ articulation strategy. Bunching and retroflexion, for example, should logically interact with the tongue shape required to produce a close front vowel /i:/ in different ways. In short, the considerable range in rhotic duration across tokens with different preceding vowel contexts demonstrates the importance of vowel quality in surface output of vowel+rhotic sequences, while interspeaker variability in correlation between specific vowel quality and rhotic duration potentially highlights the role played by speaker-specific choice of /ɹ/ articulation strategy.

5.2.2 Stressed syllabic /ɹ/

The chart in Figure 8 displays the average duration of rhoticity in tokens of stressed syllabic /ɹ/ for all subjects across the three elicitation contexts. Unsurprisingly, rhotic duration is much greater in these syllabic /ɹ/ tokens than in the /Vɹ/ tokens discussed above. The latter, in spite of considerable intra- and interspeaker variation, never have rhoticity exceeding 60% of the vowel+rhotic sequence (on average). By contrast, the former exhibit average rhotic durations which not only frequently exceed 60% of the total /Vɹ/ sequence, but which at times exceed 80% of said sequences. However, not all subjects exhibit equivalent disparities between uncontroversial /V+ɹ/ and syllabic /ɹ/. For subjects 05, 06, and 07, average cross-contextual rhotic values for the two token types are within 13pp. of one another (indeed, subjects 05 and 06 both exhibit disparity within 9pp.). Immediately evident upon visual comparison between the data in Figure 8 and that in Figure 7 is the more pronounced range

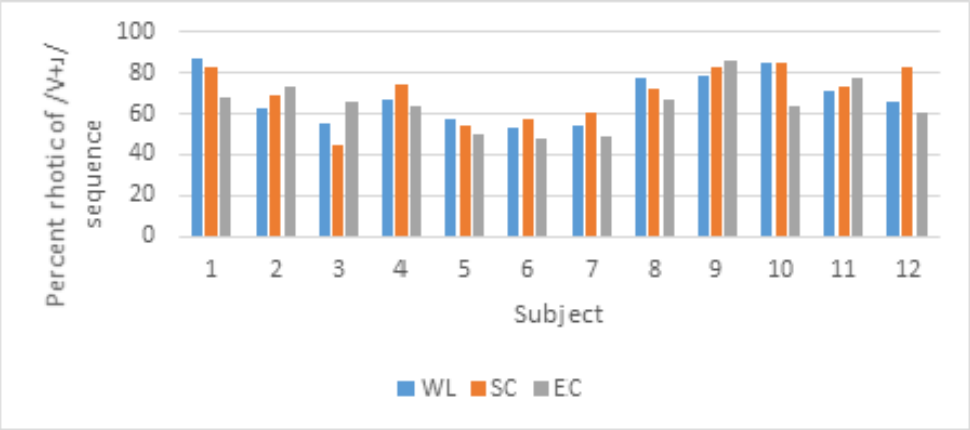

Figure 8 Average duration of rhoticity for stressed syllabic /ɹ/

of intercontextual variation exhibited by the former. On average, the range of intercontextual variation for the data in Figure 8 is approximately 13pp. The lowest average rhotic duration for stressed syllabic /ɹ/ is exhibited by subject 03 in the sentential context with a value of approximately 45%, while the highest value (approximately 87%) is found in subject 01's word list context. It is important to note that these are average values; 100% rhoticity in stressed syllabic /ɹ/ is produced with variable frequency by all subjects. For some subjects, 100% rhoticity is anomalous, often not exceeding a scattered handful of tokens (subject 03, for example, produces it only once in one of the extended context tokens). For others, however, immediate postconsonantal achievement of stable rhoticity is frequently attested: 100% rhoticity is the median value for subject 09 in 2/3 contexts and for subject 12 in 1/3. Interestingly, no one context consistently elicits stronger apparent rhoticity than the others: 4 subjects (01, 05, 08, and 10) exhibit peak rhotic duration in the word list context, 4 (04, 06, 07, and 12) do so in the sentential context, and 4 (02, 03, 09, and 11) do so in the extended context.

Based on the above, it is evident that stressed tokens of syllabic /ɹ/ exhibit a considerable range of variation in their surface realization. The degree to which /ɹ/ exhibits coarticulation with the preceding vowel (assumed to be the NURSE vowel /ɜ/) varies by both speaker and context. Perhaps most importantly, although 100% rhoticity (i.e. immediate, stable [ɜˠ]) is certainly attested, it cannot be said to be the norm produced by the majority of speakers in any context. Furthermore, the fact that 100% rhoticity is attested in a variety of consonantal contexts (and to some extent by all subjects) arguably justifies considering cases of, e.g., a 25% lag in onset of rhoticity as evidence of an underlying vowel surfacing with considerable coarticulation rather than of an underlying rhotic vowel the acoustic lag in the achievement of which is the result of transition out of the preceding consonant.

Briefly returning to the issue of [ɜˠ]/[ɹ] minimal pairs discussed earlier, the issue of ambisyllabicity in ['ɜˠ(ɹ)V] sequences was highlighted in a number of subjects' productions of a subset of stressed syllabic /ɹ/ tokens. For the majority of speakers, tokens such as 'turret' or 'juror' were produced with a perceptually (and acoustically)

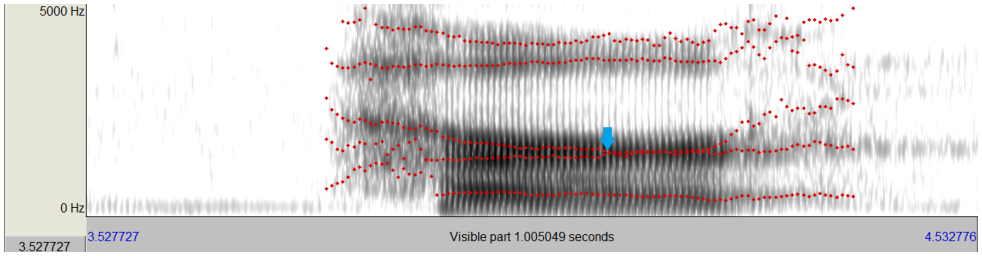

Figure 9 'turret' with ambisyllabic rhoticity. Min.F3-F2 rhoticity is strong, and [ɜˠ] is perceptually stable. There is, however, also a definite perceptual [ɹ] in the onset of the second syllable before [?], perhaps corresponding to the pronounced merger between F3 and F2 (marked)

ambisyllabic [ɹ] between [ɜ] (rhotacized to varying extents) and the following vowel. In Figure 9 above, I have provided illustration of this ambisyllabicity. The fact that this [ɹ] is so frequently produced suggests an underlying representation of /ɜɹ/ variably surfacing under the influence of allophony, connected speech processes, and syllabification considerations. The alternative, underlying /ɜˠ/ with a surface excrescent [ɹ], ignores the obvious historical presence of /ɹ/ in these sequences and (for this study) relies on a small minority of subjects who produce stable [ɜˠ] without the problematic ambisyllabic [ɹ].

5.2.3 Unstressed syllabic /ɹ/

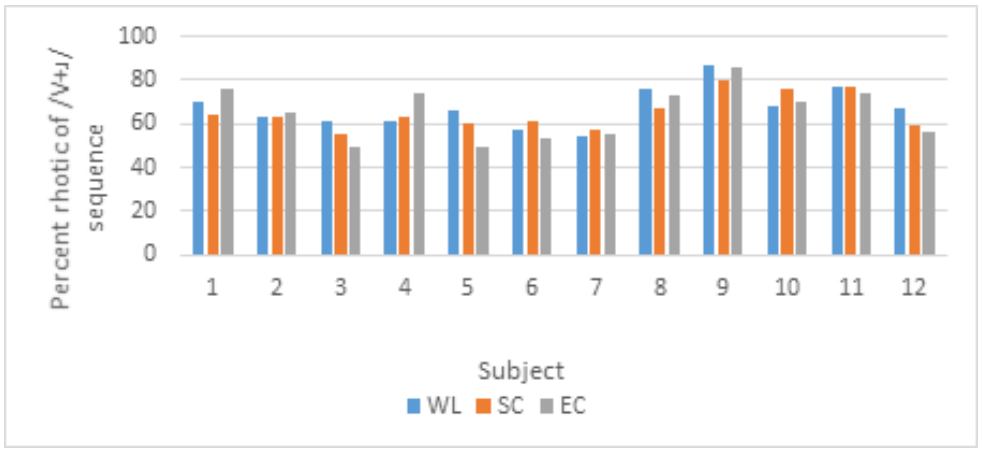

Figure 10 Average duration of rhoticity for unstressed syllabic /ɹ/

The data from subjects' production of tokens of unstressed syllabic /ɹ/ are presented above in Figure 10. Although still more pronounced than in tokens of uncontroversial /V+ɹ/, the range of variation in apparent strength of rhoticity can be seen to be narrower for unstressed tokens of syllabic /ɹ/ than for stressed ones. The average intercontextual variation across all subjects is approximately 9pp. (compared to 13pp. for stressed tokens). Average rhotic duration appears at its highest

value in different elicitation contexts, depending on the subject. For 5 subjects (03, 05, 08, 09, and 12) the word list context produced the strongest average rhoticity, for 4 (06, 07, 10, and 11) rhoticity was strongest in the sentential context, while the extended context produced the strongest rhoticity for the remaining 3 (01, 02, and 04). The context which elicited the strongest average rhoticity was only identical for stressed and unstressed tokens for 5 subjects (02, 05, 06, 07, and 08). As with stressed tokens, total (100%) rhoticity was attested to some extent by all subjects, although for the majority of subjects this was not the case for the majority of tokens. A number of subjects produced 'horror' as ['hɔ.ɚ] without medial [ɹ], unfortunately there was no context for elicitation of the potential minimal contrast with 'whore' ['hɔɹ] discussed in section 4.1 (or an equivalent contrast such as 'poor'/'poorer' ['pɔɹ]/['pɔ.ɚ]).

As illustrated in Figures 11 and 12 below, there is a positive correlation (with a Pearson correlation coefficient of .89446) between average strength of rhoticity in stressed and unstressed tokens of syllabic /ɹ/ with subjects falling into one of three broad categories of rhotic strength. Four subjects can be seen to produce relatively weak rhoticity (i.e. rhoticity not exceeding 60%, on average) and one subject produces exceptionally strong rhoticity (exceeding 80% rhoticity on average), while the remaining seven fall between these two extremes.

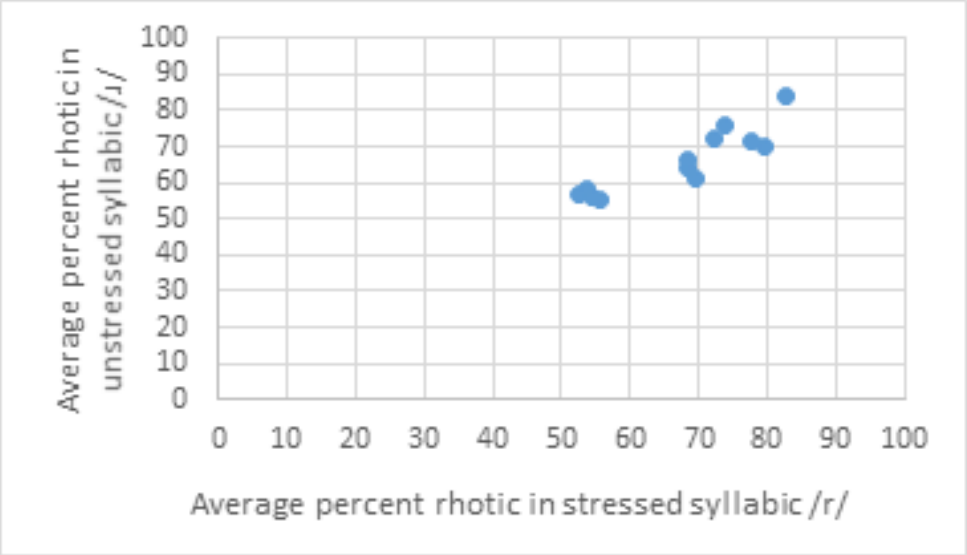

Figure 11 Percentage rhoticity in stressed vs. unstressed tokens of syllabic /ɹ/ for all subjects

5.2.4 5.2.4 Syllabic /l/

Production of the limited number of syllabic /l/ tokens proved parallel to that of unstressed syllabic /ɹ/ in terms of variability in interconsonantal schwa realization. The spectrograms in Figures 13 and 14 show productions of syllabic /l/ with and without schwa, respectively. Syllabic /l/ was elicited for 7 tokens following /t/ (2

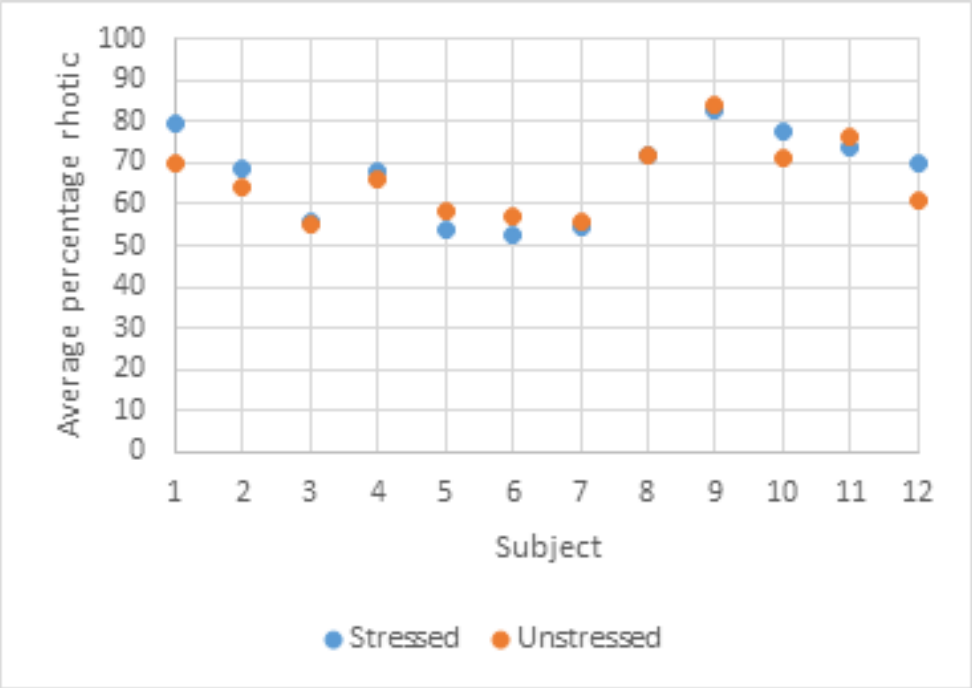

Figure 12 Comparison of percentage rhoticity in stressed vs. unstressed tokens of syllabic /r/ for individual subjects

tokens), /d/ (4 tokens), and /b/ (1 token). In nearly all cases, the purest syllabic /l/s (i.e. those without any significant trace of schwa) were produced following the /b/ of 'label'. Consistent intersubject behavior for different preceding consonant contexts contrasts with the variability found in syllabic /ɹ/, presumably resulting from /l/ having a uniform articulation strategy across all subjects. The remaining tokens elicited a considerable range in degree of pre-/l/ schwa production for all subjects. If it is assumed that /əl/ is the most appropriate underlying representation for surface [l̩] (see earlier discussion in section 2), then the attested variability in syllabic /l/ production is of particular relevance for discussion of putatively syllabic /ɹ/. Syllabic /l/ is evidently able to surface without any trace of the underlying schwa (i.e. true [l̩]) in certain contexts while surfacing with a schwa of variable duration in other contexts. Furthermore, while the largest delays in /l/-onset attested (e.g. subject 01's delay of 72% in one token) are almost certainly the result of hyperarticulation, the fact that hyperarticulation leads to an emphatic schwa rather than a prolonged liquid points (admittedly somewhat anecdotally) to an underlying representation which includes schwa. Most relevant to the present discussion is the fact that the range of variability in the surfacing of /əl/ → [l̩] is comparable to that which subjects produce for syllabic /ɹ/ in both stressed and unstressed contexts, suggesting that surface [ɚ] is just as compatible with an underlying /əC/ representation as is [l̩].

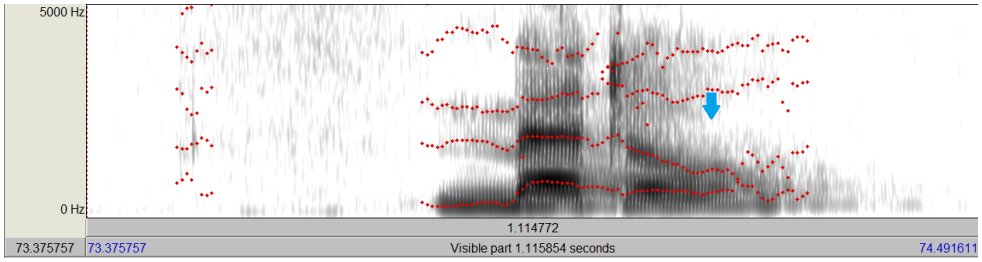

Figure 13 Schwa-ful production of potentially syllabic /l/ in 'meddle', onset of acoustically and perceptually distinct [l] labeled

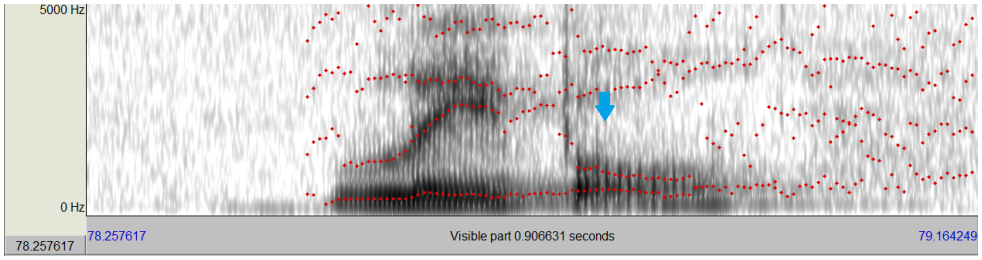

Figure 14 Schwa-less production of syllabic /l/ in 'label', with [l] onset (immediately post-plosive) identified

5.2.5 'Double' syllabic /ɹ/

Finally, there is a class of tokens which straddles stressed and unstressed syllabic /ɹ/, namely cases of two adjacent syllabic /ɹ/s in words like 'stirrer'. There were five tokens of this type presented to subjects in the elicitation task. Two of these tokens ('laborer' and 'harborer') contain two adjacent unstressed syllabic /ɹ/s, while the remaining three ('juror', 'stirrer', and 'insurer') contain a stressed syllabic /ɹ/ followed by an unstressed one. My prediction for the production of these tokens was that if syllabic /ɹ/ is underlyingly /əɹ/, then this should be reflected acoustically in cases of 'double' syllabic /ɹ/. In other words, the formant structure produced by the ideally /ə/-/ɹ/ differentiating speaker should show a relatively wide F3-F2 distance corresponding to two schwas punctuated in the middle and at the end by two points of min.F3-F2 corresponding to two target /ɹ/s. Conversely, if surface [ɚ] is actually phonemic, one would not expect this differentiation. Rather, the result should essentially be a long [ɚː]. Results were mixed; some subjects produced consistent differentiation of all double syllabic /ɹ/ sequences (i.e. something approximating [əɹəɹ]), some produced virtually no differentiation (the contrast, if produced at all, between 'stir' and 'stirrer' being a matter of rhotacized vowel length), with the remaining subjects representing something in between these two extremes. Variation in the latter class of speakers was wide, tentatively correlating (on an individual subject basis) with presence of stressed tokens, morpheme boundaries (i.e. differentiation for the relatively novel 'harborer' but not for the opaque 'juror'), and context of elicitation. In Figures 15 and 16 below, I have included the spectrograms of a differentiated production of 'stirrer' contrasted with an undifferentiated one

(produced by subjects 03 and 10, respectively). While the undifferentiated productions can be unproblematically described in an approach in which underlying /əɹ/ surfaces as [ɚ ɝ] with varying degrees of discreteness, the frequently attested differentiated tokens are not nearly as compatible with a phonemic /ɚ ɝ/ approach. This being the case, the acoustic evidence from double syllabic /ɹ/ seems to favor an underlying representation of /əɹ/.

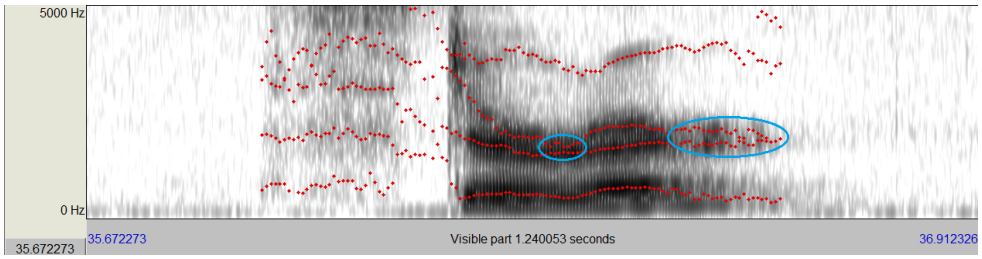

Figure 15 'stirrer' ['stɝ(ɹ)ɚ] with visibly distinct rhotic targets marked, note the widening of F3-F2 for the inter-rhotic schwa

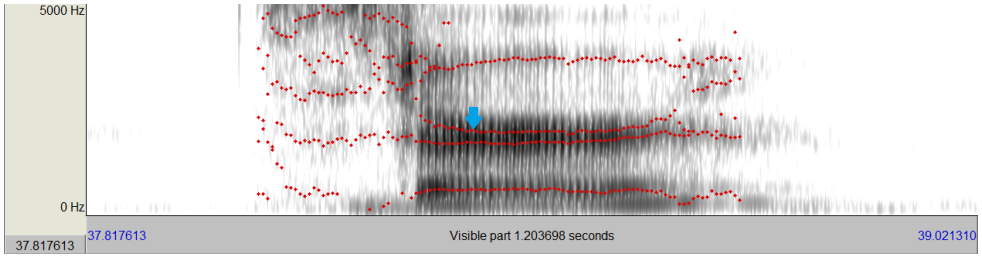

Figure 16 'stirrer' ['stɝː] with no differentiation between the two adjacent syllabic /ɹ/s, onset of min.F3-F2 marked

6 Discussion and Conclusion

Based on the above discussion, I conclude that rhotacized schwa is best analyzed as a surface phenomenon deriving from an underlying /V+ɹ/ sequence. This is the case for both stressed [ɝ] and unstressed [ɚ]; whether the hypothesized underlying vowel [ɜ] of the former is phonemic (i.e. the NURSE vowel) is irrelevant. The evidence to support this conclusion is summarized below.

As discussed in section 2, there is good reason to consider all syllabic consonants to be surface phenomena deriving from underlying /əC/ (Wells 1995, Akamatsu 2013, de Jesus Arboleda & Monroy 2015). Schwa-ful pronunciations of potentially syllabic consonants, while often perceived as hyperarticulatory or childlike, are frequently attested and can be used for lexical contrast in place of the syllabic/non-syllabic distinction. 'Lightening' as 'ligh[tən]ing' (rather than syllabic 'light[n̩]ing') vs. 'lightning? 'light[n]ing' is valid if somewhat bizarre, and is arguably no more awry than any other allophonic substitution (e.g. unaspirated /p t k/ or velarized /l/ in syllable onsets). The fact that [əC] and [C̩] are able to covary in pronunciation

supports the inclusion of /ə/ in the underlying representations of such sequences. Furthermore, degree of schwa production in these sequences has been tied to "external factors such as ... gender, accent, speech rate [and] emphasis placed upon words" (de Jesus Arboleda & Monroy 2015). Based on the above, an ideal underlying representation /əC/ of syllabic consonants was relied upon for the present study. Although the majority of work on this topic has been carried out on non-rhotic varieties of English and therefore often marginalizes discussion of syllabic /ɹ/ in favor of /l m n/, it is logical to extend the same reasoning used to discuss the latter to consideration of the underlying representation of syllabic /ɹ/ (i.e. [ɝ ɚ]).

Phonological evidence for the underlying status of rhotacized schwa/syllabic /ɹ/ is most readily reconcilable with a schwa-inclusive /ɜɹ əɹ/ approach. This is true of what limited evidence of potential [ɝ ɚ]-[ɹ] lexical contrast is available, and more obviously so for the overwhelming tendency for fluent backward talkers to reverse rhotacized schwa as [ɹɜ ɹə] (Cowan et al. 1985). It is telling that rhotacized schwa was only preserved in the reversals of Cowan et al.'s backward talkers in 7/44 tokens, 6 of which were produced by a subject who favored reversal of syllable ordering rather than phoneme ordering.

The acoustic evidence presented in section 5 aligns with results predicted for a system containing underlying /Vɹ/ rather than a primitive rhotacized schwa. Using a stricter metric for onset of rhoticity based on the achievement of a minimum distance between F3 and F2 derived from work on acoustic cues for perceptual /ɹ/ (na Dalcher et al. 2008, **?**), a consistent delay in attainment of rhoticity in putatively syllabic postconsonantal /ɹ/ is evident in a majority of subjects studied. Although for a small minority of subjects, immediate postconsonantal rhotic achievement was ubiquitous (i.e. "pure" rhotacized schwa in both a perceptual and acoustic sense), the majority of the 12 subjects produce a prerhotic vowel in a significant number of putatively syllabic tokens often corresponding to a third or more of the total vowel + rhotic sequence. Given the often cited variation in articulation strategy used to achieve /ɹ/, and given the asymmetry between these strategies with respect to strength of global coarticulatory force (Lawson et al. 2013), inter- and intrasubject variability in duration of pre-rhotic schwa (if it surfaces at all) can likely be related to individuals' choice of /ɹ/ articulation strategy and said strategy's interaction with specific phonetic context. Additionally, variation across different production contexts (word list, sentential context, and extended context) and lack of consistent correlation between a single context-type and relative strength of rhoticity is consistent with de Jesus Arboleda & Monroy's (2015) emphasis on influence of external factors on schwa production in potentially syllabic consonant realization.

An analysis of the data based on phonemic rhotacized schwa struggles to account for the overwhelming frequency with which prerhotic schwa (acoustically distinct enough to be distinguishable from formant transitions out of the preceding consonant) was seen to occur in the majority of the 12 speakers studied. A compromise approach suggesting phonemic /ɝ/ as opposed to compositional unstressed /əɹ/ → [ɚ] (i.e. Wells's (2011) NURSE vowel suggestion) is also not borne out by the data, as although stressed sequences tend to display stronger rhoticity than unstressed

ones at first glance, this trend is not symmetrically robust across all subjects nor is pure rhotacized schwa significantly more common in stressed tokens (see Figures 8 and 10 in section 5). Rather, as illustrated in the plots in Figures 11 and 12, relative strength of rhoticity is largely consistent on an intraspeaker basis between stressed and unstressed contexts.

Instead of patterning with other postconsonantal vowels in terms of target formant structure achievement, putatively syllabic /ɹ/ appears to behave like tokens of postvocalic /ɹ/. This is evident upon comparison with full vowel + /ɹ/ sequences such as that found in 'fear'. Strength of rhoticity in syllabic /ɹ/ tended to parallel general strength of rhoticity in uncontroversial /Vɹ/ sequences for all subjects. The fact that /ɝ əɹ/ exhibits more pronounced coarticulatory effects (i.e. partial or complete rhotacization of the vowel) can feasibly be attributed to schwa's acoustic and articulatory proximity to /ɹ/ by comparison to the rest of English's vowel inventory (see in particular gestural accounts relating liquids [ɹ l] to [ə ɔ] (McMahon, Foulkes & Tollfree 1994, Gick 1999, Gick, Kang & Whalen 2002)). This, in interaction with rate of speech, degree of emphasis, and speaker idiosyncrasies (including choice of /ɹ/ articulation strategy) can account for the wide variation present in the data. This variation cannot be effectively accounted for by attributing phonological status to syllabic /ɹ/, i.e. by positing phonemic rhotacized schwa. Thus, although often perceptually and, at times, acoustically, stable as rhotacized schwa, based on evidence from the present acoustic study and on previous discussion in the literature of both /ɹ/'s acoustics and articulation and of the abstract status of syllabic consonants in general, I conclude that American English syllabic /ɹ/ is best analyzed as underlyingly /V+ɹ/.

References

Akamatsu, Tsutomu. 2013. Syllabic consonants in English: phonetic and phonological effects. *Moenia* 19. 149–224.

Boersma, P. & D.Weenink. 2016. Praat: doing phonetics by computer. *Software Version 6.0.17* .

Bonilla, Joan Carles Mora. 2003. The Formation of Syllabic Consonants and Their Distribution in Southern British English. *Atlantis* 25(2). 97–112.

Chen, N.F., J.L. Slifka & K.N. Stevens. 2007. Vowel Nasalization in American English: Acoustic Variability due to Phonetic Context. *ICPhS* 16. 6–10.

Cowan, Nelson, Martin D.S. Braine & Lewis A. Leavitt. 1985. The Phonological and Metaphonological Representations of Speech: Evidence from Fluent Backward Talkers. *Journal of Memory and Language* 24. 679–698.

na Dalcher, Christine Villafa Rachel-Anne Knight & Mark J. Jones. 2008. Cue Switching in the Perception of Approximants: Evidence from two English Dialects. *University of Pennsylvania Working Papers in Linguistics* 14(2). 62–71.

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

Delattre, Pierre & Donald C Freeman. 1968. A Dialect Study of American R's by X-Ray Motion Picture. *Linguistics: An International Review* 44. 29–68.

Espy-Wilson, C. 2004. *Articulatory Strategies, Speech Acoustics and Variability* June 11-13 2004. MIT, Boston: MIT Press.

Gick, B. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16. 29–54.

Gick, B., P. Bacsfalvi, B.M. Berhardt, S. Oh, S. Stolar & I. Wilson. 2007. A motor differentiation model for liquid substitutions: English /r/ variants in normal and disordered aquisition? *Proceedings of Meetings on Acoustics* 1. 1–9.

Gick, B., A. Min Kang & D.H. Whalen. 2002. MRI evidence for commonality in the post-oral articulation of English vowels and liquids. *Journal of Phonetics* 30. 357–371.

Gimson, Afred C. 1962. *An Introduction to the Pronunciation of English.* London: Edward Arnold Ltd.

Hasegawa-Johnson, Mark, Ken Chen, Jennifer Cole, Sarah Borys, Sung-Suk Kim, Aaron Cohen, Tong Zhang, Jeung-Yoon Choi, Heejin Kim, Taejin Yoon & Sandra Chavarria. 2005. Simultaneous recognition of words and prosody in the Boston University Radio Speech Corpus. *Speech Communication* 46. 418–439.

International Phonetic Association. 1989. Report on the 1989 kiel convention. *Journal of the International Phonetic Association* 19(2). 67–80.

International Phonetic Association. 1993. Council actions of revisions of the IPA. *Journal of the International Phonetic Association* 23(1). 32–34.

de Jesus Arboleda, Immaculada & Rafael Monroy. 2015. English syllabic consonants vs. schwa at discourse level. *Revísta Española de Linguística Aplicada* 28(1). 17–47.

Jones, Daniel. 1967. *The Phoneme: Its Nature and Use.* Cambridge, London, New York, Melbourne: Cambridge University Press.

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

Klein, H.B., Maria I. Grigos, Tara McAllister Byun & Lisa Davidson. 2012. The relationship between inexperienced listeners' perceptions and acoustic correlates of children's /r/ production. *Clinical Linguistics and Phonetics* .

Kuecker, Karrie, Sarah Lockenvitz & Nicole Muller. 2013. Amount of rhoticity in schwar and in vowel+/r/ in American English. *American Journal of Speech-Language Pathology* 22. 540–553.

Laver, J. 1994. *Principles of Phonetics.* Cambridge: Cambridge University Press.

Lawson, Eleanor, Jim Scobbie & Jane Stuart-Smith. 2013. Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics* 41. 198–210.

Lockenvitz, Sarah, Karrie Kuecker & Martin J. Ball. 2015. Evidence for the distinction between 'consonantal-/r/' and 'vocalic-/r/' in American English. *Clinical Linguistics and Phonetics* 1–10.

McGowan, R.S., S. Nitrouen & C.J. Manning. 2004. Development of [r] in young, Midwestern and American children. *Journal of the Acoustical Society of America*

115. 871–884.
McMahon, A., P. Foulkes & L. Tollfree. 1994. Gestural representation and Lexical Phonology. *Phonology* 11. 277–316.
Mielke, Jeff, Adam Baker & Diana Archangeli. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory Phonology* 10. 699–730.
Scobbie, James M., Eleanor Lawson, Satsuki Nakai, Joanne Cleland & Jane Stuart-Smith. 2015. Onset vs. Coda Asymmetry in the Articulation of English /r/. *Proceedings of the 18th ICPhS* .
Wells, John. 2011. John Wells's phonetic blog. Accessed December 8, 2016. http://phonetic-blog.blogspot.co.uk/2011/12/more-syllabic-consonants.html.
Wells, John C. 1995. New Syllabic Consonants in English. In *Studies in General and English Phonetics: Essays in Honour of Professor J.D. O'Connor*, 401–414. London and New York: Routledge.

Connor McCabe
University of Cambridge
cm783@cam.ac.uk

A /ɹ/ Elicitation Materials

A.1 Word List

bird	sure	rat	peer	flower	harborer	bar	cattle
her	pearl	raw	fear	color	juror	far	meddle
fur	earl	rum	hear	healthier	stirrer	par	needle
work	early	parent	year	discover	insurer	hard	doodle
perfect	father	rarer	hair	survive	meddler	lure	saddle
sir	mister	horror	fair	orchard	reed	lore	label
servant	butter	forest	wrapper	other	pear	poor	
turret	raid	favorite	after	laborer	bear	bottle	

A.2 Sentential Context

I saw a bird today.
I saw her today.
The cat's fur is soft.
It's a lot of work, but it pays off.
This is a perfect fit.
"Ma'am" and "Sir" are more formal.
The servant has gone home.
The turret is tall.
I am sure about this.
There's a pearl in that oyster.
She married the Earl of Wessex.
We got up early in the morning.
My father is a banker.
Mister Smith will see you now.
I can't eat butter or eggs.
The empty wrapper is in the bin.
I heard from him after several days.
My front garden has a single flower.
There is no color quite as versatile as black.
He's been looking much healthier lately.
She was surprised to discover his secret.
The blossoms will survive well into autumn.
The fruit in the orchard is nearly ripe.
There is no other way to proceed.
He was a laborer for many years.
He was a known harborer of fugitives.
Each juror was carefully instructed.
I put the stirrer in my coffee.
Tell the receptionist which insurer you use.
No one likes a meddler.
My brand new reed just broke.
There was a raid on the compound.
She saw a rat in the kitchen.
Vegetables are eaten raw or cooked.
I'll have a rum and coke.
As a parent, I strongly agree.
It certainly is rarer these days.
With horror, I realized my mistake.
He ventured into the forest for firewood.
Chocolate is his favorite snack.
The child tried to peer through the window.
I have a fear of heights.
I can't hear a word he's saying.
It's been a year since the incident.
There's dog hair on the furniture.
I think that's fair to say.
When the pear is ripe, I'll eat it.
I saw a bear on the road.
I went to the bar for a drink.
We're far away now.
That's par for the course.
Some mattresses are hard, others soft.
The shiny lure attracted the fish.
Folklore is often part of an oral tradition.
The working poor suffered under the policy.
Does anyone have a bottle opener?
Grass-fed cattle tend to be healthier.
I don't like to meddle in the affairs of others.
It's like a needle in a haystack.
She proceeded to doodle all over her notes.
The old saddle still fit nicely.
I don't like to label myself or others.

A.3 Extended Context

Mister Johnson, who passed early this morning, was a kind and decent man who will surely be missed. To the horror of many, the healthy man of fifty died from ingesting raw, unpasteurized butter, having failed to read the warning label on the wrapper. Johnson worked as a laborer for twenty-five years, and was proud to have built his own home – the house with the turret on the edge of the forest. After a hard day's work, Johnson was often to be found at the bar, working his way through a bottle of rum while trying to meddle in community affairs.

Although never a parent, many of you will recall that Johnson once raised two bear cubs which he discovered abandoned in his pear orchard.

It is fair to say that smiles will be somewhat rarer as our community copes with this loss. He is survived by his poor wife Jane, who works as a home insurer, and his beloved pet bird. Jane has requested that only red flowers be sent to the memorial, as it was by far his favorite color.