Moraic Onsets in Arrernte

Nina Topintzi & Andrew Nevins[*]

Abstract:
The Australian language Arrernte has been argued by Breen & Pensalfini (1999) and Evans & Levinson (2009) to present a case of VC syllabification (vowel-consonant constituency, with coda maximization), rather than CV syllabification (consonant-vowel constituency, with onset maximization). In this paper we demonstrate that greater insights for a number of phenomena are achieved when analyzed with CV syllabification and onset consonants that are moraic, a possibility independently proposed for a wide range of languages by Topintzi (2010). Previous analyses were obliged to posit an underlying fleeting initial schwa for surface forms beginning with CV at the left edge; we demonstrate that once the full range of phenomena are considered, no such schwa is desirable, and that these words are underlyingly CV-initial. We review a range of prosodic morphology and external evidence from phonetic studies, acquisition, and musicology that points towards a CV syllabification in Arrernte and provide an analysis for the allomorphy, stress assignment, reduplication, and the transpositional language game 'Rabbit Talk' in terms of reference to moraic structure. The results lend themselves to new directions in the analysis of Arrernte and provide further evidence for moraic onsets in prosodic morphology.

1. Overview: The Universality of CV Syllabification

The question of the universality of the syllable incites particular excitement among phonologists, at least since Jakobson's (1962) proposal of the ubiquity of CV syllabification. In a recent paper, Hyman (2011a) recants an earlier view (Hyman 1983) that Gokana does not have syllables, arguing instead that a restriction on the maximum prosodic stem size in verbs is best accounted for by assuming bisyllabic trochees. In a related facet of the same kind of debate, Labrune (2012) asserts that Japanese does not have syllables, suggesting that only moras play a role in phonotactic and metrical phenomena (however, see Kawahara (2016) for a rejoinder). In the present paper, we focus on a similar question of universality, although with respect to syllabification itself, and in particular the nature of the Onset as a preferred constituent, which in turn favors CV rather than VC syllables.

The CV syllabification universal under examination is the one that favors CV, by the principle of onset maximization (Steriade 1982, Itô 1986), under which coda consonants are allowed, but only as a last resort, when consonants cannot be incorporated as an onset to a following vowel or made into a syllabic consonant. We thus do not limit this to the strict-CV phonology of Lowenstamm (1996), under which codas are claimed never to exist, with empty nuclei following them, but rather the broader claim that a VCV sequence will always be syllabified as V.CV under CV syllabification.

By contrast, a putative VC syllabification algorithm would pursue 'Coda Maximization', whereby onset consonants are only allowed as a last resort, and a VCV sequence will be syllabified as VC.V. The case study at hand is Arrernte, a language previously studied by Breen & Pensalfini (1999), Henderson (2013), and a number of other authors, but brought to the spotlight of the universals debate by Evans & Levinson (2009), who contend that as a result of adopting Breen & Pensalfini's analysis, "The status of the CV assumption in any model of UG must be revised." Evans & Levinson thus have in mind an analysis of the VC type, in which consonants are preferentially syllabified as codas, rather than onsets. Of particular note in this connection, however, is the fact that Breen & Pensalfini (1999: 10; henceforth B&P) themselves state "if there is a viable alternative to VC(C) syllabification of Arrernte, it should be preferred". Our goal in this paper is to provide exactly such an alternative, consistent with CV syllabification and in fact, across many of the phenomena in question, more straightforward than the VC analysis.

[*] For very helpful suggestions and discussion, we thank Myf Turpin, David Wilkins, Rob Pensalfini, Marija Tabain, Toni Borowsky, Michael Becker, Leo Wetzels, Moira Yip, and Bob Ladd; and the editors, associate editor, and anonymous reviewers of *Phonology.*

Investigation of the nature of syllabification in Arrernte involves processes such as stress assignment, prosodically-conditioned allomorphy, ludlings, and reduplication, a set of phenomena that are often crosslinguistically weight-sensitive, as extensively catalogued in McCarthy & Prince (1986). Our argument in this paper is that these processes are weight-sensitive in Arrernte as well. What makes Arrernte somewhat novel from the perspective of crosslinguistic typology is the fact that the weight-bearing unit to which such processes are sensitive – the mora – can be assigned to Onset consonants, an analysis that has been proposed for a number of unrelated languages (e.g. Davis 1988, Topintzi 2010). Seen from this perspective, the patterning of weight-sensitive processes in Arrernte exhibits a principled conformity to a formal universal, once certain consonants are moraic. Attention to the statement of formal universals in this case thus provides a clear methodological guide to understanding the nature of variety in substantive properties. The universal in this case pertains to the set of data structures that learners use to encode sound patterns: moras, and only moras are the formal unit that can be referred to by weight-sensitive properties.

Before proceeding to the analysis of Arrernte, a bit of contextualization may be helpful. The possibility of VC syllabification in Australian languages was brought to the world's attention in a series of papers by Sommer (1970, 1981) on the language Oykangand, subsequently responded to by Dixon (1970), Darden (1971), and McCarthy & Prince (1986). Arrernte itself, analyzed as part of the Aranda family (Pama-Nyungan), has received generative analyses in Davis (1988) and Downing (1998), based on the description in Strehlow (1942). There is certainly some historical precedent for the impression of VC syllabification, based on looking at word-edges alone: Hale (1964) and Blevins (2001) have studied the widespread loss of initial consonants throughout Australian languages, attributing it to a shift of stress from the initial to the second syllable. While Arrernte was apparently no exception to this sweeping change, nonetheless, "25% of Arrernte words are pronounced in isolation with an initial consonant" (B&P: 2). In fact, to account for stems such as *mp*^{w}*aɻ* 'make' and *taŋk* 'be pleased', B&P have to propose that these words have an underlying hidden initial vowel, an initial red-flag to any clear demonstration that the language disallows consonantal onsets.

Beyond distributional observations about widespread onsetless word-initial sequences, the specific set of arguments that B&P propose in favor of VC-only syllables in Arrernte follows (as summarised in Hyman 2011b).

(1) *Arguments that have been used to support VC-only syllables in Arrernte:*
(a) all words begin with a vowel and end with a consonant
(b) stress is assigned within a word to the first nucleus that is preceded by a consonant
(c) allomorphy: for the plural suffix, C-initial monosyllables pattern with polysyllables
(d) would-be "onsets" do not undergo reduplication, which copies a VC(C) constituent
(e) Rabbit Talk (RT) transposes a VC(C) constituent

We will challenge each of these arguments throughout this paper, providing alternative analyses, evidence, and argumentation. Before turning to the analyses of individual phenomena, however, we wish to stress the main claim of the present paper: *Arrernte is a language with Onset Maximization like any other, but prefers moraic onsets over moraic codas, from which all other irregularities follow*. In fact, moraic codas seem to be banned (see the discussion in §4.1-4.2), and thus Arrernte may fill the gap in Hyman (2011b: 8), who asks whether "there is a language which has moraic onsets and (only) non-moraic codas". The Arrernte phenomena we discuss in the following sections, while perhaps unfamiliar in terms of their surface phonological typology, can in fact be straightforwardly understood in terms of a constellation of well-documented ingredients that have received empirical support from elsewhere, namely Prosodic Morphology (McCarthy & Prince 1986), moraic onsets (Topintzi 2010) and cross-anchoring (Itô, et al. 1996).

2. A puzzle: allomorphy in Arrernte and the need for moraic onsets

We begin with a discussion of a seemingly puzzling set of facts within the morphophonology of Arrernte, which displays allomorphy for the plural and the dual/reciprocal suffixes that is dependent on the shape of the base (2). Bisyllabic bases (2a), as well as C-initial monosyllables (2b) select the allomorphs *-əwaɾ* and *-iɾ* respectively, whereas V-initial monosyllables (2c) and sub-syllabic C-stems (2d) choose the allomorphs *-əɾiɾ* and *-əɾ* instead. Setting for the moment the case in (2d) aside, this distribution is rather surprising; one would expect that monosyllables as a whole would pick one allomorph and disyllables another one, instead of the situation highlighted here whereby monosyllables behave differently depending on whether they are C- or V-initial. Hyman (2011b), however, analyzes the allomorphy as syllable-based. Subscribing to the idea that Arrernte employs .VC. rather than .CV. syllables, he unites bisyllabic (2a) with – superficially – monosyllabic (2b), by suggesting that the forms in (2b) actually begin with *ə* (even if they are pronounced without it), a fact that renders them bisyllabic too. The same reasoning applies to (2d), rendering C-only stems monosyllabic alongside the true monosyllables in (2c). Monosyllables thus take the other allomorph.[1] The Arrernte data below are transcribed into IPA, accompanied by the forms in the orthographic system used in schools and by the government (see Appendix A for the correspondences between the two systems). The latter system – shown in italics – is only included the first time the data are presented; we do so here for the ease of readers familiar with a broad range of transcription conventions. In addition, note that all Arrernte words are spelled with a final *-e*, which we omit, since its presence is predictable at the phrase level, as explained later in this section.

(2) *Allomorphy data* (B&P; Henderson (2013: 211-212))[2]

	Base verb			*Plural*		*Dual /Reciprocal*		
a.	inŋel̪		*inngelh-*	inŋəl̪-əwaɾ	*inngelh-ewarr*	inŋəl̪-iɾ	*inngelh-irr*	'be like'
	ak^{w}əɳ		*akwern-*	ak^{w}əɳ-əwaɾ	*akwern-ewarr*	ak^{w}əɳ-iɾ	*akwern-irr*	'insert'
b.	mp^{w}aɻ	/əmp^{w}aɻ/	*empwar-*	mp^{w}aɻ-əwaɾ	*empwar-ewarr*	mp^{w}aɻ-iɾ	*empwar-irr*	'make'
	taŋk	/ətaŋk/	*etangk-*	taŋk-əwaɾ	*etangk-ewarr*	taŋk-iɾ	*etangk-irr*	'be pleased'
	k^{w}əɳ	/ək^{w}əɳ/	*ekwern-*	k^{w}əɳ-əwaɾ	*ekwern-ewarr*	k^{w}əɳ-iɾ	*ekwern-irr*	'swallow'
c.	aɻ		*ar-*	aɻ-əɾiɾ	*ar-errirr*	aɻ-əɾ	*ar-err*	'see'
	aŋk		*angk-*	aŋk-əɾiɾ	*angk-errirr*	aŋk-əɾ	*angk-err*	'talk'
	at̪		*ath-*	at̪-əɾiɾ	*ath-errirr*	at̪-əɾ	*ath-err*	'grind'
d.	t̪	/ət̪/	*eth-*	t̪-əɾiɾ	*eth-errirr*	t̪-əɾ	*eth-err*	'poke'

But what if the initial schwa of the bases in (2b&2d) is not there? Then the generalization that bisyllabic bases receive one type of allomorph and monosyllables the other one can no longer be upheld, since the bases in (2b) should now falsely pattern like the monosyllables of (2c) and not the bisyllables of (2a). An account that insists on .VC. syllabification throughout could then take one of two reasonable options: at best, it could employ an ad hoc mechanism that only admits onsets in truly C-initial words; at worst, it would apply no modification and simply accept failure to analyze these examples altogether. Either way, this approach provides no explanation for this pattern of allomorphy.

We propose instead that a uniform analysis of the allomorphy data dependent on the shape of the base is possible, but it relies on mora-counting rather than syllable-counting. A key component to our proposal however is the admission of moraic onsets (Topintzi 2010). We review the empirical evidence in favor of this position in Section 3 below. For the time being, it suffices to mention that we apply standard syllabification to Arrernte words, and assign a mora to each syllable onset. The assumed moraicity of the bases is shown in (3).

(3) *Allomorphy facts when initial ə is absent*

	Base verb	*Plural*	*Dual/Reciprocal*	
a.	i$_{\mu}$n.ŋ$_{\mu}$ə$_{\mu}$l̪	inŋəl̪-əwaɾ	inŋəl̪-iɾ	'be like'
	a$_{\mu}$.k$^{w}_{\mu}$ə$_{\mu}$ɳ	ak^{w}əɳ-əwaɾ	ak^{w}əɳ-iɾ	'insert'
b.	mp$^{w}_{\mu}$a$_{\mu}$ɻ	mp^{w}aɻ-əwaɾ	mp^{w}aɻ-iɾ	'make'

[1] In fact, the same analysis would be sustainable under a more ordinary syllabification that prefers CV over VC syllables. In this view, (3a-b) take the same allomorph because they are bisyllabic, e.g. [a.k^{w}əɳ] and [ə.taŋk], as opposed to shorter stems as in (3c-d).

[2] The data in (2)-(3) are all non-finite forms. When pronounced, these forms contain finiteness markers such as the PRESENT suffix *-əm*, which we omit following the practice of B&P. Thanks to Myf Turpin (pers. comm. October 2016) for pointing this out.

	t$_{\mu}$a$_{\mu}$ŋk	taŋk-əwaɾ	taŋk-iɾ	‘be pleased’
	k$^{w}_{\mu}$ə$_{\mu}$ɳ	k^{w}əɳ-əwaɾ	k^{w}əɳ-iɾ	‘swallow’
c.	a$_{\mu}$ɻ	aɻ-əɾiɾ	aɻ-əɾ	‘watch/see’
	a$_{\mu}$ŋk	aŋk-əɾiɾ	aŋk-əɾ	‘talk’
	a$_{\mu}$t̪	at̪-əɾiɾ	at̪-əɾ	‘grind’
d.	t̪	t̪-əɾiɾ	t̪-əɾ	‘poke’

All that must be said for this allomorphy is that bases of maximally one mora (3c-d) take the alternants *-əɾiɾ / -əɾ*. Longer bases (3a-b) take *-əwaɾ / -iɾ*.[3] Instrumental to this generalization is that onsets are moraic. Their moraicity is exactly what allows C-initial monosyllables (3b) to pattern differently from V-initial monosyllables (3c). We note that the form of the allomorphs themselves does not reflect a particular connection to the shape of their host, a pattern that is not unique to Arrernte; for example, in Axininca Campa, the genitive suffix is [-ni] after bimoraic stems, and [-ti] after longer stems (Bye 2008). While many cases of allomorphy are phonologically optimizing (see Nevins 2011 for an overview), there is no necessarily optimizing connection between the form of each allomorph and its environment, even though the latter can crucially rely on factors such as mora count. Importantly, the present mora-counting approach is advantageous compared to Hyman's (2011b) account. For the latter to work, it was critical that the bases in (3b) included an initial *ə*. Without it and under the non-moraic-onset-hypothesis that Hyman endorses, it seems impossible to group *3a&b* together vs. *3c&d*, contrary to fact.

In sum, if initial schwas are present in the structure, then either a mora- or, more plainly, a syllable-counting approach works without the need for contentious VC syllables. If, however, initial schwas are missing, then only a mora-counting account with moraic onsets proves workable. However, recall that the adoption of initial schwas is everything but uncontroversial. Wilkins' (1989) grammar of Arrertne explicitly adopts the hypothesis that surface C-initial words in Arrernte are indeed C-initial, without an underlying vowel before them. Evans (1995) and Tabain (2009) indicate the schwa's fleeting nature as observed by variable schwa-epenthesis. Evans comments that "[The word for] 'sits', for example, can be pronounced [anə́mə] (three syllables), [nə́mə] (two), [anə́m] (two), or [nəm] (one). Yet the number of VC syllables stays constant at two" (1995: 745). Presumably Evans is referring to different intonational and contextual environments, and his initial [a] is a variant of epenthetic schwa. Commenting directly on the same example, Henderson (2013: 269) asserts "The variation in the number of surface syllables is directly due to the treatment of word-initial and phrase-final [ə] at the phrase level". Initial schwas – and also final ones – should be taken to be the effect of a late, variable, intonationally-phrase-conditioned process.[4] On these final schwas, specifically, Tabain & Breen (2011:70) assert "The presence or absence of the final vowel is completely non-contrastive, and its presence is assumed to provide additional phonetic cues to consonant identity."

The status of initial (and final) schwas is rendered more complex by considering additional evidence, not present in the central arguments in B&P's paper, but which are brought to bear on this question from 'external evidence' presented in Kiparsky (2013). The first is from loanwords, where it is reported that for borrowings from English such as *paddock*, their adaptation into Arrernte as *parrike* suggests that initial schwas are not necessarily required, again casting doubt on the view that all apparent C-initial words have an underlying schwa at the beginning; instead a CV analysis looks compatible with these forms. Conversely, at the right edge of the word, there is evidence for a CV analysis as well: in connected phrasal speech, epenthesis is frequent in phrase-final position, although less so after nasal consonants (Pensalfini, pers. comm, Tabain & Breen 2011).

Some of the most revealing external evidence for truly CV-initial bias in Arrernte comes from

[3] Our discussion of Rabbit Talk justifies the non-moraicity of medial codas (§4.2). While we have no clear arguments regarding the moraicity or lack thereof of final codas, for simplicity, we maintain that they are non-moraic too. Even if final codas were moraic, the statement would be simply revised to: "maximally bimoraic bases take *-əɾiɾ / -əɾ*. Longer ones take *-əwaɾ / -iɾ*." Readers may be curious about verbs longer than two syllables. While Wilkins' (1989) glossary contains a number of C-initial polysyllabic verbs, they are all arguably morphologically complex verb-verb compounds, as in fact their allomorphy patterns are based on the second member of the compound (e.g. nthar-iw- 'to pinch-throw' takes the *-əɾiɾ* suffix as in (3c)).

[4] One could imagine that this schwa is actually a phonetic reflex (that can be missing on the surface) of an input mora, thus [mp] or [əmp] would map from underlying /$^{\mu}$mp/.

musicological evidence. In her work on Alyawarr (a closely related language) song traditions, Turpin (2015) documents the textsetting rule of taking a line-final consonant and moving it to the beginning of a vowel-initial word. This kind of ‘consonant transfer’ (Hale 1984) is widespread in Arandic songs as a way of providing an onset consonant from adjacent or line-final words in order to syllabify onsetless words as CV. As Turpin (2015:73) puts it, such a process serves to ‘create a full consonant+vowel syllable on the first (strong) beat of a bar’. In other words, a CV syllable in Arandic aligns with a strong beat, precisely because, as we will argue, CV syllables are heavy given moraic onsets. These phenomena not only point to the fact that initial schwas may not be required in all surface forms (and in fact, go a long way to support the hypothesis that CV syllabification, or a word-initial Onset constraint, manifest an emergence-of-the-unmarked effect; McCarthy & Prince 1994), but also that our CV-based analyses above are certainly compatible without invoking hidden leading schwas for every C-initial word.

For these reasons, in the remainder of the paper we propose that schwas-at-edges, and especially the initial schwas, are postlexically epenthetic and not phonologically active for the processes in question. On this view, the initial schwa, optionally added at the left edge in Arrernte, is no different from the stylistic prepausal final schwa optionally added at the right edge in ‘young’ Parisian French forms such as bonjour[ə] (Hansen 1997), for which there is no underlying or etymological source – and indeed in which *mer* ‘sea’ and *mère* ‘mother’ can both be optionally pronounced [mɛʁə]. The initial schwa found in Arrernte may be comparable to paragogic vowels found in prepausal position in Sardinian (Pittau 1991, Ladd pers.comm.), or more broadly, to non-phonological vocalic elements investigated in Ridouane & Fougeron’s (2011) study of Berber; indeed, the rich work emerging from such studies may be applicable to eventual phonetic studies of Arrernte’s fleeting schwas, a point to which we return in the conclusion.

Given the evidence surveyed above, however, we reject the analysis of all surface C-initial words as beginning with a hidden underlying schwa, and in what follows, reference to hidden schwas is made only for comparison between our pro-CV analysis and the B&P pro-VC account or to highlight any possible effects their admission would have had. We maintain that our analysis overall fares better than its contenders, as these founder as soon as initial schwas are no longer underlying, just as in the case of allomorphy just described. Elsewhere, the CV and moraic onset analysis remains equally successful.

3. Moraic onsets in stress, compensatory lengthening, and word-minimality

In a recent large-scale study, Topintzi (2010) has gathered evidence from stress, compensatory lengthening (CL), reduplication, and gemination from a wide range of languages that point towards the possibility of moraic onsets. The existence of metrical processes displaying sensitivity to onset consonants has been a topic of research for many years; see Davis (1988), Downing (1998), Everett and Everett (1984), Goedemans (1998), and Gordon (2005), who discuss onset-sensitivity of stress in languages ranging from English and Italian to Pirahã and Iowa-Oto. Bagemihl (1998) also entertains mora-bearing obstruents (as opposed to sonorants) for word-minimality effects in Bella Coola, as discussed further below.

An initial example comes from Pattani Malay (Yupho 1989; Abramson 1999, 2003; Hajek & Goedemans 2003; Topintzi 2008), which distinguishes between initial singleton versus geminate consonants as shown in (5). Geminates are usually the result of initial syllable or morpheme reduction, as in (4):

(4) *Initial syllable deletion varies with initial geminate in Pattani Malay* (Yupho 1989: 130)
buwi ~ wːi ‘give’
sɨdadu ~ dːadu ‘police’
pɨmatɔ ~ mːatɔ ‘jewelry’

In the examples in (4), the initial syllable deletes, generating a free variant whose original second onset now geminates. In the standard account that views compensatory lengthening as mora preservation (cf. Hayes 1989 and many since), this geminate is thus rendered moraic through compensatory lengthening, as a means to preserve the single mora of the originally initial syllable. As a result of processes giving way to (4), the language contrasts forms with initial singletons vs. geminates, as in (5):

(5) *Initial geminates vs. singletons in Pattani Malay* (examples from Abramson 1999, 2003, who provides extensive phonetic description of the acoustic cues employed to achieve this contrast):

Singletons		Geminates	
pagi	‘morning’	pːagi	‘early morning’
tido	‘to sleep’	tːido	‘put to sleep’
makɛ	‘to eat’	mːakɛ	‘to be eaten’
labɔ	‘to profit’	lːabɔ	‘cause to be late’
ɣatɔ	‘comprehensive’	ɣːatɔ	‘to spread out’
sɛpaʔ	‘to kick’	sːɛpaʔ	‘to be kicked’
cabɛ	‘branch’	cːabɛ	‘side road’
butɔ	‘blind’	bːutɔ	‘kind of tree’

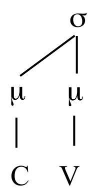
In addition to compensatory lengthening, the stress pattern of Pattani Malay corroborates the moraicity of the initial geminate. In words lacking geminates, primary stress is word-final, with all remaining syllables receiving secondary stress (6a), unless the syllable contains the schwa-like vowel [ɨ], in which case, it remains stressless (6b). However, in words starting with a geminate, primary stress appears initially with secondary stress later in the word (6c), even if the geminate is followed by the weak [ɨ] (6d). Assuming that Pattani stress is regulated through quality, i.e. avoidance of stress on low-sonority [ɨ], and quantity considerations, that is, heavy (i.e. geminate-containing) syllables receive stress, with the latter condition presiding over the former, then assignment of primary stress on the first (heavy) syllable rather than the default final is easily captured.

(6) *Stress in Pattani Malay* (Yupho 1989: 133-135)

a. ˌmãˌkɛˈnɛ̃	‘food’
b. bɨˌlaˈkɛ	‘back’
kɨˈda	‘shop’
c. ˈmːaˌtɔ	‘jewellery’
ˈɟːaˌlɛ	‘to walk’
d. ˈkːɨˌda	‘to the shop’

Initial geminates then in Pattani Malay are represented as moraic onsets in the manner shown in (7).

(7) *Moraic onsets*



Topintzi (2010) utilizes this representation for certain onsets in other languages as well, including Marshallese, Trukese, Pirahã, Karo and Bella Coola. We propose that onsets in Arrernte are likewise moraic, and thereby directly contribute to stress assignment and allomorph selection, as well as indirectly participating in a variety of other prosodic morphology processes.

While Pattani Malay shows that moraic onsets can arise in the case of geminates, this is not the only environment in which onsets contribute to weight. Moraic onsets can also be found depending on the sonority properties of the consonants involved. Consider the stress pattern of Pirahã, as analyzed in Topintzi (2010: 76), where voiceless obstruents pattern differently from all voiced consonants.

(8) *Pirahã weight scale*

PVV > BVV > VV > PV > BV [P=voiceless obstr.; B=voiced obstr. & sonorants]

Pirahã is a tonal language, and in the examples below, acute accent marks high tone, whereas boldface indicates the stressed syllable. See K. Everett (1998) for a discussion of the acoustic correlates of stress in Pirahã.

(9) *Pirahã examples* (E&E=Everett&Everett 1984, E=D. Everett 1988)

a. PVV > BVV

káo.bá.baí ‘almost fell’ [E88: 239]
pa.**hai**.bií ‘proper name’ [E&E: 708]
pii.bi.gái ‘deep water’ [D. Everett, pers comm.]

b. BVV > VV

bii.oá.ii ‘tired [literally: being without blood]’ [D. Everett, pers comm.]
poo.**gáí**.hi.aí “banana” [E&E: 709]

c. VV > PV

pia.hao.gi.so.**ai**.pi ‘cooking banana’ [E&E: 710]

d. PV > BV

ʔa.ba.gi ‘toucan’ [E&E: 710]
ti.**po**.gi ‘species of bird’ [E&E: 710]

e. rightmost heaviest stress

ʔa.ba.**pa** ‘Amapá’ (city name) *__ʔa__.ba.pa [E&E: 710]
ho.áo.**ii** ‘shotgun’ *ho.**áo**.ii [E&E: 710]
ti.**po**.gi ‘species of bird’ *__ti__.po.gi [E&E: 710]
paó.hoa.**hai** ‘anaconda’ *paó.**hoa**.hai/*__paó__.hoa.hai [E&E: 707]

The generalization about weight assignment and thereby stress placement is that it depends on three factors: long vs. short vowels (VV vs. V); presence vs. absence of onsets (CV vs. V), and quality of onsets (PV vs. BV). Note that there are no codas to potentially contribute to syllable weight. As evident above, moraic onsets in Pirahã are only the voiceless obstruents. As a result, this presents a case in which only a subset of onsets contribute to syllable weight.

In Arabela, as analyzed in Topintzi (2010: 85), obstruents are moraic regardless of voicing.[5] In this language, stress is rhythmic and normally creates trochees from left to right. The rightmost stress is the primary one. Degenerate monosyllabic feet are also admitted. However, Payne & Rich (1988) state that “if a word-final syllable that would have received stress has a voiced [=sonorant] onset, and the immediately preceding syllable has a voiceless onset, then the syllable with the voiceless onset is stressed”. Topintzi (2010) argues that this stress shift of primary stress is due to the moraicity of the voiceless onset in the penult.

(10) *Arabela stress* (Payne & Rich 1998)

a. *Normal pattern*

(tèna)(kári) ‘afternoon’
(sàma)(rú) ‘spirit’
(hùwa)(hàni)(yá) ‘peaceful’

b. *Stress shift*

(nòwa)(ʃì)(ʃáno) *(nòwa)(ʃìʃa)(nó) ‘brightened’
(sàpo)(hò)(sáno) *(sàpo)(hòsa)(nó) ‘deceived’
(mwèra)(tì)(tyénu) *(mwèra)(tìtye)(nú) ‘cause to be seen’

[5] More precisely, what Topintzi (2010) suggests is that *voiceless* obstruents are moraic as opposed to the non-moraic sonorant onsets. Voiced obstruents appear as variants of the voiceless obstruents under certain conditions (see Topintzi 2010: 85 and references therein), but do not arise phonemically.

Arabela has codas, but they are irrelevant for weight and stress purposes, e.g. *(mòkoʃ)(tyáka)* 'palm fruit' instead of **(mò)(kòʃ)(tyáka)*, which indicates that codas are not moraic. A similar, though more complex pattern is found in Umutina (Wetzels et al. 2014), where, if the last syllable contains a plosive onset, it will be stressed, and otherwise vowel sonority will decide between the last two syllables. Summarizing, we have thus far seen moraic onsets contributing to syllable weight when geminates (Pattani Malay), voiceless (Pirahã), or obstruent (Arabela).[6] Finally, English secondary stress, as argued by Davis (1988), can be sensitive to the presence of obstruent onsets. Consider the -*ative* suffix in American English. In (11a), the two bases have an identical stress pattern, but their derived forms differ in secondary stress. Only the forms with an obstruent onset attract secondary stress (11b) vs. (11c).

(11) *Sensitivity of secondary stress to obstruency of onsets* (Davis 1988)

a. ínnovate, óperate	ínnovàtive, óperative
b. invéstigàtive, quálitàtive	
c. nóminative, génerative, cúmulative	

The contribution of onsets to syllable weight is not restricted to the domain of stress assignment, however. Onsets may not only undergo, but also cause compensatory lengthening. Deletion of /r/ word-initially as well as in complex onset clusters produces lengthening of the following vowel in Samothraki Greek, as discussed in Topintzi (2010: 102-103) after Katsanis (1996).

(12) CL environment I: *Deletion of /r/ initially and lengthening* (Katsanis: 50-51)

ra > a:	rafts > á:fts	'tailor (masc.)'
ri > i:	ríɣaɲ > í:ɣaɲ	'oregano'
ru > u:	rúxa > ú:xa	'clothes'
re > e:	rema > é:ma	'stream'
ro > o:	róɣa > ó:ɣa	'nipple, berry (of a grape)'

(13) CL environment II: */r/ in onset CrVC: deletion and lengthening*

a. *biconsonantal clusters* (Katsanis: 54-56)

pr+o > po:	prótus > pó:tus	'first'
vr+i > vi:	vris' > ví:s'[7]	'tap'
fr+e > fe:	frena > fé:na	'brakes'
xr+o > xo:	xróma > xó:ma	'colour'
kr+a > ka:	krató > ka:tó	'I hold'
ɣr+a > ɣa:	ɣráfo > ɣá:fu	'I write'
θr+o > θo:	θrónos > θó:nus	'throne'
ðr+a > ða:	ðrákos > ðá:kus	'dragon'
br+e > be:	ɣabré > ɣabé:	'bridegroom'
dr+u > du:	ðédru > ðédu:	'tree'
tr+u > tu:	metrún > mitú:n	'they count'

b. *triconsonantal clusters* (Katsanis: 58-59)

spr+a > spa:	áspra > áspá:	'white'
xtr+a > xta:	éxtra > éxta:	'hostility'
zdr+u > zdu:	sidrufçá > zdu:fçá	'company (of people)'

[6] There are distinct possibilities for a sonority-based account of onset moraicity, where less sonorous onsets favor being moraic, versus the view adopted in Topintzi (2010) who argues for a voicing-based account. The differences arise based on the status of intermediate patterns (e.g. in no language do obstruent and nasal onsets pattern as moraic vs. non-moraic liquids and glides). In some patterns, like Karo (Gabas 1999; analyzed by Topintzi 2010 and Blumenfeld 2006), it seems to be voiceless obstruents and sonorants that pattern as moraic onsets vs. the voiced obstruents. However, English and Umutina, as discussed in the text, group all obstruents, voiced or voiceless, as moraic. Further exploration of whether moraic onset restrictions can be stated in terms of sonority-alone, voicing-alone, or whether both are needed is a task not taken up presently.

[7] The apostrophe expresses palatalisation due to an underlying /i/ that has deleted on the surface.

ftr+a > fta:	ráftra > á:fta:	'tailor (fem.)'

Potentially similar cases, usually subject to specific morphemes, are reported in Topintzi (2010: 205-6). These patterns demonstrate that the loss of an onset /r/ preserves its associated mora, transferring it to the vowel nucleus, and therefore that onsets can be moraic, paralleling the original arguments for moraicity of codas (Hayes 1989), can be found in codas. Conversely, onset lengthening as a response to CL occurs in Pattani Malay (as discussed above), as well as potentially Trique and Trukese.

We have reviewed the evidence for moraic onsets in terms of their contribution to stress assignment and compensatory lengthening. In Bella Coola (Bagemihl 1998), a bimoraic word minimum admits CV words, but not C-only or V-only words (Topintzi 2010: 146).

(14) *Minimal word-size in Bella Coola* (Bagemihl 1998)

Word Shape	Examples			
a. V	*			
b. VV	ya	'good'		
c. VC	n̩ƛ'	'dark, night'		
d. C	*			
e. CV / CR	ƛ'i	'fast' /	c'm̩	'index finger'
f. CC	tk'	'sticky'		
g. CCC	sχp	'to tie a knot'		
h. CCCC	p'χʷɬt	'bunchberries'		

Further discussion of the range of moraic onsets can be potentially found in Kpelle, which treats certain moraic onsets as tone-bearing units (Topintzi 2010: 207)

Recall now the central puzzle with which we started in Arrernte, armed with moraic onsets. One of B&P's most straightforward arguments about the nature of Arrernte syllabification comes from the process of stress assignment. Stress is assigned within a word to the first nucleus that is preceded by a consonant (B&P: 3), or in Evans' (1995) words, "In words of two or more VC-syllables, stress falls on the second VC syllable".

(15) *Stress in Arrernte with VC syllables* (syllabification after B&P: 3)

a. əmpʷ.áɻ.əm	*empwarem*	'is making'	əᵗn.əɳ.ɪɾ.əm	*etnyernirrem*	'are standing'[8]
b. ikʷ.əɲtʲ	*ikwenyty*	'policeman'	atʷ.əɾ.əm	*atwerrem*	'is fighting'

Syllabification of a VCV sequence on the other hand as either V.CV or VC.V is not directly accessible from the phonetic signal, and so the deduction of which syllabification pattern a particular utterance contains depends on particular phonological processes that refer to syllables. The behavior of different syllable types relies on their internal weight structure. As we develop an analysis of weight-sensitive stress in Arrernte onwards in this paper, we endorse (16) as a formal universal that regulates such patterns:

(16) Stress assignment, weight-sensitive allomorphy, compensatory lengthening and prosodic morphology, when sensitive to quantity distinctions in different syllable-types, refer to the representational unit of weight called the mora

The phonological universal in (16), developed by Hyman (1985), McCarthy & Prince (1986) and Hayes (1989), is formal, not substantive in nature: it delimits the nature of data structures that can be referred to by morphophonological processes. It is neutral on the substantive question of which segments can bear moras and does not constitute the only potential formal property to which these processes may refer; cf. sonority-driven stress, tone-stress interactions and so forth.[9]

[8] Henderson (1998/2013: 69) transcribes the present form of "stand" as: /əʈɳ-əm/.

[9] Further enrichments of models of syllable weight may be found in Broselow, Chen & Huffmann (1997), Curtis (2003), Gordon (2006), and Ryan (2014).

Of particular interest is the typology of mora-bearing consonants and the nature of the relation between onset moraicity and coda moraicity. In languages surveyed above such as Pirahã, voiceless onsets are argued to be moraic, while there are no coda consonants at all (and thus moraic codas are not a prerequisite for moraic onsets). Languages such as Pattani Malay have moraic onset geminates, but others, such as Arabela, have moraic onsets but not geminates – the kind of contrast also found in the coda position, where some languages have moraic codas that are singletons, and others have word-final geminates. Turning to the properties of the consonants themselves, in languages with moraic codas, either all consonants are moraic, or there can be a difference between sonorant consonants (moraic) and obstruents (non-moraic) (Zec 1995). Conversely, in the languages under discussion here, all onset consonants are moraic (as argued for Arrernte and Pattani Malay), or only a subset of them – typically voiceless obstruents or obstruents as a whole – are moraic (as found in English, Arabela, and Pirahã). In short, whether or not onsets are moraic (and which onsets are) in a given language involve largely the same variables as those determining coda moraicity.

4. Moraic onsets in Arrernte

In this section, we apply the moraic onset analysis to Arrernte. In Section 4.1, we demonstrate that stress assignment is straightforward if all word-initial onsets are moraic: stress falls on the leftmost heavy syllable. In Section 4.2, we investigate the transpositional language game Rabbit Talk, demonstrating that the pattern of what is transposed depends on moraic onsets as well. Finally, in Section 4.3, we examine reduplication patterns in Arrernte and propose that the selection of reduplicant material indirectly refers to mora-preservation of the segments involved.

4.1. Stress

To begin, consider the stress data in (17), collected from a range of sources. Descriptively speaking, C-initial words present initial stress (17a) and V-initial ones peninitial stress (17b).

(17) *Stress in Arrernte*

a.	mpʷáɻəm	*mpwarem*	‘is making’	B&P: 3
	ᵗnə́ɳirəm	*tnyernirrem*	‘are standing’	B&P: 3
	máʈəm	*martem*	‘is closing’	Wilkins 1989: 87
	kʷáʈ	*kwárt*	‘egg’	Turpin et al. 2014: 57
b.	ikʷə́ɲtʲ	*ikwenyty*	‘policeman’	B&P: 3
	atʷə́rəm	*atwerrem*	‘is fighting’	B&P: 3
	ipə́ʈ	*ipert*	‘hole’	Henderson 2013: 219
	a.ᵖmə́ɻ	*apmer*	‘camp, home’	Turpin et al. 2014: 60

A variety of proposals have been offered for Arrente stress assignment, including initial(-segment) extrametricality (Halle & Vergnaud 1987) and its variant of stress avoidance on vowel-initial words, as they are considered to be placed outside the prosodic word (Downing 1998); this is also the account adopted in Kiparsky (2013). Conversely, other accounts speak of an imperative to align the foot-edge or the stressed syllable with an onset (variants of this position appear in Goedemans 1998; Smith 2005; Hyde 2007; Topintzi 2010) or acknowledge the possibility that onsets pose their own requirements in stress placement rules (Davis 1988, in a grid-based framework). While all these proposals satisfactorily account for Arrernte stress, they fail to straightforwardly extend to other phenomena such as allomorphy, where, as exemplified in Section 2, the use of moraic onsets has proved indispensable in the absence of initial hidden schwas.

We argue that the moraic onset analysis unifies allomorphy with stress and other phenomena from unrelated domains of the grammar, to be shown below. What is more, it moves towards a demonstration that there are no longer any compelling data in favor of a VC.V analysis of Arrernte.

If vowels and onset consonants, and not coda consonants, are moraic in Arrernte (see also §4.2), the statement of stress assignment becomes quite straightforward in the light of moraic onsets (as shown in (7)): the leftmost bimoraic syllable receives stress (Davis 1988). In practice, this amounts to the first syllable in the word, if that is onsetful (18a), or to the second syllable if the first one constitutes a light onsetless syllable (18b).

(18) *Stress in Arrernte with moraic onsets* [1st heavy σ is underlined]

a.	mpʷá.ɻəm	'is making'	ᵗnə́.ɳi̪.ɾəm	'are standing'
	má.ʈəm	'is closing'	kʷáʈ	'egg'
b.	i.kʷə́ɲtʲ	'policeman'	a.tʷə́.ɾəm	'is fighting'
	i.pə́ʈ	'hole'	a.ᵖmə́ɻ	'camp, home'

The technical implementation of this pattern can be straightforwardly stated in a model like that of Topintzi (2010). Arrernte onsets are rendered moraic on the surface through the ranking WBYP-ONS >> *μ/ONS. In such a model, WBYP is a generic constraint that assigns weight to consonants, with WBYP-ONS being a specific version that assigns weight to onsets (and WBYP-CODA to codas), while *μ/ONS prohibits onsets from acquiring moras.

Given moraic onsets and the Weight-to-Stress principle, the Arrernte stress pattern can be captured as shown in two tableaux; the first displays the evaluation for light-initial words (19) and the second for heavy-initial words (20). Only candidates that fulfil the moraicity requirements are considered here, and the Weight-to-Stress-Principle (WSP) is categorically evaluated so that the lack of stress on a potentially trimoraic syllable like [ɾ$_\mu$e$_\mu$m$_\mu$] (not shown here) and the lack on a bimoraic one such as [tʷ$_\mu$e$_\mu$] are equally assessed.

(19) *stress on peninitial heavy in light-initial word*

		/atʷəɾəm/	WSP	ALIGN-L
☞	a.	a$_\mu$.tʷ$_\mu$ə́$_\mu$.ɾ$_\mu$ə$_\mu$m	*	*
	b.	á$_\mu$.tʷ$_\mu$ə$_\mu$.ɾ$_\mu$ə$_\mu$m	**!	
	c.	a$_\mu$.tʷ$_\mu$ə$_\mu$.ɾ$_\mu$ə́$_\mu$m	*	**!

(20) *stress on initial in heavy-initial word*

		/ᵗnəɳi̪ɾəm /	WSP	ALIGN-L
☞	a.	[ᵗn$_\mu$ə́$_\mu$.ɳ$_\mu$i̪$_\mu$.ɾ$_\mu$ə$_\mu$m]	**	
	b.	[ᵗn$_\mu$ə$_\mu$.ɳ$_\mu$í$_\mu$.ɾ$_\mu$ə$_\mu$m]	**	*!
	c.	[ᵗn$_\mu$ə$_\mu$.ɳ$_\mu$i̪$_\mu$.ɾ$_\mu$ə́$_\mu$m]	**	*!*

Recall that B&P instead assume that the words of type (18a) begin with an initial schwa and thus pattern with the truly V-initial ones of (18b), as shown in (21a) and (21b), respectively. Stress is then assigned within a word to the first nucleus that is preceded by a consonant (B&P: 3), or as Evans (1995) puts it, "In words of two or more VC-syllables, stress falls on the second VC syllable". In our account, no statement of that sort is required. Even if initial schwas were truly present, our moraic onset account would still predict stress on the first heavy syllable from the left, which would always happen to be the second in the word. (21) illustrates this possibility for concreteness.

(21) *Stress in Arrernte with moraic onsets* AND *initial schwas* [1st heavy σ is underlined]

a.	ə.mpʷá.ɻəm	'is making'	ə.ᵗnə́.ɳi̪.ɾəm	'are standing'
	ə.má.ʈəm	'is closing'	ə.kʷáʈ	'egg'
b.	i.kʷə́ɲtʲ	'policeman'	a.tʷə́.ɾəm	'is fighting'
	i.pə́ʈ	'hole'	a.ᵖmə́ɻ	'camp, home'

In summary, stress assignment in Arrernte is straightforward given moraic onsets which contribute to syllable weight and are independently found in conditioning prosodically-determined suffix allomorphy. We now turn to a more intricate data set, involving a transpositional language game.

4.2. Rabbit Talk data and CROSS-ANCHORING

Rabbit Talk (RT) “is a language game, that involves transposing the initial portion of a word to the end of the word, not unlike the Pig Latin of English” (B&P: 7). Transposition affects forms such as those in (23a-k), all of which appear as variants of a basic pattern that can be informally stated as follows.

(22) *RT informally stated*
Shift to the end of the word all the material up to and including the first moraic onset[10]

(23) *Rabbit talk data after B&P: 7-8* (adapted so that there are no initial schwas; see Appendix B for the original data)

		Arrernte		Rabbit Talk		
I.	a.	məɳ	*mern*	ɳəm	*rnem*	‘plant food’
	b.	kʷəɳətʲək	*kwernetyek*	ɳətʲəkəkʷ	*rnetyekekw*	‘to put in’
	c.	jan̪	*yanh*	an̪əj	*anhey*	‘there (nearby)’
	d.	lʲaʈ	*lyart*	aʈəlʲ	*artely*	‘now, today’
	e.	kəl	*kel*	lək	*lek*	‘right, OK’
	f.	n̪t̪əm	*nhthem*	mən̪t̪	*menhth*	‘giving’
II.	g.	itiɾəm	*itirrem*	iɾəmit	*irremit*	‘thinking’
	h.	aɾaŋkʷ	*arrangkw*	aŋkʷaɾ	*angkwarr*	‘no’
	i.	iŋʷən̪t̪	*ingwenhth*	n̪t̪iŋʷ	*nhthingw*	‘tomorrow’
III.	j.	ulkəʈ	*ulkert*	ʈulk	*rtulk*	‘perentie’
	k.	alpətʲək	*alpetyek*	tʲəkalp	*tyekalp*	‘to go back’
IV.	l.	aʈʷ	*artw*	jaʈʷ	*yartw*	‘initiated man’
	m.	iŋk	*ingk*	jiŋk	*yingk*	‘foot’
	n.	mp	*mp*	jəmp	*yemp*	‘come on’

The difference between Patterns I-III has to do with the exact location of the first onset. In Pattern I (23a-f), it is word-initial, in Pattern II, (23g-i) it is postvocalic, whereas in Pattern III (23j-k)), it is postconsonantal. The net effect is a divergence in terms of the amount of the transposed material, as explained below. Finally, when the base word is onsetless and monosyllabic, a different strategy is preferred. Instead of transion, an epenthetic [j-] is added word-initially (23l-n; Pattern IV). This pattern is discussed after the transposed cases have been examined in detail.

Crucially, and rather uncontroversially (e.g. Itô et al. 1996; Vogt 2009; Borowsky 2010), the transpositional game applies on fully-fledged surface words whose moraification and syllabification has already applied. Onset moraicity is critical in determining how much material will be transposed. This can be expressed by means of an alignment constraint specific to the transposition game (24).

(24) ALIGN-R (GAME, C_μ): The right edge of the game form must end in a moraic consonant

Some schematic transpositions are shown in (27). Moreover, the systematic lack of initial schwas is captured through the constraint ranking in (25), where we include the surface constraint *[ə, which is consistent with the explicit statement in Tabain et al. (2004) that “the consonant of an utterance-initial /əC/ syllable becomes onset at the surface to the following syllable due to deletion of initial schwa”. Of course, this is a word-level constraint, and we make no commitment here to its ranking at the phrase-level phonology, which as we have noted above, Henderson (2013: 269) argues to be the locus of variation for the appearance of initial schwa. Similarly, Pensalfini (1998) develops a different

[10] An anonymous reviewer points out that it could equally be the case that it is the material left over that could move to the *beginning* of the word. That is indeed true and there is no principled reason to choose between the two options. However, it can be argued that the proposal in (27) is conceptually simpler since it involves the same material being chosen and transposed at the end of the word. In the alternative, some material word-initially has to be selected and then the remainder, i.e. a different sequence, must shift to the beginning.

word-level and phrase-level analysis for Arrernte whose rankings differ precisely in their treatment of initial and final schwa.

The ranking in (26) is intended to express the fact that a schwa will nonetheless be occasionally inserted word-medially to e.g. avoid the lack of a nucleus. We also assume that there are specific phonotactic considerations that regulate which clusters are acceptable and which are not. For example, [lk] in [ʈulk] is fine and does not induce *ə*-insertion. Since our focus is not on an exhaustive set of constraints governing allowable clusters, we refer to these conditions collectively as PHONOTACTICS (PT) in (26), constraints impelling schwa-insertion within clusters as in (27a-c).

(25)	*[ə >> MAX-BG[11]	*[ə: No word-initial schwas MAX-B(ASE)G(AME): No deletion between base and game
(26)	PT >> DEP-ə-BG	PT: collective phonotactic constraint for ‘Have nucleus’, ‘Avoid certain medial clusters’, and related restrictions DEP-ə-B(ASE)G(AME): No insertion of *ə* between base and game

(27) *Generation of RT forms schematically* [transposed material in grey; only moraicity of 1st onset shown]

a. m$_\mu$əɳ → *əɳm [*due to (25)*], **ɳm* [*due to (26)*]; *hence* ɳəm$_\mu$ — Pattern I: #CV…
b. k$^w_\mu$əɳət^{j}ak → *əɳət^{j}akk^w [*due to (25)-(26)*]; *hence* ɳət^{j}akək$^w_\mu$ — Pattern I: #CV…
c. n̪t̪$_\mu$əm → *əmən̪t̪ [*due to (25)*], *mn̪t̪ [*due to (26)*]; *hence* mən̪t̪$_\mu$ — Pattern I: #CV…
d. it$_\mu$iɾəm → iɾəmit$_\mu$ — Pattern II: #VCV…
e. ulk$_\mu$əʈ → *əʈulk [*due to (25)*]; *hence* ʈulk$_\mu$ — Pattern III: #VCCV…

In (27a), a shift of [m] would create *[ənm], a form avoided since it contains an initial schwa (25). [ɳm] is impossible, as it lacks a nucleus, and thus a medial schwa is inserted. Similarly, a shift of the onset in (27b) produces a sequence *kk^w* that violates phonotactics, therefore a schwa must be inserted to provide a nucleus. As usual, the initial schwa is also deleted. (27c) works in roughly the same way, however [nt] is treated as a complex consonant under the onset node. The form *[əmən̪t̪] cannot survive due to (25), leaving [mən̪t̪], again resulting in insertion of a medial schwa. (27d-e) demonstrate that longer material besides a single onset is transposed in V-initial polysyllables. In particular, a VC- sequence is transposed in (27d) vs. a VCC in (27e). Data such as these had led B&P to propose the literal transposition of the first VC(C) syllable of the word at its end. We argue instead that this is merely a superficial side-effect of the moraic onset transposition.

We proceed towards formally expressing the mechanism of the transposition itself; effectively, as we have seen, in RT, *strings* at the left edge of the base word are transposed at the right. To capture this effect, we propose to employ the CROSS-ANCHOR constraint that Itô et al. (1996) introduce for a Japanese Argot game.[12] The important intuition this captures is that “reversal does not affect single edge segments, but rather edge strings” and this requires a modification of correspondence “from the level of elements to the level of strings, such that a string *x* can have a string *y* as its (string-) correspondent” (Itô et al. 1996: 247). Let a string *x* begin/end a string *S*, iff *x* is an initial/final substring of *S*. Consequently, the *beginnings* of *S* are the set of strings that begin *S* and the *endings* of *S* are the set of strings that end *S* (ibid.: 250). By means of illustration, (28) [Itô et al.’s example (74)] presents the various substrings that can serve as beginnings or endings for the word *karaoke*. In the argot game, *karaoke* surfaces as *okekara*.

(28)

$$\text{beginnings (karaoke)} = \left\{ \begin{array}{l} \varnothing \\ \text{k} \\ \text{ka} \\ \text{kar} \\ \text{kara} \\ \text{karao} \\ \text{karaok} \\ \text{karaoke} \end{array} \right\} \quad \left\{ \begin{array}{r} \varnothing \\ \text{e} \\ \text{ke} \\ \text{oke} \\ \text{aoke} \\ \text{raoke} \\ \text{araoke} \\ \text{karaoke} \end{array} \right\} = \text{endings (karaoke)}$$

[11] B(ase)G(ame) is inspired by Itô et al. (1996), who refer to this correspondence as a Base-Argot one, since they focus on a Japanese Argot game. To expand on its generality, we replace *Argot* with *G(ame)* and refer to B(ase)G(ame) relationships.
[12] B&P (17-18) also utilize a version of CROSS-ANCHOR in their brief examination of alternative accounts to their own.

What reversal establishes is that some string *x*∈*beginnings* of the Base corresponds to some string *x'*∈*endings* of the Game, and conversely, some string *y*∈*endings* of the Base corresponds to some string *y'*∈*beginnings* of the Game, as schematized in (29).

(29) String reversal (cf. ibid: example (75))

Base: [x . . . . y]

Game: [y' . . . x']

Formally, this requirement is implemented through the CROSS-ANCHOR constraint (ibid.: 251).

(30) *CROSS-ANCHOR Definition*

A game form G is *cross-anchored* to a base B iff there exist strings *x, y, x', y'* such that

(i) $x \in$ beginnings(B), $y \in$ endings(B), $x' \in$ endings(G), $y' \in$ beginnings(G)

(ii) $x \, \Re_s \, x'$ and $y \, \Re_s \, y'$, where $\Re_s$ refers to the correspondence relation of strings

(iii) $x', y' \neq G \,;\; x', y' \neq \varnothing$

With this much background, we return to the Arrernte data. The basic pattern (Pattern I) involves the interplay between CROSS-ANCHOR and the alignment constraint (24), specific to the game. Finally, a Base-Game (BG) version of the anti-metathesis LINEARITY constraint ensures that no excessive material will be transposed.

The following tableaux maintain the grey-shading vs. lack thereof in candidates as a means to indicate the corresponding cross-anchored strings. The input is taken to be the base form upon which RT is formed. The Base, as explained before, is already syllabified and moraified, along the lines sketched in §4.1. Candidates are presented in pairs, with the Base on the left and the RT candidates on the right. We first consider Pattern II, as this requires the bare minimum of the constraints introduced. Pattern II transposes what looks like a VC sequence.

(31) ***Pattern II***: [itiɾəm]$_B$ – [iɾəmit]$_G$

[$i_{\mu}.t_{\mu}i_{\mu}.ɾ_{\mu}ə_{\mu}m$]			CROSS-ANCHOR	ALIGN-R (GAME, C_{μ})	LIN-BG
☞ a.	itiɾəm	iɾəmit			iɾəm < **it**
b.	itiɾəm	tiɾəmi		*!	tiɾəm < **i**
c.	itiɾəm	itiɾəm	*!		
d.	itiɾəm	əmitiɾ			əm < **itiɾ** !

Candidate (31c) is quickly ruled out as it fails to satisfy CROSS-ANCHORING, unlike the winner (31a), whereby the correspondence between *Base* [itiɾəm] and *RT/Game* [iɾəmit] satisfies CROSS-ANCHORING. But the remaining pairs are also compatible with it, e.g. { itiɾəm - tiɾəmi } or { itiɾəm - əmitiɾ }. However, (31b) can be excluded on the grounds that no RT-output forms are ever V-final, under the assumption that the output forms in RT are regulated by the high-ranking ALIGN-R (GAME, C_{μ}). The second possibility (31d), incurs unnecessary LINEARITY violations.

(32) LINEARITY-BG: Base segments should not be misplaced from their original position to a different position in the game form

LINEARITY is not computed over strings, rather over segments, as traditionally implemented. In particular, it counts how many segments have been misplaced / metathesized from their original position (reading this from L-to-R). Violations of LINEARITY are placed in bold. In (31a) two segments have shifted position, while in (31d) four have.

Pattern I exemplifies the effect of the rankings in (25) and (26), with the deletion of an initial schwa in the game form and the addition of an epenthetic schwa, respectively. The latter is warranted by the requirement to provide syllable nuclei (33) or to break up unacceptable clusters as in (23c-d). In the example below, the inserted-schwa is shown in bold and ALIGN-R (GAME, C$_\mu$) is omitted.

(33) ***Pattern I:*** [məɳ]$_B$ – [ɳ**ə**m]$_G$

[m$_\mu$ə$_\mu$ɳ]				CROSS-ANCHOR	*[ə	PT	DEP-ə-BG	MAX-BG
☞	a.	m$_\mu$ə$_\mu$ɳ	ɳ$_\mu$**ə**$_\mu$m$_\mu$				*	*
	b.	m$_\mu$ə$_\mu$ɳ	ɳ$_\mu$m$_\mu$			*!		
	c.	m$_\mu$ə$_\mu$ɳ	m$_\mu$ə$_\mu$ɳ	*!				
	d.	m$_\mu$ə$_\mu$ɳ	ə$_\mu$ɳ$_\mu$m$_\mu$		*!			

Consideration of Pattern III highlights the importance of employing moraic onsets. All the candidates below satisfy CROSS-ANCHOR, which is henceforth omitted. The Base [alpət^{j}ək]$_B$ for example corresponds to the actual RT winner [t^{j}əkalp]$_G$ (35b) instead of the possible [alpət^{j}ək]$_B$ – [pət^{j}əkal]$_G$ pairing (35d). The latter form transposes less material and consequently should be preferred, because it would incur fewer violations of LIN-BG. It is crucial that the RT alignment constraint requires that the forms end in a *moraic* consonant, rather than simply a consonant. Alternatively, consider an RT form, virtually identical to (35d), but with insertion of a mora on the final C (cf. (35c)). This is ruled out by another constraint, DEP-C$_\mu$(BG), that bans insertion of consonantal moras on the RT form that were not present in the first place in the Base. In fact, all RT forms violate this constraint once, since the Base *k* acquires an onset in the RT form by virtue of being an onset. It is just that (35c) additionally assigns a mora on the final *l* in the RT form. Finally, candidate (35a) loses early on due to violation of dominant *[ə; conversely, (35e) escapes violation of the high-ranking constraints, but eventually fails, as LIN-BG penalizes it more severely due to superfluous transposition.

(34) DEP-C$_\mu$(BG): No insertion of moras on consonants between the base and the game forms

(35) ***Pattern III:*** [alpət^{j}ək]$_B$ – [t^{j}əkalp]$_G$

[a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k]				*[ə	ALIGN-R (GAME, C$_\mu$)	DEP-C$_\mu$(BG)	MAX-BG	LIN-BG
	a.	a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k	ə$_\mu$t$^j_\mu$ə$_\mu$.k$_\mu$a$_\mu$lp$_\mu$	*!		*		ət^{j}ək < **alp**
☞	b.	a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k	t$^j_\mu$ə$_\mu$.k$_\mu$a$_\mu$lp$_\mu$			*	*	t^{j}ək < **alp**
	c.	a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k	p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$.k$_\mu$a$_\mu$l$_\mu$			**!	*	pət^{j}ək < **al**
	d.	a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k	p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$.k$_\mu$a$_\mu$l		*!	*	*	pət^{j}ək < **al**
	e.	a$_\mu$l.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k	k$_\mu$a$_\mu$l.p$_\mu$ə$_\mu$t$^j_\mu$			*	*	k < **alpə!t^j**

Pattern III thus transposes what looks like a [VCC] sequence. Such a string is not a syllabic constituent, but rather the emergent result of high-ranking constraints. It should be noted that this analysis rests on the assumption that medial-codas in Arrernte are non-moraic, but onsets are. Had medial codas been moraic, then we would predict an RT pairing [a$_\mu$l$_\mu$.p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$k] – [p$_\mu$ə$_\mu$.t$^j_\mu$ə$_\mu$.k$_\mu$a$_\mu$l$_\mu$], which would fare better than the actual winner (35b), by virtue of better constraint satisfaction overall. On the other hand, it should be pointed out that the analysis as shown in (35) works equally well whether the final codas are moraic or not, which would mean that Arrernte either has a version of high-ranking WBYP-CODA relativized to final codas that only admits final moraic codas or ranks WBYP-CODA low overall, respectively. To maintain the most straightforward analytic option, we have considered all Arrernte codas as non-moraic.

We can now turn to Pattern IV (23l-n) which affects onsetless monosyllables and involves no transposition. Due to the absence of an onset, RT cannot actively enforce any change in the form. Assuming further that the output of the RT-formation cannot be the null parse (a ban formalized here in terms of MPARSE), we include a high-ranking constraint requiring that the exponent of RT is non-identical to the base form it is associated with (calling this RT≠BASE). This makes intuitive sense, since the output of the language game, essentially a disguise form, *has* to be somehow different from the word it derives from. A solution of last resort then is the addition of *j*-.

The reasoning behind the lack of transposition is particularly evident in examples such as [aʈ^w]→ [jaʈ^w] instead of *[ʈ^{w}a] (23l). The RT alignment constraint would promote the candidate [aʈ^w] itself, but that would violate both CROSS-ANCHORING as well as RT≠BASE (and DEP-C$_\mu$(BG) if final codas are non-moraic, as assumed throughout). Epenthesis is thus favored. The lack of transposition however is harder to see in other cases, especially if final codas *are* after all moraic. Consider for instance the mapping [iŋk]$_B$ - [jiŋk]$_G$ (23m). What precludes the RT-form *[kiŋ] from arising? We claim that the crucial observation relates to the nature of the clusters involved. We take homorganic [ŋk] or [mp] to behave as (partial) geminates, hence resistant to splitting, as reflected by GEMINATE INTEGRITY.[13] In turn, this suggests that onsetless monosyllabic words containing other final clusters might work differently. A form that would be interesting to test is the word [alp] "return-IMP" (Henderson 2013: 48). While this appears later too in Henderson's discussion of RT (p.267), in that instance, it emerges suffixed, thus no longer suitable to answer this empirical question. Other monosyllabic examples with non-homorganic clusters would further refine this aspect of the analysis. The situation just outlined is exemplified in the tableau below.

(36) ***Pattern IV:*** [iŋk]$_B$ – [jiŋk]$_G$

[i$_\mu$ŋk]			RT≠BASE	MPARSE	GEMINATE INTEGRITY	CROSS-ANCHOR	DEP-BG
☞ a.	i$_\mu$ŋk	j$_\mu$i$_\mu$ŋk				*	*
b.	i$_\mu$ŋk	i$_\mu$ŋk	*!			*	
c.	i$_\mu$ŋk	∅		*!			
d.	i$_\mu$ŋk	ki$_\mu$ŋ$_\mu$			*!		

Notably, the replacement of transpositio, with glide-epenthesis when the former is not possible meshes well with typological preferences. Glide-insertion is very common cross-linguistically, although rare word-initially, where ʔ-insertion is preferred (Uffmann 2007). Still, the choice of *j* is justified; Arrernte lacks *ʔ* (B&P: 20). Uffmann (2007: 458) suggests that "glides are inserted to minimise the contrast to the following or preceding vowel" (in terms of consonant-vowel transitions) so arguably *j*-insertion is the most cost-effective and least intrusive way of ensuring satisfaction of RT through disguise, while at the same time allowing the base form to be easily recoverable.

Before concluding this section, we briefly consider various ways of capturing the RT data. The intuition pursued here is that Rabbit Talk transposes everything up to the first moraic onset. Given the work on prosodic morphology pivots in works such as Nevins & Vaux (2003) and Yu (2007), statements such as 'everything before the second vowel' are not crosslinguistically recurrent and arguably not expressible in the formal vocabulary of such operations. Thus, to attempt to recast Rabbit Talk without reference to moraic onsets might be possible, our proposal develops a specific analysis of the transposition operation in terms of extant constraints operative in other language games and in terms of a target that ends in a moraic consonant.

4.3 Reduplication patterns in Arrernte

Arrernte displays multiple instances of reduplication (Henderson 2013: 227-264), some of which include so-called linker morphs, e.g. /-əlp/ (Inkelas & Zoll 2005: §2.2.3). The most productive

[13] See Appendix C for the clusters Arrernte permits according to Henderson (2013). GEMINATE INTEGRITY refers to the finding that geminates resist splitting under processes such epenthesis (Hayes 1986) or reduplication (Hume et al. 1997). Arrernte manifests an additional pattern, whereby (partial) geminates maintain their integrity under transposition.

instances of reduplication involve the attenuative, the iterative and the frequentive, a few examples of which appear in (37).[14]

(37) *Arrernte Reduplication* (underlying forms according to Henderson 2013: 233-238)

root-PRES	i. Frequentive	ii. Attenuative	iii. Iterative
	/root-FREQ-PRES/	/ATTEN-root-PRES/	/ITER-root-PRES/
a. /ət̪-əm/	/ə.t̪-ə.p-**ə.t̪**-əm/	/**ə.t̪**-əl.p-ə.t̪-əm/	
eth-em	*eth-ep-eth-em*	*eth-elp-eth-em*	
'poke- PRES'			
b. /əmp^{w}aɻ-əm/			/**əmp^{w}aɻ**-əmp^{w}aɻ-əm/
empwar-em			*empwar-empwar-em*
'make-PRES'			'do things repeatedly in quick succession'

Henderson (2013) proposes that *descriptively* the reduplicant includes a VC-string (or two sequences of VC-strings), a fact which he views as copying of VC-*syllables* (p.227). Thus, "the frequentive is marked by a suffix which consists of /-əp/ followed by a monosyllabic reduplicant which copies from the end of the preceding stem" (p. 237, cf. (37i)), "the attenuative is marked by a disyllabic form which consists of a monosyllabic template followed by /-əlp/" (p. 232, cf. (37ii)), while iterative reduplication "involves a disyllabic reduplicant template which precedes and compounds with a verb word" (p. 235, cf. (37iii)).

What should make us wary of this interpretation of the facts, however, is in fact Henderson's own admission that a form such as 'ATTEN-poke-PRES' emerges on the surface as [**t̪**-əlp-ət̪-əm] (p.233), effectively acknowledging the lack of initial schwas. His assumption that reduplication is formed at a level where the initial ə of the root is present seems to be driven more by analytical than by empirical choices. Our account on the other hand, which rejects initial underlying schwas, is consequently much closer to the actual pronounced forms. The data in (37) are accordingly repeated in (38) alongside more examples.

(38) *Arrernte Reduplication* (SRs after Henderson 2013; URs according to our proposal)

root-PRES	i. Frequentive	ii. Attenuative	iii. Iterative
	/root-FREQ-PRES/	/ATTEN-root-PRES/	/ITER-root-PRES/
a./t̪-əm/	[t̪-əp-**ət̪**-əm]	[**t̪**-əlp-ət̪-əm]	
th-em	*th-ep-eth-em*	*eth-elp-eth-em*	
'poke- PRES'			
b. /mp^{w}aɻ-əm/			[**mp^{w}aɻ**-əmp^{w}aɻ-əm]
mpwar-em			*mpwar-empwar-em*
'make-PRES'			'do things repeatedly in quick succession'
c. /at̪-əm/		[**at̪**-əlp-at̪-əm]	
ath-em		*ath-elp-ath-em*	
'grind-PRES'			
d. /aɲc-əm/		[**aɲc**-əlp-aɲc-əm]	
anty-em		*anty-elp-anty-em*	
'climb-PRES'			

A common feature in the forms above is that the reduplicant itself transcends syllable boundaries, as in [**mp^{w}a.ɻ**-ə.mp^{w}a.ɻ-əm]. Comparable data are found with reduplication in Oykangand, a Cape York language of Australia.

(39) *Oykangand VC-prefixing reduplication* (Sommer 1981)

/eder/	[**ed**-eder]	'rain'
/algal/	[**alg**-algal]	'straight'

[14] Here and below, the reduplicant is placed in bold and linker morphs are highlighted in grey; note that the later suffixation of the present suffix [-əm] is irrelevant and hence we omit it. Where Henderson 2013 does not include full glosses in cases of the reduplicated forms we state the morphemes involved, and include glosses when explicitly provided in the original source.

/igu-/ [**ig**-igun] 'go'

Given the VC(C) output of reduplication, Sommer (1981) has syllabified Oykangand words as *alg.al* instead of *al.gal*, on the arguably flawed assumption that the reduplicated portion ***alg*** in ***alg**-algal* should correspond to a constituent. In fact, this proposal underlies Henderson's (2013) analysis of VC-syllables in Arrernte. But it has long been known that reduplication does not copy a constituent; rather, it aims to fill a particular templatic target (McCarthy & Prince 1986, 1993). For example, in Mokilese, reduplication must result in a prefix which is a heavy syllable, and to achieve this, will fill up the heavy syllable template in any way it can, either copying a long vowel intact if there is one (40a), copying an onset consonant into coda position to create a heavy syllable (40b), or lengthening a vowel if there is no other source of heaviness to recruit (40c) (McCarthy & Prince 1986).

(40) *Mokilese Reduplication*

a.	/kookɔ/	[**koo**-kookɔ]	'grind coconut'
b.	/wadek/	[**wad**-wadek]	'read'
	/poki/	[**pok**-poki]	'beat'
c.	/pa/	[**paa**-pa]	'weave'

McCarthy & Prince (1986, 1993) present an analysis of Oykangand reduplication in which ***alg**-algal* involves copying a contiguous string of segments to syllabic positions, so that ***al*** is mapped as a core CV-syllable to the prefixal reduplicant position, but in which the extra /g/ is also retained in the copied material ***al.g***. This apparent "overcopying" of an extra segment, according to McCarthy & Prince (1986), occurs precisely to provide an onset for the following syllable, but it also serves an additional purpose: satisfaction of the reduplication constraint STROLE (structural role in McCarthy & Prince 1993, see also Bennett 2015), which ensures matching syllable roles between correspondents in base and reduplicant. In the reduplicated form ***al.g**al.gal*, both members of the cluster [lg] present correspondence to the syllable roles they held in the base. This is not the case in hypothetical ****a.l**al.gal* where onset satisfaction is achieved but STROLE is not, since [l], originally a coda in the base, emerges as an onset in the reduplicant. Alternatively, one might draw parallels with maximization of clusters in correspondence, as shown in the formation of German hypocoristics such as *Gor.ba.chov* → *Gorbi* (Itô & Mester 1997), and in English 'millenial' abbreviations such as *presumptuous* → *presumpsh* (Spradlin 2015). In these truncatory formations, correspondence aims to copy as much as possible of a consonant cluster, even when transcending syllable boundaries. Such data would not be brought to bear on an argument of German or English as VCC languages; rather, the prosodic morphology processes in question copy more than 'is needed'.

We claim that a similar "trans-junctural syllabification in reduplication contexts" (McCarthy & Prince 1993: 134) applies to Arrernte.[15] Prosodically speaking then, the template for the frequentive and attenuative looks like a mora, whereas for the iterative it is a syllable. The monomoraic template is demonstrated more clearly by examples such as [**t̪**-əl.p-ə.t̪-əm] (38a.ii), where the reduplicant is merely a moraic onset. The same in fact may be said for (38a.i), where the preceding schwa is epenthetic to satisfy phonotactic considerations (namely, the presence of a nucleus).

Let us start with an analysis of the attenuative form [**t̪**-əl.p-ə.t̪-əm], which fills the reduplicative template of a single mora and requires correspondence between the left edges of base and reduplicant (41).[16] Note that we assume throughout that the linker morph is not part of the reduplicated portion and thus is ignored for the purposes of FAITH-BR; any markedness violations, however, are computed as usual. The fully faithful candidate (42d) is eliminated due to violation of high-ranking phonotactics of Arrernte (26), which bans sequences like *[lpt̪], and others similar to it. For this reason, an epenthetic schwa appears in the base to break up the cluster between *p* and *t* (cf. our analysis of medial epenthetic schwas in (33) above); if the schwa is also copied in the reduplicant, as in (42b-c), L-ANCHOR-BR – defined in (41) – is satisfied, but at the cost of the top-ranking *[ə constraint (cf.

[15] Comparable cases exist in other Australian languages, such as Kuuk Thaayorre, Mangarrayi, Jingulu, Kuku Nganhcara which alternate between CV-/VC- reduplication (Round 2013), but also beyond Australia in Tzeltal and even in Mokilese itself, as in [an.d-an.dip] 'spit'(McCarthy & Prince 1986: 18).

[16] The same analysis can apply to the frequentive, whose formation is near-identical to that of the attenuative.

(25)). The winning form (42a) merely consists of a reduplicant [**t̪**], i.e. a moraic onset. Its failure to conform to left-anchoring requirements (since the element initial in the reduplicant, i.e. *t̪* at the left edge does not match the epenthetic schwa at the left edge of the base) nonetheless permits it to fully satisfy the dominant phonotactic constraints of Arrernte.

(41) L-ANCHOR-BR: The left edge of the base and the left edge of the reduplicant are in correspondence (McCarthy & Prince 1993, 1995)

(42) *Arrernte Attenuative*: [**t̪**əl.pə.t̪əm] 'ATTEN-poke-PRES'

	/RED-əlp-t̪-əm /	*[ə	PT	L-ANCHOR-BR	ONSET	RED=μ
☞	a. **t̪**-əl.p-ə.t̪-əm			*		
	b. **ə.t̪**-əl.p-ə.t̪-əm	*!			*	*
	c. **ə.**-əl.p-ə.t̪-əm	*!			**	
	d. **t̪**-əl.p-t̪-əm		*!			

This analysis would be further corroborated if a single moraic onset were to arise in reduplication of forms such as C-initial disyllabic roots, but this unfortunately proves hard to test. As mentioned in footnote 3, forms such as /nthariw/ 'to pinch-throw' which seemingly fulfil the desired profile are arguably morphologically complex. Thus /nthar-iw/ is treated as a compound verb, whose first part is difficult to identify, but whose second part – here /iw-/ 'to throw away' – corresponds to a free verb form (Wilkins 1989: 266-7). More generally, Wilkins (pers. comm. June 2016) points out that Arrernte seems to lack C-initial polysyllabic verb roots that are unambiguously simplex. With this caveat in mind, it is notable that a polysyllabic, complex stem such as [n̪t̪ak-iw-] *nthakiwe* 'to char fur/scales off the game' forms the attenuative in precisely the manner expected, i.e. [**n̪t̪**-əlp-ən̪t̪akiw-əm] *nthelpenthakiweme* (Wilkins 1989: 267 and pers. comm, June 2016; forms in italics as provided by Wilkins), with mere copying of the first moraic onset.

The importance of L-ANCHOR-BR and ONSET become apparent with roots that are underlyingly vowel-initial. In such cases, no epenthetic schwas appear in the base. These forms demonstrate the effect of lower-ranked, but still vital, constraints in the phonology of Arrernte reduplication.

(43) *Arrernte Attenuative*: [**a.t̪**əl.pa.t̪-əm] 'ATTEN-grind-PRES'

	/ RED-əlp-at̪-əm /	*[ə	PT	L-ANCHOR-BR	ONSET	RED=μ
☞	a. **a.t̪**-əl.p-a.t̪-əm					*
	b. **a.**-əlp-at̪-əm				*!	
	c. **t̪**-əl.p-a.t̪-əm			*!		

The winner (43a) fully satisfies L-ANCHOR-BR, unlike the contender in (43c), a form comparable to the winner in (42). Candidate (43b) on the other hand exactly fills the templatic constraint (RED=μ) through copying of a single vocalic mora. Nonetheless, the winner presents extension of the reduplicant by means of *overcopying* of [**t̪**]. We argue that such overcopying is motivated by ONSET satisfaction, while overcopying of clusters such as [**ɲc**] in (38d.ii) serves to also satisfy STROLE. These points are illustrated below.

Undercopying, as illustrated by the iterative form [**mpʷa.ɻ**-əmpʷaɻ-əm] (44a) compared to [**əmpʷa.ɻ**-əmpʷaɻ-əm] (44b), is also motivated by general markedness principles of the language, specifically the avoidance of initial schwas. Candidate (44c) is ruled out on phonotactic grounds, resting on the assumption that the cluster [ɻ.mpʷ] is illegal, and thus an epenthetic schwa must break it up. If the cluster is after all deemed acceptable, then the elimination of (44c) becomes more difficult to be accounted for. A solution can be given by the constraint STROLE, which we claim is independently needed in light of examples such as (45). In particular, if STROLE >> L-ANCHOR-BR, then (44c) will be dispreferred not due to phonotactic violations, but because it assigns different syllable roles to the [ɻ] in the reduplicant and base, namely coda vs. onset, respectively, a fact that (44a) with schwa-insertion manages to avoid. The exclamation marks in (44c) mirror these two possible analyses.

(44) *Arrernte Iterative*: [**mp^{w}a..ɻ**ə.mp^{w}a..ɻəm] ‘ITER-make-PRES’

/RED-mp^{w}aɻ-əm /	*[ə	PT	STROLE	L-ANCHOR-BR	ONSET
☞ a. **mp^{w}a..ɻ**-ə.mp^{w}a..ɻ-əm				*	
b. **ə.mp^{w}a..ɻ**-ə.mp^{w}a..ɻ-əm	*!				*
c. **mp^{w}aɻ.**-mp^{w}a..ɻ-əm		*!	*!		

Tableau (45) illustrates both the effects of STROLE as well as of ONSET. All forms depicted satisfy anchoring, which is assumed to be dominated by STROLE. Minimal copying, as in (45c), satisfies RED=μ the best, but incurs a fatal violation of ONSET, as previously explained. The interesting candidates then are the former two. Although both extend the prosodic template to better satisfy ONSET, it is the first candidate with some extra copying that wins. We attribute this to satisfaction of STROLE with respect to the coda in the first syllable, which the rival (45b) violates. While one could instead suggest that (45a) better satisfies MAX-BR compared to (45b), this cannot be generally true of this reduplication pattern, since elsewhere (cf. (42)), absolutely minimal reduplicants are preferred instead.

(45) *Arrernte Attenuative*: [**aɲc**-əlp-aɲc-əm] ‘ATTEN-climb-PRES’

/ RED-əlp-aɲc-əm/	STROLE	L-ANCHOR-BR	ONSET	RED=μ
☞ a. **aɲ.c**-əl.p-aɲ.c-əm			*	*
b. **a.ɲ**-əl.p-aɲ.c-əm	*!		*	*
c. **a.**-əl.p-aɲ.c-əm			**!	

What this brief illustration is meant to demonstrate is that the hypothesis we have been entertaining all along against the presence of underlying initial schwas conforms to the empirical facts of Arrernte reduplication in a straightforward manner. Conversely, proponents of the opposing view such as Henderson (2013) need to assume that reduplication happens at some prior level of representation, because output forms actually emerge without schwas initially. No argumentation is given for this move, however. In our reduplication analysis, we have employed analytical ingredients adopted earlier on, namely the absence of *initial* schwas and the insertion of medial schwas. Note that we have marked schwas in affixes and linker morphs as underlying, but this was done for convenience only to avoid burdening the representations too much. They could equally be epenthetic, like other medial schwas.

In sum, STROLE and ONSET can determine the size of the reduplicant, vindicating the role of CV syllabification and moraic onsets in the phonology of Arrernte. Reduplication underscores a recurring theme throughout the paper: like so many other languages, Arrernte utilizes a prosodic template for reduplication. This template (moraic or syllabic), combined with the regulating effects of markedness constraints such as ONSET and STROLE, may cause overcopying of consonants. Thus, even in a language that may so often present onsetless syllables, to the point that some theorists have taken this as evidence for a .VC. syllabification, in fact we do identify an “emergence of the unmarked” effect in its reduplication (McCarthy & Prince 1994).

5. Discussion and Concluding remarks

Thus far, it should be evident that we have addressed four out of the five arguments in favor of a pro-VC analysis of Arrernte, as presented by Hyman (2011b, cf. (1)). As we have demonstrated, all the phenomena of allomorphy (§2), stress (§4.1), Rabbit Talk (§4.2), and reduplication (§4.3) are not only amenable to an analysis that utilizes classic CV-syllabification with moraic onsets, but often such analysis proves superior to its VC-based alternative.

The remaining argument for VC-syllabification relies on the observation that “all words begin with a vowel and end with a consonant”. Apart from this statement being somewhat inaccurate (recall that “25% of Arrernte words are pronounced in isolation with an initial consonant” (B&P: 2)), it does not logically follow that this fact would entail VC-syllabification. Consider for a moment Yapese (Broselow 2003); like Arrernte, Yapese allows for both open and closed syllables word-medially, but word-finally it only permits closed syllables, a fact that can be attributed to a high-ranked FINAL-C constraint that would impose final coda consonants. Once FINAL-C is treated as a licensing constraint that licenses a marked structure at the R-word-edge, the possibility opens up for something like INITIAL-V as a mirror licensing constraint at the L-word-edge, but not internally. Alternatively, or even complementarily, the V-initial effect could be ascribed to the well-known Australian C-drop (§1). The upshot is that a language of the Arrernte type, i.e. #V…C#, is entirely expected given licensing (and historical) considerations without making recourse to an otherwise unsubstantiated type of syllabification.

Two questions are open for discussion about the Arrernte moraic onsets vs. onsetlessness debate: phonetic and diachronic. How did Arrernte come to have moraic onsets? Is it due to the consonant inventory or the vowel inventory? Gordon (2005) suggests that weight-sensitivity in rhymes is dispreferred in languages without vowel length distinctions, and found a greater perceptual energy difference between Arrernte CV and V than between CV and CVC. He argues that the small vowel inventory allows for open syllable lengthening, rendering the overall energy difference between CV and CVC smaller and hence less reliable than the CV vs. V difference. Gordon's claim is thus that given the relative consistency of variability, the CV vs V difference is a more reliable phonetic hook to hang one's phonological hat on, which in our terms, paves the way for phonologization of onsets as moraic.

Is there any phonetic evidence for VC syllabification? In short, no. Tabain et al. (2004) looked for acoustic evidence of planned coarticulation and reduced variability (CV vs. VC) in English, Arrernte, and also Yanyuwa and Yindjibarndi "for which there is no explicit hypothesis of underlying VC syllables". They found that "It might be noted (contrary to our hypothesis) on the rare occasion results are significant for the Aboriginal data, it is the VC context which shows more variability than the CV context" (p.185). Tabain (2009) looked at articulatory kinematics, but "results show no differences between English and Arrernte jaw movement". In short, once we adopt the present analysis, there is neither phonological nor phonetic analysis for VC syllabification. In addition, Henderson (2013: 270) remarks the following: "Non-contrastive aspects associated with stress provide some evidence that supports a CV analysis. Firstly, in emphatic pronunciation, stops preceding stressed vowels may show aspiration, something not noted for stops preceding unstressed vowels. The traditional analysis of similar phenomena across languages interprets this as evidence that the target for stress placement is a CV sequence. The fact that consonants preceding primary stressed vowels are of significantly greater duration than in other contexts [Butcher forthcoming] may also be interpreted in this way."

Our argument overall, therefore, is that Arrernte is much as it appears on the surface; taking stock of the fact that plenty of words start with CV at the left edge (alongside arguments against placing a fleeting underlying schwa at the left edge of such words to save the VC hypothesis) and the evidence for CV syllabification at the left and right edge in connected and sung speech, we have provided: a straightforward account for allomorphy and stress assignment in terms of mora-counting among CV syllables with moraic onsets, an account of reduplication that depends on the presence of an ONSET constraint in the language, and a cross-anchoring account of Rabbit Talk in terms of moraic alignment.

Of special interest for the CV hypothesis of Arrernte is the emergence of the unmarked effect (McCarthy & Prince 1994) that occurs in Arandic Baby Talk (ABT), as described by Turpin et al. (2014). In ABT, speakers deliberately truncate the word-initial vowel of truly V-initial words such as *iwenh* [ɪˈwʊn̪ə] into *wunh* [wʊn̪ɐ] 'what' and *aperley* [aˈpeɭɪ] into *pelyey* [ˈpɪʎɪ] 'father's mother'. If Arrernte had any trace of VC-syllabification, baby talk would then be employing a highly marked structure in these CV transformations, an implausible conclusion. We therefore take this as additional evidence that CV structure is unmarked, and that the impulse for CV syllabification in the baby talk register directly manifests this structure. Indeed, in terms of models of the development from child language to adult language that involve increased promotion of faithfulness (Gnanadesikan 2004), to move from a fully CV-syllabified form such as [ˈpɪ.ʎɪ] towards an adult form of the same word with an initial onsetless vowel like [aˈpe.ɭɪ] would involve a re-ranking of faithfulness above markedness for the initial vowel, but no imaginable resyllabification to a *more* marked configuration could be at play.

The analysis of Arrernte as having moraic onsets within a CV phonology yields the potential for a range of future phonetic studies of their realization. Interestingly, the onset consonants lost in the great Australian C-drop were primarily nasals and glides, which are exactly the mirror image of what are the world's most favored codas (Zec 1995, Morén 2001). This would suggest that above and beyond the general preference for onsets to be low-sonority and codas to be high-sonority, there is a set of preferences on weight, whereby moraic onsets are preferentially low-sonority (cf. Topintzi 2010 for a related proposal), while moraic codas prefer high-sonority – where these latter preferences can only be observed if the language allows moraic vs non-moraic codas, or moraic vs. non-moraic onsets. This latter finding is perhaps ultimately phonologized from the phonetic effect, discussed by Giavazzi (2010), that stressed vowels prefer to align with less sonorous onsets as the latter enhance the loudness of the former. In terms of durational correlations, our claim that Arrernte onsets are moraic (while codas are not) can be directly translated into predictions about the phonetic realization of mora-bearing consonants, in a manner parallel to the crosslinguistic study of Broselow et al. (1997); see also Ryan (2014) for a discussion of the contribution of onsets to weight in gradient terms. More broadly, we contend that among the most fruitful directions for future phonetic studies of Arrernte could be on the relation between initial moraic consonants and the phonetic realization of the preceding fleeting schwa. Arguably, laboratory phonology approaches to Arrernte could yield rich findings through investigating the nature of the initial schwa, paralleling the valuable findings conducted on related phenomena that have been conducted on non-phonological vocalic elements in languages such as Berber (Ridouane & Fougeron 2011), and on intrusive, non-phonological vocalic elements more generally (Hall 2006).

While there is still a lot more to be discovered about Arrernte, in a way, the fact that we have been able to reach the present conclusions at all is precisely the result of the fact that the description of Arrernte has been extremely thorough in terms of morphophonological phenomena, language games, and other types of external evidence that were developed in the context of a coherent competing approach to their syllabification. We have argued that once moraic onsets, independently argued for based on a broad range of phenomena by Topintzi (2010), are adopted, these phenomena follow straightforwardly from the types of processes that typically refer to moras. There may in fact be more languages like Arrernte, with prosodic morphology phenomena that refer to moraicity in such a way that statements like "stress falls on the first vowel preceded by a consonant" becomes "stress the first heavy syllable". Continued descriptions of phenomena such as reduplication and language games are a rich source of discovering the interface between morphology and syllable weight, with the computation of the latter determined in certain language-specific ways that rely both on consonantal quality and syllable position.

References

Abramson, Arthur. 1999. Fundamental Frequency as a Cue to Word-initial Length: Pattani Malay. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 591–594, San Francisco.

Abramson, Arthur. 2003. Acoustic Cues to Word-Initial Stop Length in Pattani Malay. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, August 3-9, 2003, 387–390. Universitat Autònoma de Barcelona.

Bagemihl, Bruce. 1998. Maximality in Bella Coola (Nuxalk). In *Salish Languages and Linguistics, Theoretical and Descriptive Perspectives*, ed. Ewa Czaykowska-Higgins & M. Dale Kinkade, 71–98, Berlin & New York: Mouton de Gruyter.

Bennett, William. 2015. *The Phonology of Consonants: Harmony, Dissimilation and Correspondence*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2001. Where have all the onsets gone? Initial consonant loss in Australian aboriginal languages. In *Forty years on: Ken Hale and Australian language*, ed. Jane Simpson, David Nash, Mary Laughren, Peter Austin & Barry Alpher, 481–492. (Pacific Linguistics 512), The Australian National University.

Blumenfeld, Lev. 2006. *Constraints on Phonological Interactions*. Ph.D. diss., Stanford University.

Borowsky, Toni. 2010, Language disguise in OT: reversing and truncating, *Indigenous language and social identity* (Pacific Linguistics 626), Canberra, ACT.

Breen, Gavan. 2009. Arrernte. In *Concise Encyclopedia of Languages of the World*, ed. Keith Brown & Sarah Ogilvie, 72–75. Oxford: Elsevier.
Breen, Gavan & Robert Pensalfini. 1999. Arrernte: A language with no syllable onsets. *Linguistic Inquiry* 30.1: 1–26.
Breen, Gavan & Veronica Dobson. 2005. Central Arrernte. *Journal of the International Phonetic Association* 35.2: 249–254.
Broselow, Ellen. 2003. Marginal phonology: phonotactics on the edge. *The Linguistic Review* 20: 159–193.
Broselow, Ellen, Su-I Chen & Marie Huffman. 1997. Syllable weight: a convergence of phonology and phonetics. *Phonology* 14: 47–82.
Bye, Patrick 2008. Allomorphy – Selection, not optimization. In *Freedom of Analysis?* ed. Sylvia Blaho, Patrick Bye, and Martin Krämer, 63–92. Mouton de Gruyter.
Curtis, Emily. 2003. Geminate weight: case studies and formal models. Doctoral dissertation, University of Washington.
Darden, Bill. 1971. A note on Sommer's claim that there exist languages without CV syllables. *International Journal of American Linguistics* 37:126–128.
Davis, Stuart. 1988. Syllable onsets as a factor in stress rules. *Phonology* 5.1: 1–20.
Dixon, R.M.W. 1970. Olgolo syllable structure and what they are doing about it. *Linguistic Inquiry* 1: 273–276.
Downing, Laura. 1998. On the Prosodic Misalignment of Onsetless Syllables. *Natural Language and Linguistic Theory* 16: 1–52.
Evans, Nicholas. 1995. Current Issues in Australian phonology. In *The Handbook of Phonological Theory*, ed. John Goldsmith, 723–761. Blackwell.
Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–448.
Everett, Daniel. 1988. On Metrical Constituent Structure in Pirahã Phonology. *Natural Language and Linguistic Theory* 6: 207–246.
Everett, Keren. 1998. Acoustic correlates of stress in Pirahã. *Journal of Amazonian Languages* 1.2: 104–162.
Everett, Daniel & Keren Everett. 1984. On the relevance of syllable onsets to stress placement. *Linguistic Inquiry* 15. 705–711.
Gabas Jr., Nilson. 1999. *A Grammar of Karo (Tupi)*. Ph.D. diss., University of California, Santa Barbara.
Giavazzi, Maria. 2010. *The phonetics of metrical prominence and its consequences for segmental phonology*. Ph.D. diss., MIT.
Gnandesikan, Amalia (2004). Markedness and faithfulness constraints in child phonology. In: R. Kager, J. Pater & W. Zonneveld (eds.), *Constraints in phonological acquisition*. Cambridge University Press. 73-108.
Goedemans, Rob. 1998. *Weightless Segments*. The Hague: Holland Academic Graphics.
Gordon, Matthew. 2005. A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23: 595–653.
Gordon, Matthew. 2006. *Syllable Weight: Phonetics, Phonology, Typology*. New York: Routledge
Hajek, John & Rob Goedemans. 2003. Word-initial Geminates and Stress in Pattani Malay. *The Linguistic Review* 20: 79–94.
Hale, Ken. 1942. Classification of northern Paman languages, Cape York Peninsula, Australia: A research report. *Oceanic Linguistics* 12: 255–302.
Hale, Ken 1984. Remarks on creativity in Aboriginal verse. In *Problems and solutions: occasional essays in musicology presented to Alice M Moyle*, ed. J.C. Kassler & J. Stubington, 254–262. Sydney: Hale & Iremonger.
Halle, Morris & Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, MA: MIT Press.
Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23: 387–429.
Hansen, A. B. 1997. Le nouveau [ə] prépausal dans le français parlé à paris. In *Polyphonie pour Iván Fónagy*, 173–198. Paris: L'Harmattan.
Hayes, Bruce. 1986. Inalterability in CV phonology. *Language* 62: 321–335.
Hayes, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20: 253–306.
Henderson, John. 2013. *Topics in Eastern and Central Arrernte Grammar*. Originally 1998 Ph.D. diss., the University of Western Australia. München: Lincom.
Hume, Elizabeth, Jennifer Muller & Aone van Engelenhoven. 1997. Non-Moraic Geminates in Leti. *Phonology* 14: 371–402.
Hyde, Brett. 2007. Issues in Banawá Prosody: Onset Sensitivity, Minimal Words, and Syllable Integrity. *Linguistic Inquiry* 38.2: 239–285.
Hyman, Larry M. 1983. Are there syllables in Gokana? In *Current approaches to African linguistics*, ed. Jonathan Kaye, Hilda Koopman, Dominique Sportiche & André Dugas, Vol. 2, 171–179. Dordrecht: Foris.
Hyman, Larry M. 1985. *A theory of phonological weight*. Dordrecht: Foris. (Reprinted, 2003, Stanford: CSLI).
Hyman, Larry M. 2011a. Are there really no syllables in Gokana? Or: What's so great about being universal? *Phonology* 21: 55–85.
Hyman, Larry M. 2011b. Phonological Universals vs. Phonological Diversity. Paper presented at the Australian National University, December 14 2011.
Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: doubling in morphology*. Cambridge: Cambridge University Press.

Itô, Junko. 1986. *Syllable Theory in Prosodic Phonology*. Ph.D. diss. University of Massachusetts, Amherst.
Itô, Junko & Armin Mester. 1997. Sympathy Theory and German Truncations. *University of Maryland Working Papers in Linguistics* 5: 117–138.
Itô, Junko, Yoshihisa Kitagawa & Armin Mester. 1996. Prosodic faithfulness and correspondence: evidence from a Japanese argot. *Journal of East Asian Linguistics* 5: 217–294.
Jakobson, Roman. 1962. *Selected writings I: Phonological Studies*. The Hague: Mouton.
Labrune, Laurence. 2012. Questioning the universality of the syllable: evidence from Japanese. *Phonology* 29: 113–152.
Lowenstamm, Jean. 1996. CV as the Only Syllable Type. In *Current Trends in Phonology Models and Methods*, Jacques Durand & Bernard Laks (eds.), European Studies Research Institute, University of Salford, pp. 419-442.
Katsanis, Nikolaos [Κατσάνης, Νικόλαος]. 1996. *The dialect of Samothraki* [Το γλωσσικό ιδίωμα της Σαμοθράκης]. Municipality of Samothraki.
Kawahara, Shigeto. 2016. Japanese has syllables: a reply to Labrune (2012). *Phonology* 33.1: 169–194.
Kiparsky, Paul. 2013 On the empirical bases of phonological typology. Talk at the workshop on Phonological Typology, University of Oxford, Somerville College, 11-13 August 2013.
McCarthy, John J. & Alan Prince. 1986. *Prosodic Morphology 1986*. Technical Report #32, Rutgers University Center for Cognitive Science, 1996. Available at: http://works.bepress.com/john_j_mccarthy/54.
McCarthy, John J. & Alan Prince. 1993. Prosodic Morphology I: Constraint Interaction and Satisfaction. Technical Report #3, Rutgers University Center for Cognitive Science. Available at: http://works.bepress.com/john_j_mccarthy/53. Version of 2001.
McCarthy, John J. & Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistics Society* 24: 333–379. Amherst, MA: GLSA. Available at: http://works.bepress.com/john_j_mccarthy/43.
McCarthy, John J. & Alan Prince. 1995. Faithfulness & Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. Jill Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 249–384. Amherst, MA: Graduate Linguistic Student Association.
Morén, Bruce. 2001. *Distinctiveness, Coercion and Sonority: a Unified Theory of Weight*. London: Routledge. Originally 1999 Ph.D. diss., University of Maryland at College Park.
Nevins, Andrew. 2011. Phonologically conditioned allomorph selection. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice, 2357–2382. Malden, MA: Blackwell.
Nevins, Andrew and Bert Vaux. 2003. Metalinguistic, Shmetalinguistic: The Phonology of Shm- Reduplication'. *The Proceedings of the 39th Chicago Linguistics Society (2003*). 702-721.
Payne, David and Furne Rich. 1988. Sensitivity to Onset in Arabela Stress. ms., Información de Campo 743. Yarinacocha, Pucallpa, Perú: Instituto Lingüístico de Verano.
Pensalfini, Rob. 1998. The Development of (Apparently) Onsetless Syllabification: A Constraint-Based Approach. In: *Proceedings of the Chicago Linguistics Society* (*Annual Meeting 1998)*, 167–178. Chicago.
Pittau, Massimo 1991. *Grammatica della lingua sarda*. Carlo Delfino editore.
Ridouane, Rachid & Cécile Fougeron. 2011. Schwa elements in Tashlhiyt word-initial clusters. *Journal of Laboratory Phonology* 2: 1–26.
Round, Erich. 2013. Why reduplicate VC? Kuuk Thaayorre answers a lingering question. Paper presented at the 44th Annual Conference of the Australian Linguistics Society, Melbourne.
Ryan, Kevin. 2014. Onsets contribute to syllable weight: statistical evidence from stress and meter. *Language* 90: 309–341.
Smith, Jennifer L. 2005. *Phonological Augmentation in Prominent Positions*. New York and London: Routledge.
Sommer, Bruce. 1970. An Australian language without CV syllables. *International Journal of American Linguistics* 36: 57–58.
Sommer, Bruce. 1981. The Shape of Kunjen. In *Phonology in the 1980's*, ed. D. L. Goyvaerts, 231–244. Ghent: Story-Scientia.
Spradlin, Lauren. 2015. OMG the Word-Final Alveopalatals are Cray-Cray Prev: A Morphophonological Account of Totes Constructions in English. Paper presented at the Penn Linguistics Colloquium, March 22, 2015.
Steriade, Donca. 1982. Greek Prosodies and the Nature of Syllabification. Ph.D. dissertation, Massachusetts Institute of Technology.
Strehlow, T.H.G. 1942. Arandic Phonetics. *Oceania* 3.2: 248–264.
Tabain, Marija. 2009. A preliminary study of jaw movement in Arrernte consonant production. *Journal of the International Phonetic Association* 39.1: 33–51.
Tabain, Marija, Gavan Breen & Andrew Butcher. 2004. VC vs. CV syllables: a comparison of Aboriginal languages with English. *Journal of the International Phonetic Association* 34.2: 175–200.
Tabain, Marija & Gavan Breen. 2011. Central vowels in Central Arrernte: A spectrographic study. *Journal of Phonetics* 39: 68-84.
Topintzi, Nina. 2008. On the Existence of Moraic Onset Geminates. *Natural Language and Linguistic Theory* 26.1: 147–184.
Topintzi, Nina. 2010. *Onsets: Suprasegmental and prosodic behaviour*. Cambridge: Cambridge University Press.
Turpin, Myfany. 2015. Alyawarr women's song-poetry of Central Australia. *Australian Aboriginal Studies* 1: 66–96.

Turpin, Myfany, Katherine Demuth & April Ngampart Campbell. 2014. Phonological aspects of Arandic baby talk. In Rob Pensalfini, Myfany Turpin and Diana Guillemin (eds.) *Language Description informed by Theory*. Amsterdam: John Benjamins. 49-80.

Uffmann, Christian. 2007. Intrusive /r/ and optimal epenthetic consonants. *Language Sciences* 29: 451–476.

Wetzels, W. Leo, Stella Telles & Ben Hermans. 2014. Onsets and syllable prominence in Umutina. *DELTA* 30: 703–720.

Wilkins, David P. 1989. *Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar*. Ph.D. diss., Australian National University.

Vogt, Barbara. 2009. *Prosodische Morphologie in deutschen Geheimsprachen*, Ph.D. diss., University of Verona.

Yu, Alan. 2007. *A natural history of infixation*. Oxford: Oxford University Press.

Yupho, Nawanit. 1989. Consonant Clusters and Stress Rules in Pattani Malay. *Mon-Khmer Studies* 15: 125–137.

Zec, Draga. 1995. Sonority Constraints on Syllable Structure. *Phonology* 12: 85–129.

APPENDIX A: Arrernte transcription

Arrernte transcription conventions are at times confusing given the multitude of systems that have been used in the literature. Beyond the typical orthographic one (e.g. Breen 2009: 73), we find IPA-based ones (e.g. Breen & Dobson 2005 and in some cases in Henderson 2013) or a mixture of those (e.g. B&P: Appendix A). We present the consonantal sounds of Arrernte in IPA alongside their correspondents in the orthographic system. The table below largely follows Breen & Dobson 2005.

Correspondence between IPA (left) and Australianist orthography (right)

	Bilabial	Dental laminal	Alveolar apical	Retroflex/Post-alveolar laminal	Alveopal. Apical	Velar
plosive	p *p*	t̪ *th*	t *t*	ʈ *rt*	tʲ/c *ty*	k *k*
nasal	m *m*	n̪ *nh*	n *n*	ɳ *rn*	nʲ/ɲ *ny*	ŋ *ng*
pre-stopped nasal	ᵖm *pm*	ᵗn̪ *thn*	ᵗn *tny*	ᵗɳ *rtn*	ᵗnʲ/ᶜɲ *tny*	ᵏŋ *kng*
lateral		l̪ *lh*	l *l*	ɭ *rl*	lʲ/ʎ *ly*	
approximant	w *w*			ɻ *r*	j *y*	
tap/trill			ɾ *rr*			

Also, some combinations as they appear in Henderson (2013):
tʷ = tw ɲc = nty ɳp = rnp ɭk = rlk ʎc / ʎɟ = lty

APPENDIX B: Rabbit talk forms as given in B&P

Rabbit talk data as presented in B&P: 7-8, but re-numbered and adapted using IPA transcription (cf. Breen & Dobson 2005)

		Arrernte	Rabbit Talk	
I.	a.	əməɳ	əɳəm	'plant food'
	b.	əkʷəɳətʲək	əɳətʲəkəkʷ	'to put in'
	c.	itiɾəm	iɾəmit	'thinking'
	d.	aɾaŋkʷ	aŋkʷaɾ	'no'
	e.	iŋʷən̪t̪	ən̪t̪iŋʷ	'tomorrow'
	f.	əjan̪	an̪əj	'there (nearby)'
	g.	əlʲaʈ	aʈəlʲ	'now, today'
	h.	əkəl	ələk	'right, OK'
II.	i.	ən̪t̪əm	əmən̪t̪	'giving'
	j.	ulkəʈ	əʈulk	'perentie (type of large lizard)'
	k.	alpətʲək	ətʲəkalp	'to go back'
III.	l.	aʈʷ	əjaʈʷ	'initiated man'
	m.	iŋk	əjiŋk	'foot'
	n.	əmp	əjəmp	'come on'

APPENDIX C: Clusters in Arrernte

Medial clusters

Henderson (2013: 22) lists the following *surface-medial* clusters.

		C2										
		p	t̪	t	c	ʈ	k	m	ŋ	^{p}m	^{c}ɳ	^{k}ŋ
	m	mp$^{(w)}$										
	n̪		n̪t̪$^{(w)}$									
	n	np$^{(w)}$		nt$^{(w)}$	nc$^{(w)}$		nk$^{(w)}$	nm	nŋ$^{(w)}$			
	ɲ				ɲc$^{(w)}$							
C1	ɳ	ɳp$^{(w)}$				ɳʈ$^{(w)}$	ɳk$^{(w)}$	ɳm	ɳŋ			
	ŋ						ŋk$^{(w)}$					
	l̪		l̪t̪$^{(w)}$									
	l	lp$^{(w)}$		lt$^{(w)}$	lc		lk$^{(w)}$			l^{p}m$^{(w)}$		l^{k}ŋ$^{(w)}$
	ʎ	*(ʎ p$^{(w)}$)			ʎc$^{(w)}$							
	ɭ	ɭp$^{(w)}$				ɭʈ$^{(w)}$	ɭk$^{(w)}$			ɭ^{p}m$^{(w)}$		ɭ^{k}ŋ$^{(w)}$
	ɾ	ɾp$^{(w)}$	ɾt̪$^{(w)}$		ɾc$^{(w)}$		ɾk$^{(w)}$			ɾ^{p}m$^{(w)}$	ɾ ^{c}ɲ	ɾ ^{k}ŋ$^{(w)}$

Superscripted ^{w} indicates rounding; it does not count as a separate consonant. Certain rounded counterparts are missing: *[nm^{w}], *[ɳm^{w}], *[ɳŋ^{w}], *[lc^{w}], [ɾ ^{c}ɲ^{w}]. This is presumably an accidental gap. Henderson states that in most cases, it cannot be identified whether the first or the second part of the cluster bears the rounding, thus he represents rounding – when it appears – as affecting both consonants. However, we follow here B&P's notation who only specify it on the second of the two.

B&P (1999: 21) do not list the clusters with prestopped nasals (here in dark gray), but they do mention clusters such as [lm$^{(w)}$], [lŋ$^{(w)}$], [ɭm$^{(w)}$], [ɭŋ$^{(w)}$], [ɾm$^{(w)}$], [ɾŋ$^{(w)}$] and [ɾɲ] which look very similar to the corresponding sequences with prestopped nasals in Henderson (2013). We thus take these to be notational variants and roughly the same. The cluster [(ʎp$^{(w)}$)] is only mentioned as a probable one in some Northern speakers in Henderson (2013) only.

Our claim in Section 4.2 was that the NC clusters in light gray behave as partial geminates and cannot be split by processes such as Rabbit Talk (GEMINATE INTEGRITY). The other clusters in principle could be split.

Initial and final clusters

While we were not able to find a concrete discussion of final clusters, for initial ones Henderson reports (2013: 23) that "the set of clusters which can occur immediately after word-initial underlying /ə/ is more restricted than the set found in other environments… For example, *arrkene* /aɾkənɐ/ "playing", but */əɾk…/ *[ɾk…]". Initially, we find words with an initial singleton C or with a partial geminate, e.g. [(ə)ɻ] "3S.NOM" or [(ə)mp] "come on", respectively, but not with an initial heterorganic cluster, e.g. *[(ə)lk]. We may treat these partial geminates as tautosyllabic (which captures the non-splitting of the cluster, and is also in line with the fact that Arrernte doesn't have true onset clusters); if the only clusters found initially can be the partial geminate ones, then that would provide additional evidence for syllabification wholly in the onset. Recall bases like [(ə)mp] which are compatible with this. This would be because in the former case, where onsets are present, the epenthetic final-schwa could be added to offer them a nucleus even in the absence of the initial schwa, thus producing [mpə] and [ɻə], but the same mechanism would not salvage *[lkə], since this would require the impossible coda-onset initial cluster.