THE NATURE OF VOCOIDS ASSOCIATED WITH SYLLABIC CONSONANTS IN TASHLHIYT BERBER

John Coleman
University of Oxford, UK

ABSTRACT

Tashlhiyt Berber has entered the phonological folklore as an unusual language in which any consonant can be syllabic, many words consisting entirely of consonants. I shall argue for an alternative analysis, according to which the epenthetic vowels which frequently accompany syllabic consonants are the phonetic realizations of syllable nuclei. Where no epenthetic vowel is evident, it can be regarded as hidden by the following consonant, according to a gestural overlap model. On this view, Tashlhiyt syllable structure is a quite unmarked CV(C(C)), and syllabification is unproblematic.

1. INTRODUCTION

Many influential phonology texts (e.g. [9, 13]) repeat Dell and Elmedlaoui's analysis of Tashlhiyt [3], which holds that all Tashlhiyt phonemes, even obstruents, are sometimes syllabic, and that numerous words consist only of consonants (e.g. /tχ̩zn̩tn̩t/ 'and store them'). If that is correct, the cross-linguistic syllabification algorithm of [8, 10], according to which syllable nuclei are projections of vowels, is inapplicable to Tashlhiyt. [3] also proposed a syllabification procedure in which relative, rather than absolute, sonority plays the main part in locating syllable nuclei. However, that procedure rather implausibly requires nine iterative passes through each word. In this paper I shall present evidence in support of an alternative proposal [2], according to which syllabic consonants are phonologically analyzed as a vowel+consonant sequence. In defence of this view I shall show that in Mohamed Elmedlaoui's speech, syllabic consonants are frequently (but variably) spoken as vowel+consonant sequences. I shall confirm Dell and Elmedlaoui's claim [3, 4, 5] that the quality of these vowels (which they call 'transitional schwas') is predictable on the basis of the neighbouring consonants, and I shall show that the openness of the next vowel is also a factor. The vowels in question are thus epenthetic. I shall show that their distribution is better accounted for by the hypothesis that they occur to fill syllable nucleus positions that would otherwise be empty than by the alternative proposal of [3, 4, 5].

The phonemes of Tashlhiyt are: voiceless stops /t tˤ k kʷ q qʷ/, voiced stops /b d dˤ g gʷ/, voiceless fricatives /f s sˤ ʃ ʃˤ χ χʷ ħ h/, voiced fricatives /z zˤ ʒ ʒˤ ʁ ʁʷ/, nasals /m n nˤ/, liquids /l lˤ r rˤ/, approximants /j w ʕ/, and vowels /i u a/. I shall accept the claim in [3] that /j/ and /w/ are not separate phonemes, but non-nuclear variants of /i/ and /u/. A word containing a pharyngealized phoneme exhibits pharyngealization throughout, and has backer, slightly more open vowel qualities. Labialization of velar and uvular obstruents is usually contrastive (e.g. /ikti/ 'hot' vs. /ikʷti/ 'recall'). /r/ is pronounced as [ɾ], [r], or [ʀ], with brief vocoids before and sometimes after the tap, and between the individual closures of trills. Consonants may be contrastively geminate, even word-initially (e.g. /ttgʷgʷa/ 'be washing clothes'). Typically for an Afroasiatic language, many words are formed by intercalating independent vowel and consonant melodies, e.g. singular /aʃakʷʃ/ 'pile of stones', plural /iʃukaʃ/. I shall refer to phonemic /i/, /u/, and /a/ as 'lexical vowels' and to the so-called 'transitional schwas' as 'epenthetic vowels'.

2. MATERIALS

2.1. Recordings

In the absence of a Tashlhiyt dictionary, an unsorted list of 555 words was collated from various papers. Each word was written down in the frame sentence "ini za — yat tklit adˤnˤinˤ" ('please say — again'), and the list of sentences was read aloud by two native speakers. The list was recorded twice by each speaker. The specific questions to be investigated were not known at the time of recording, so the database is unbalanced, aiming more for breadth of coverage than numerous repeats. In each session, 311 words occurred only once, 218 words were spoken twice, 15 words were spoken three times and 11 words were spoken four times, giving 895 tokens per session. The words in [3] were recorded from two to four times by each speaker, so that words containing syllabic consonants are well-represented in the database. Due to constraints on research time, only data from one session with one speaker, Prof. Elmedlaoui, is discussed below. Since he is the principal informant for prior publications [3, 4, 5], it is hoped that this limitation of the present study will be excused.

Digital audio recordings were made in a sound-treated booth at Oxford University Phonetics Laboratory using an AKG C451E microphone, an AKG B46E preamplifier, and a Sony DTC-1000 ES PRO DAT recorder at 48,000 samples/s. The recordings were digitally transferred to the hard disk of a Silicon Graphics Indy computer using Silicon Graphics *DATman* software, and down-sampled to 16,000 samples/s. The sample resolution was 16 bits. All subsequent waveform editing and acoustic analysis was performed using Entropics *waves* speech processing software.

2.2. Acoustic Measurements

In each word containing syllabic consonants (429 tokens), the following were measured: 1) duration of epenthetic vowels, 2) duration of lexical vowels, 3) duration of syllabic consonants (except /r/ and /rˤ/), 4) frequencies of the first three formants. F1–F3 of some lexical vowels were also measured, not as a control, but as indicators of the periphery of the speaker's vowel space. Epenthetic vowels were also impressionistically transcribed.

3. DISTRIBUTION OF EPENTHETIC VOWELS

3.1 Method of Analysis

The database contained 929 sequences of two consonant phonemes (dyads), the environment in which epenthetic vowels may occur. Epenthetic vowels occurred in 368 (40%) of these. The observed probability of epenthesis was calculated for every dyad and for every pair of consonant classes listed in section 1.

Exceptions to Dell and Elmedlaoui's model [5] and to mine [2] were found, necessitating an analysis of their relative correctness. Signal detection theory [11] was used to compare the goodness-of-fit between the observed distribution of epenthetic vowels and their expected occurrence according to the two models. Each dyad was classified according to each model as either a 'hit' (i.e. epenthesis occurred as expected), a 'miss' (epenthesis was not expected even though it occurred), a 'false alarm' (epenthesis was expected but it did not occur), or a 'correct rejection' (epenthesis was not expected and did not occur). Note that there are two different ways in which a model can be wrong, and two ways in which it can be right. The hit rate was expressed as the probability of occurrence and the false alarm rate as the probability of non-occurrence, and these probabilities translated into z-scores using the inverse z-transform. The difference between z(hits) and z(false alarms), d', provides a measure of how well each model distinguishes between the occurrence and non-occurrence of epenthetic vowels. The higher d', the better the model's performance.

In Model A, epenthetic vowels were expected to occur according to the account given in [5]: in brief, they are not expected 1) between two voiceless consonants, 2) between homorganic stops (including the two parts of a geminate), 3) between a stop and a homorganic noncontinuant with a different value for the feature [sonorant] (e.g. /tn/, /dn/, /tl/, /dl/), or 4) after fricatives. Epenthetic vowels are expected to occur 1) between two non-homorganic stops (including nasal stops), especially if both are voiced, and 2) after liquids.

In Model B, epenthetic vowels are expected to occur 1) in a syllable nucleus that is not filled by a lexical vowel, and 2) after /r/. Syllable nuclei are expected to occur according to the template CV(C(C)), where single-consonant codas are preferred over empty codas and coda clusters, and the second coda consonant may not be more sonorous than the first. For example, the first syllable of /tχzntnt/ is /tVχ/ (where V denotes an empty nucleus), not /tVχz/, because /z/ is more sonorous than /χ/. The second syllable is /zVn/, not /zVnt/, because single consonant codas are preferred, whereas a consonant cluster is acceptable in the coda of the third syllable, /tVnt/, because the final /t/ cannot form a syllable by itself.

3.2. Results

The probability of epenthesis (compared case-by-case to absence of epenthesis) was above chance before or after certain consonants, e.g. p(epenthesis) = 0.68 before /l/, 0.79 before /dˤ/, 0.86 after /ʁʷ/, 0.88 after /qʷ/ and 0.91 after /gʷ/. It only exceeded 0.95 before /r/ and /rˤ/ and in the dyad /lg/. Epenthesis occurred without variation only in dyads where the number of instances was very small; these cases cannot be taken as evidence of environments in which epenthesis is obligatory. In contrast, there are some well-represented environments in which epenthesis was never observed, i.e. before /f/, /ʃ/ and /ʒ/, and between two voiceless consonants (except in one instance of /qs/, but the word-list spelling may have been wrong in this case). Other than these, the probability of epenthesis was never below 0.05, though before /s/ it comes close ($p = 0.07$).

In the comparative evaluation of models, both performed well. Model A predicted 75% of the 368 observed epenthetic vowels, but also expected epenthesis in 41% of the cases in which it did not occur, giving a d' of 4.34. The false alarm rate of Model B was similar, 43%, but the hit rate was higher, 89%, yielding a better d' of 4.74. Although the two scores are quite close, Model B should be favoured on three counts: 1) its better coverage at explaining occurring instances of epenthesis; 2) the fact that its statement of where epenthesis is expected is much more succinct than the list of separate cases given in [5]; 3) epenthesis in empty syllable nuclei is paralleled in many other languages [7, 12].

4. QUALITY OF EPENTHETIC VOWELS

4.1. Variation in Quality

[2] reports that a variety of epenthetic vowel qualities are found. In this study, [ɪ], [ɪ̈], [ë], [ɛ̈], [æ̈], [ɑ̈], [ʌ̈], [ɤ̈], [ɐ], [ə̟], [ə], [ɵ], [ö], [ɔ̈], [ʊ], and [ʏ] were recorded. The first three formant frequencies were carefully estimated from a 14-pole discrete Fourier transform of a portion of signal centred on the amplitude peak of each epenthetic vowel and windowed with a 25.6 ms Hanning window. Vowels for which estimation of F1 or F2 proved problematic were disregarded. Figure 1 is a scatter plot of the F1 and F2 of epenthetic vowels. The allophones listed above lie in overlapping clusters which fill the vowel space. A distinction between a more open and a closer group of vowel qualities is evident, however.

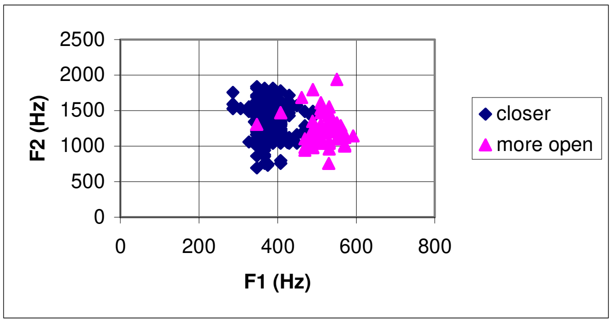

Figure 1. Epenthetic vowel qualities.

In their reply to [2], Dell and Elmedlaoui suggest that these variations in epenthetic vowel quality are due to a "requirement that an articulator must follow the shortest possible path when moving from one target to the next" [4, 5]. In other words, epenthetic vowel quality is determined by the preceding and following consonants. I shall call this the C–C model. If it is correct, we would not expect to find any dyads separated by different epenthetic vowels. In fact, some instances of dyads with different epenthetic vowels do occur in my database, such as the following. (None of these examples contains a pharyngealized consonant which would affect the vowel qualities of the entire word.) Example (4) instantiates the F1 contrast of Figure 1.

1a)	dr̩	Unrounded: F1 ≈ 390 Hz, F2 ≈ 1510 Hz	tajdr̩t [dərət] 'ear of wheat'
b)		Rounded: F1 ≈ 410 Hz, F2 ≈ 1410 Hz	tudr̩t [dɵrət] 'life' [n.f.]
2a)	gl̩	Unrounded: F1 ≈ 360 Hz, F2 ≈ 1700 Hz	trgl̩t [ɡ̈ilʲ] 'you locked'
b)		Rounded: F1 ≈ 390 Hz, F2 ≈ 1225–1350 Hz	jugl̩ [gʊl] 'he hung'
3a)	kr̩	Unrounded: F1 ≈ 390 Hz, F2 ≈ 1550 Hz	tṣk[ə̟]rt 'do' [v.perf.2sg.]
b)		Rounded: F1 ≈ 390 Hz, F2 ≈ 1205–1350 Hz	juk[ʊ]r 'he stole'
4a)	kr̩	Closer, fronter: F1 ≈ 390 Hz, F2 ≈ 1550 Hz	tṣk[ə̟]rt 'do' [v.perf.2sg.]
b)		Opener, backer: F1 ≈ 560 Hz, F2 ≈ 1290 Hz	ak[ʌ]rkur 'blackbird'

4.2. Variation in F1

In order to determine whether the local consonantal context is sufficient to explain the F1 variation of Figure 1, a General Linear Model of epenthetic vowel F1 was constructed, with preceding consonant place, following consonant place, preceding consonant manner and following consonant manner as fixed factors. Consonant manners were the same as the consonant classes of section 1, and the place categories were labial, coronal, dorsal (i.e. velars and uvulars), labiodorsal (i.e. labialized velars and uvulars), and pharyngeal/laryngeal. Of these, preceding and following manner were highly significant ($p < 0.005$; $F(6,76) = 3.442$ and $F(6,76) = 7.409$, respectively), preceding place was significant ($p < 0.05$, $F(4,76) = 2.904$) and preceding place * preceding manner was a highly significant interaction ($p < 0.001$, $F(8,76) = 3.631$). Post hoc multiple comparisons using Tamhane's T2 test were conducted to identify the significant manner distinctions. The largest difference is that the mean F1 of epenthetic vowels is 76 Hz higher after voiced fricatives than voiced stops. This is probably because the only voiced fricative after which epenthesis frequently occurred is /ʁʷ/ and uvular constrictions have a high F1 [12].

The mean F1 of epenthetic vowels was 42 Hz higher before liquids than voiced stops or nasals. This may be because the most frequent liquid, /r/, has the most open oral cavity of the sonorant consonants (apart from /ʕ/, which did not occur after epenthesis). The high incidence of /rˤ/ may also raise the mean F1 of liquids, due to its pharyngeal constriction.

These results do not explain why different qualities are sometimes found in the same consonantal context, nor why some epenthetic vowels are quite open, as in (4), nor why the difference in mean F1 of the two clusters in Figure 1 is 110 Hz. An obvious alternative hypothesis is that neighbouring vowels may also affect epenthetic vowel quality; that is, they may be underspecified or "targetless schwas", whose quality is determined in part by neighbouring vowel gestures, as proposed for English by [1]. I call this the V–V model. By this hypothesis, the epenthetic [ʌ] in /akr̩kur/ 'blackbird' is open because of the open vowel of the preceding syllable, and the epenthetic [ɵ] in /tudr̩t/ 'life' and the epenthetic [ʊ]'s in /jugl̩/ 'he hung' and /jukr̩/ 'he stole' are rounded because of the preceding /u/.

In order to test this hypothesis, a second GLM of epenthetic vowel F1 was constructed with preceding and following vowel as the fixed factors. /j/ and /w/ were counted as instances of /i/ and /u/, respectively. In this model, only *following* vowel was found to be a significant factor ($p < 0.001$, $F(2,8) = 8.778$). Post hoc multiple comparisons using Tamhane's T2 test showed that the distinction between /a/ and /i/ was significant, with mean F1 27 Hz higher before /a/ than before /i/. This gives some support to the idea that vowel context is influential, at least as far as the feature [low] is concerned, but it is a relatively small effect compared with neighbouring consonant manner, and does not explain the four examples in the previous paragraph. A third GLM was constructed which combined the significant factors of the previous two models. In this case, following manner and following vowel were the only significant factors ($p < 0.001$, $F(6,65) = 7.567$ and $p < 0.05$, $F(2,65) = 3.822$, respectively): preceding consonant manner drops out in this combined model. Thus, the main influences on the F1 of epenthetic vowels are the following consonant and the lowness of the following vowel.

4.3. Variation in F2

In order to understand F2 variation, two GLM's were constructed with the same fixed factors as for the F1 models, one instantiating the C–C model, the other the V–V model. For the latter, labialized dorsal obstruents were regarded as containing /u/, because of the effect of labialized consonants on epenthetic vowels, as we shall see. In the C–C model, all four factors were significant, as well as the interaction between preceding and following manner. Post hoc Tamhane tests showed that the mean F2 of epenthetic vowels is significantly lower after voiced fricatives than liquids (which are 231 Hz higher), voiced stops (233 Hz higher), voiceless fricatives (238 Hz higher) and voiceless stops (239 Hz higher). This is probably because the only voiced fricative after which epenthesis frequently occurred is /ʁʷ/. A few instances of preceding /ʁ/ add to the general low mean F2 of epenthetic vowels after voiced fricatives. The F2-lowering effect of voiced fricatives was also evident when they followed epenthetic vowels, in which F2 is on average 182 Hz higher before liquids and 245 Hz higher before voiced stops. Again, /ʁʷ/ and /ʁ/ are mainly responsible.

Place of articulation of neighbouring consonants has a strong effect on the F2 of epenthetic vowels. Mean F2 is 223 Hz higher after coronals and labials than after labiodorsals, 307 Hz higher after dorsals than labiodorsals, and 568 Hz higher after pharyngeals and laryngeals than labiodorsals. We see evidence here for a "gravity hierarchy", with labiodorsals more grave than dorsals, labials and coronals, and laryngeals and pharyngeals least grave. The F2 lowering effect of labialized consonants and dorsal consonants is expected from acoustic theory, and is also seen in the influence of following consonant place: the mean F2 of epenthetic vowels is 225 Hz higher before coronals than before labiodorsals, and 259 Hz higher before dorsals than labiodorsals.

In the V–V model, preceding and following vowels were highly significant ($p < 0.001$, $F(2,8) = 23.783$ and 14.359, respectively). This time, however, it is the distinction between /u/, on the one hand, and /i/ and /a/, on the other, that is criterial. The mean F2 of epenthetic vowels is 209 Hz lower after /u/ (including the labiality of labiodorsal consonants, recall) than after /a/, and 224 Hz lower after /u/ than /i/. Likewise, mean F2 is 125 Hz lower before /u/ (or a labiodorsal consonant) than before /i/ (or /j/).

4.4. Variation in F3

In the C–C model of F3, preceding consonant place and manner and following consonant manner were significant. In the V–V model, preceding vowel was significant. When those four factors were combined, preceding manner remained highly significant ($p < 0.001$, $F(6,89) = 3.735$) and following manner was also significant ($p < 0.05$, $F(6,89) = 2.822$). With this selection of factors, preceding consonant place and preceding vowel were not significant, though preceding place * preceding manner, preceding vowel * following manner, and preceding place * following manner were also significant ($p < 0.05$). The effect of preceding manner on F3 of epenthetic vowels is partly due to fricatives, especially voiced fricatives (/ʁʷ/ and /zˤ/ in particular). Mean F3 of epenthetic vowels is 335 Hz higher after voiced fricatives than after approximants (mainly /w/). F3 is relatively low after the labialized consonants /qʷ/, /kʷ/, /gʷ/, /w/, and /χʷ/. F3 is also low before labialized consonants, apart from /ʁʷ/. Before the voiced fricatives /ʁʷ/, /ʁ/, and /z/, F3 is relatively high.

Vowel	Mean duration (ms)	N	SD (ms)
/a/	68	207	16
/u/	65	79	17
/i/	65	103	21
‘ə’	35	322	10
‘ɐ’	43	44	15

Table 1. Means and standard deviations of vowel durations.

5. DURATIONS OF EPENTHETIC VOWELS

The durations of epenthetic and lexical vowels were compared in various ways. First, means and standard deviations of the three lexical vowel categories, the closer epenthetic vowels [ɪ], [ɪ̈], [ë], [ɤ̈], [ə̝], [ə], [ɵ], [ö], [ʊ], and [ʏ] (labelled ‘ə’) and the more open epenthetic vowels [ɛ̈], [æ̈], [ɑ̈], [ʌ̈], [ɐ], [ɜ̈], (labelled ‘ɐ’) were calculated (Table 1). Though ‘ə’ and ‘ɐ’ are shorter than lexical vowels, their duration distributions overlap with those of lexical vowels. A one-way analysis of variance of vowel duration was conducted with the five vowel categories of Table 1 as the independent variable. Post hoc Tamhane tests showed that the mean durations of /a/, /u/ and /i/ were not significantly different from each other, but they are all significantly different (at the 0.05 level) from ‘ə’ and ‘ɐ’, which are significantly different from each other.

In 9% of words containing an epenthetic vowel and a lexical vowel, epenthetic vowel duration is actually greater than lexical vowel duration. Although they were not studied as closely as epenthetic vowels, lexical vowels may be centralized, shorter and lower in amplitude, sometimes making them more ‘reduced’ than epenthetic vowels.

At their shortest (≤25ms), epenthetic vowels may not be long enough to develop voicing and were not amenable to spectral analysis. My hypothesis is that epenthetic vowels may be eclipsed by adjacent consonants to a greater or lesser degree. Where the degree of overlap is extensive, no appreciable vowel remains. Observations of duration variation are consistent with this idea, which thus allows for the apparent absence of any nuclear vowel, giving rise to surface syllabic consonants.

Finally, variation in vowel duration does not explain the variation in F1 noted earlier, as in a model in which epenthetic vowels exhibit articulatory undershoot towards an open target: duration and F1 are uncorrelated.

6. CONCLUSION

I have presented evidence from the speech of one Tashlhiyt speaker in support of the following claims: 1) Epenthetic vowels are better explained as the phonetic realization of syllable nuclei unfilled by a lexical vowel than by reference to a list of various kinds of C–C combinations. 2) Their quality is to some extent predictable from preceding and following segmental context. 3) However, two clusters of spectrally and durationally distinct epenthetic vowels can be identified. 4) The separation of these clusters is not fully explained by local segmental context.

ACKNOWLEDGMENTS

I am grateful to Jana Dankovičová for transferring the audio recordings to disk and dividing them into files; to Mohamed Elmedlaoui for his patience in making the recordings, as well as for engaging, together with François Dell, in an exceptionally well-tempered debate.

REFERENCES

[1] Browman, C. P. and L. Goldstein. 1992. “Targetless” schwa: an articulatory analysis. In Docherty, G. J. and D. R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody.* Cambridge: CUP.

[2] Coleman, J. 1996. Declarative syllabification in Tashlhit Berber. In [6].

[3] Dell, F. and M. Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics,* 7, 105–130.

[4] Dell, F. and M. Elmedlaoui. 1996. Nonsyllabic transitional vocoids in Imdlawn Tashlhiyt Berber. In [6].

[5] Dell, F. and M. Elmedlaoui. 1996. On consonant releases in Imdlawn Tashlhiyt Berber. *Linguistics,* 34, 357–395.

[6] Durand, J. and B. Laks (eds). 1996. *Current Trends in Phonology: Models and Methods.* Vol. 1. Manchester: University of Salford.

[7] Itô, J. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory,* 7, 217–259.

[8] Kahn, D. 1996. *Syllable-based generalizations in English phonology.* Bloomington: Indiana University Linguistics Club.

[9] Kenstowicz, M. 1994. *Phonology in Generative Grammar.* Oxford: Blackwell.

[10] Levin, J. 1985. *A metrical theory of syllabicity.* MIT PhD dissertation.

[11] MacMillan, N. A. and C. D. Creelman. 1991. *Detection Theory: A User’s Guide.* Cambridge: CUP.

[12] McCarthy, J. J. 1984. The phonetics and phonology of Semitic pharyngeals. In Keating, P. A. (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III.* Cambridge: CUP.

[13] Prince, A. and P. Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* Piscataway, NJ: Rutgers Centre for Cognitive Science.