VC vs. CV syllables: a comparison of Aboriginal languages with English

Marija Tabain
Macquarie University, Sydney
m.tabain@uws.edu.au

Gavan Breen
Institute for Aboriginal Development, Alice Springs
dictionaries@iad.edu.au

Andrew Butcher
Flinders University, Adelaide
Andy.Butcher@flinders.edu.au

Traditionally, phonological theory has held that the CV syllable is the basic syllable type across the world's languages. Recently however, Breen & Pensalfini (1999) have challenged the primacy of the CV syllable in phonological theory with data from Arrernte, an Aboriginal language spoken in central Australia. In this study, we set out to see if there is any acoustic phonetic basis to Breen & Pensalfini's claim. We examine real-word data from one speaker of Arrernte, five speakers of English, and three speakers each of Yanyuwa and Yindjibarndi (these are two other Aboriginal languages). Using F2 and F3 measures of the consonant, and locus equation measures, we find that CV does show more stability than VC in the English speakers' data, but that for the Aboriginal language speakers' data, there is a parity between the CV and VC measures. We suggest that this greater parity may be a necessary constraint on languages which have multiple places of articulation (six in the case of the Aboriginal languages studied here). We propose an alternative view of suprasegmental organization, and we suggest that more work is needed in order to understand the phonetic bases of suprasegmental structure.

1 Introduction

1.1 CV vs. VC syllable structure: phonological data from Arrernte

It is almost taken for granted in phonological theory that the basic syllable type in the world's languages is CV (Jakobson 1962, Blevins 1995). The CV syllable is considered to be at the top of the implicational hierarchy: a language does not have a VC or a V or a CVC syllable unless it has a CV syllable.

The primacy of the CV syllable in phonological theory has recently been challenged by Breen & Pensalfini (1999) with phonological data from Arrernte, a language spoken in central Australia. At the surface, Arandic languages certainly have CV syllables: an utterance can consist of just a CV syllable, such as [ma] 'here, take it!'. However, there are some features of Arandic languages that suggest that the traditional syllable is not as central to the structure of these languages as it is of most others. For example, the pronunciation of stressed /i/ is dependent on the following consonant: approximately [ɛ] before an alveolar apical, [i] before a retroflex or an alveo-palatal and in between for others. The evidence for the ubiquity of VC syllables in Arrernte comes from a play language called 'Rabbit Talk', certain reduplication rules, and a stress rule. We use as the basis for our argument McCarthy & Prince's (1986: 2)

Journal of the International Phonetic Association (2004) 34/2
DOI: 10.1017/S0025100304001719

claim that 'It is a commonplace of morphology that rules count moras (m), syllables (s), or feet (F) but never segments'.

Before we outline the phonological arguments, a brief word is in order regarding vowels in Arrernte. A postulated sound change in the history of the Arandic languages resulted in the transfer of the feature of roundness from vowels to the adjacent consonants. This resulted in the loss of the contrast between the two high members of the earlier three-vowel system. The resulting two-vowel system, which seems to be still maintained in one or two dialects (but not in that dealt with by Breen & Pensalfini 1999) has a low vowel /a/ and a featureless vowel /ə/ whose surface features are dependent on the adjacent consonants. Most dialects have innovated a high front vowel /i/ and perhaps a high back rounded vowel /u/, whose status, however, is still unclear (see Breen 2001: 51–52). These are not of high frequency.[1]

The rules for the use of Rabbit Talk and the phonetic form of the words produced by these rules give support to the idea that /ə/ is also an underlying vowel that does not appear on the surface utterance-initially, and that it is present whenever a surface form begins with a consonant.

The main rule of Rabbit Talk is (informally) that the first consonant or consonant cluster and any vowel that immediately precedes it are shifted to the end of the word. If this leaves /ə/ at the beginning of the word, it does not appear on the surface and the surface form is consonant-initial. But if the part of the word that is moved does not include a surface vowel, /ə/ will appear on the surface, separating it from the original final consonant of the word. For example, beginning with /əməɳ/ 'tucker (vegetable food)', pronounced (as a citation form) [məɳɐ],[2] transpose /əɳ/ to the beginning of the word, which would give /əɳəm/, pronounced [ɳəm]. Some other examples are listed in table 1a.[3]

It can be seen that for the first three words in table 1a, which are (surface) consonant-initial and have /ə/ as the vowel following this consonant (or cluster) the Rabbit form has the same number of (surface) syllables as the regular form (disregarding any added non-phonemic final vowel). This is also the case for the next three, which are vowel-initial and have /a/ or /i/ following the first consonant (or cluster). The next three, which are vowel-initial and have /ə/ as the vowel following the first consonant or cluster, have a Rabbit form which is

[1] The consonant phonemes are as follows, using IPA symbols:

	peripheral		laminal		apical	
	bilabial	velar	dental	alveolar	alveolar	post-alveolar
stop	p	k	t̪	tʲ	t	ʈ
nasal	m	ŋ	n̪	nʲ	n	ɳ
pre-stopped nasal	pm	kŋ	t̪n̪	tʲnʲ	tn	ʈɳ
lateral			l̪	lʲ	l	ɭ
tap					r	
glide	w	ɰ		j		ɻ

[2] The final non-phonemic vowel [ɐ] that is a common feature of citation and utterance-final forms, especially of short words, in Arrernte is much less frequent in Rabbit Talk.

[3] Note that when presenting Arrernte examples, we use the voiceless stop symbols for the stop series in accord with Arrernte orthography and convention, whereas later in the paper we use voiced stop symbols in order to facilitate comparison with the English data (which consist of voiced stops). Note also that /tʲ/ is referred to as /ɟ/ later, i.e. the voiced palatal stop. Arrernte and the other two Aboriginal languages examined here do not contrast voiced and voiceless stops.

Table 1 (**a**) Rabbit Talk examples from Arrernte. In the second column, the shifted portion is underlined. (**b**) Frequentative reduplication examples from Arrernte. In the second column, the frequentative morpheme is in bold and the reduplicated portion is underlined. (**c**) Attenuative reduplication examples from Arrernte. In the second column, the attenuative morpheme is in bold and the reduplicated portion is underlined.

(**a**)

regular form	rabbit form	English
(ə)kəɻ	(ə)ɻək	meat
(ə)n̪t̪əm	(ə)mən̪t̪	giving
(ə)kʷəɳətʲək	(ə)ɳətʲəkəkʷ	to swallow
aratʲ	atʲar	right
itirəm	irəmit	thinking
atniŋk	iŋkatn	many
akəlʲ	(ə)lʲak	small
itətʲək	(ə)tʲəkit	to light (a fire)
awənk	(ə)nkaw	young woman
(ə)waɻ	aɻəw	only
(ə)jan̪	an̪əj	that (nearby)
(ə)lʲat	atəlʲ	now

(**b**)

present tense form	reduplicated form	English
(ə)tnəm	(ə)tn**əp**ətnəm	standing
untəm	unt**əp**untəm	running
atʷəm	atʷ**əp**atʷəm	hitting
atʷərəm	atʷər**əp**ərəm	fighting
(ə)mpʷaɻəm	(ə)mpʷaɻ**əp**aɻəm	making
(ə)kəmirəm	(ə)kəmir**əp**irəm	getting up
anʲəntəl̪iləm	anʲəntəl̪il**əp**iləm	putting together

(**c**)

present tense form	reduplicated form	English
aɻəm	aɻ**əlp**aɻəm	looking
(ə)tnəm	(ə)tn**əlp**ətnəm	standing
iləm	il**əlp**iləm	telling
atʷəm	atʷ**əlp**atʷəm	hitting
(ə)mpʷaɻəm	(ə)mpʷ**əlp**əmpʷaɻəm	making
(ə)kutʲəm	(ə)k**əlp**əkutʲəm	gathering
itirəm	it**əlp**itirəm	thinking

Note that in all cases, the forms given here represent underlying forms. For ease of interpretation, brackets denote that the vowel is not usually realized on surface form.

one syllable shorter (on the surface) than the regular form. Finally, the last three, which are consonant-initial and have /a/ (or it could be /i/) as the first vowel, have a Rabbit form which is one syllable longer than the regular form.

Clearly, formalizing the rule stated informally above would be a complex matter with conventional syllabification. However, if we assume that the underlying syllable is VC(C) and that underlying /ə/ surfaces only between consonants,[4] the rule is as in (1).

(1) Transpose the first syllable to the end of the word.

The VC(C) syllable approach leads to a substantial simplification of rules also in the case of reduplication. We will illustrate first for frequentative reduplication of verbs. For example, the present tense form /aɻəm/ 'is looking' (with root /aɻ/ and suffix /əm/) may be reduplicated to /aɻəpaɻəm/, which may be translated 'keeps looking'. Other examples are given in table 1b.

The rules, ASSUMING NO INITIAL UNDERLYING VOWELS AND A CONVENTIONAL SYLLABIFICATION, would seem to be:

(2) Suffix /əp/ to the stem.
(3) Repeat the final consonant or consonant cluster of the stem and any immediately preceding vowel; if there is no preceding vowel insert /ə/ before the repeated consonant (cluster).

ASSUMING VC(C) SYLLABLES as above, the rules are:

(4) Suffix /əp/ to the stem.
(5) Repeat the final syllable of the stem.

Various other reduplication examples can be dealt with in a similar way. Somewhat different, however, is attenuative reduplication, which generally conveys the meaning 'starting to do' or 'doing to a lesser degree'. In this case the morpheme added follows the first consonant of the stem. For example, /ətnʲəm/ 'falling' becomes /ətnʲəlpətnʲəm/ 'staggering'. Some other examples are given in table 1c. In a CONVENTIONAL ANALYSIS, the rule for attenuative reduplication could be written as follows:

(6) Repeat the initial vowel (if any) and the first consonant or consonant cluster of the stem.
(7) Insert /əlp/ after the first consonant or consonant cluster of the reduplicated stem.

Note that this rule, and others given above on the assumption of conventional syllabification, go against the generalization that reduplication operates with prosodic units (moras, syllables or feet) and not segments. The rule WITH VC(C) SYLLABIFICATION is as follows:

(8) Repeat the initial syllable.
(9) Add /əlp/ after the first syllable of the reduplicated stem.

Importantly, the MOST WIDESPREAD STRESS RULE in Arandic, USUALL*y* STATED as in (10), does not have a simple statement in terms of conventional syllables, but does WITH VC(C) SYLLABLES, as in (11).

(10) Main stress falls on the first vowel that follows a consonant.
(11) Main stress falls on the second syllable.

Henderson (1990) found also that plural and reciprocal morphology is sensitive to whether stems are monosyllabic or disyllabic, but the rule is straightforward only with the VC(C) model.

Before we move on to a consideration of the phonetics of syllable structure, we should briefly discuss the question of re-syllabification between the underlying and surface forms. As implied above, there may be re-syllabification at a phonetic level, although this surface syllable structure is not completely clear. However, at least one generalization can be made and some tendencies noted. The consonant of an utterance-initial /əC/ syllable becomes onset at the surface to the following syllable due to deletion of initial-schwa. The second consonant of a cluster, especially if it is heterorganic, may also do this. Utterance-final consonants may become onsets to (usually optional) added non-phonemic final vowels. It is possible that syllables containing schwa are more likely to have onsets and less likely to have codas than syllables containing other vowels; however, the data presented below do not allow us to analyze schwa and non-schwa vowels separately due to an imbalance in the tokens. Moreover,

[4] Whether those consonants could form a permissible cluster or not is irrelevant.

the present study does not claim to illuminate the issue of underlying vs. surface syllable structure in Arrernte. Our data are by definition of a phonetic nature, and we only aim to describe any phonetic differences that may exist between Arrernte, which we have claimed has an underlying VC syllable structure, and English, for which the syllable structure is deemed to be CV.[5]

1.2 Phonetic measures of CV vs. VC syllable structure

To our knowledge, there has been no proposal regarding the phonetic correlates of a CV vs. a VC syllable structure. For this reason, the measures we adopt here reflect our reading of the phonetic literature, especially as it relates to coarticulation.

We propose the following: if a consonant is found to be LESS affected by the following vowel context than by the preceding vowel context, that consonant is organized as part of a CV syllable. We would argue that if coarticulation is at a minimum, the speaker has planned that sequence carefully, attempting to maximize the identity of both the consonant and the vowel.

There is some phonetic evidence from European languages which supports such a view. In their landmark study, Kozhevnikov & Chistovich (1965) argued that the unit of speech planning is the CV syllable – speakers plan the transition from the consonant to the vowel, and plan the duration from one CV syllable to another. The VC transition is, by contrast, produced in the time that remains between successive CV syllables, and is hence much more variable. Tuller & Kelso (1990) found that the phasing of laryngeal opening with oral stop closure/release is more stable at faster rates of speech if the syllable is CV rather than VC (so that faster rates of speech favour a stop release burst, which is a cue to consonant identity). More recently, Kochetov (in press) has presented electromagnetic midsagittal articulometry (EMA) data which demonstrate clear differences in the nature and organization of articulatory gestures in syllable onset and coda positions in Russian. Onset gestures are more tightly controlled – greater in magnitude and sometimes in duration. Coda gestures are more variable and often reduced in magnitude and duration. For further examples of greater stability in syllable onset gestures, the reader is referred to Byrd (1994), Turk (1994) and, for a review, Krakow (1999).

The corollary of these articulatory findings is that from a perceptual point of view, syllable-initial consonants are expected to be more easily identifiable and distinguishable than syllable-final ones. Redford & Diehl (1999) have shown that listeners to English CVC sequences make more errors identifying final consonants than initial ones. Discriminant analysis of acoustic measures confirmed a better performance for initial than for final consonants, ‘suggesting that the perceptual advantage for initial consonants may be attributable to their greater acoustic distinctiveness’ (p. 1563). In these cases, the phonetic data were taken to support a view of an underlying CV syllable structure. Redford & Diehl, for example, suggest that their results are ‘consistent with the hypothesis that cross-language preferences for initial over final consonants emerge through selection for communicative value’ (p. 1560). Benkí (2003) reports similar results for perception of syllables in noise, with CV sequences being much more robust than VC sequences.

It may seem counterintuitive to suggest that less coarticulation between the consonant and the vowel implies a tighter relationship between the two, since a high degree of coarticulation might imply a merging between the consonant and the vowel, the ultimate ‘tight relationship’. However, we would argue that by minimizing coarticulation, the speaker is aiming to provide

[5] It is worth noting that Yi (1999) has argued for a non-canonical syllable structure in Korean, based on psycholinguistic data. Yi suggests that the Korean CVC syllable is more accurately described as having a CV + C (‘body + coda’) structure, rather than a C + VC (onset + rhyme) structure. Although this does not imply that Korean has a VC syllable structure, it does suggest that psycholinguistic and phonetic investigations of syllables may provide a more complex description than grammars would suggest.

as much information as possible to the hearer regarding the syllable (Lindblom 1990). For example, formant values for the consonant are clear, as are formant values for the vowel; moreover, duration of the formant transition is more likely to be longer, thereby giving the hearer more time to integrate the sequence of phonemes. By contrast, with a high degree of coarticulation, the duration of the transition between the two phonemes is more likely to be reduced, thereby giving the hearer less time to integrate the sequence, and formant values for the individual segments are less likely to be distinctive and unique to those particular segments.

Of course, phonetic research rarely shows equal effects across all consonants and vowels. Much of the literature on coarticulation is concerned with the resistance or non-resistance to coarticulation of different consonants and vowels (Recasens 1985, 2002) in addition to the question of anticipatory vs. carryover coarticulation (whether a segment is more affected by the following segment or by the preceding segment, cf. Farnetani 1997). Recasens (1989) has argued that anticipatory coarticulation reflects planning of the speech sequence, and that carryover coarticulation reflects lingering effects from the preceding segment, and is therefore not planned or controlled. Moreover, Recasens (1984) found evidence for greater carryover effects than anticipatory effects in Catalan. Such a view in relation to a single consonant would suggest that the CV unit is planned and that the VC unit is simply the 'leftover', in line with Kozhevnikov & Chistovich's claims. However, the behaviour of individual consonants with respect to anticipatory vs. carryover coarticulation is an area which requires further study, and we hope to be able to provide some additional data on this question in the present study.

We might note finally that all of the studies mentioned above have been carried out on European languages (Catalan, Russian, as well as other studies on Italian, German and English; see, for example, Butcher & Weiher 1976, Farnetani 1990, Farnetani & Recasens 1993). To our knowledge, no work has been carried out on anticipatory vs. carryover coarticulation in any non-European languages. We hope to fill this gap with a preliminary investigation of the influence of preceding vs. following vowel context on a given consonant in Arrernte, and in two other Aboriginal languages (Yanyuwa and Yindjibarndi, which are described in a little more detail below).

One of the constraints which may restrict the coarticulatory effect of consonants on vowels in European languages (and especially English) is the need to maintain perceptual distinctions amongst vowels. Traunmüller (1999) suggests that, in languages with minimally contrastive vowel systems, such as in the North-Western Caucasus and Northern China (2 or 3 central vowels distinguished by degree of openness) it is the absence of such a constraint which allows for quite massive C → V coarticulatory effects (cf. Choi 1992 for Marshallese). Clearly Arrernte is a potential member of this category of languages. Traunmüller proposes that listeners automatically REATTRIBUTE to the consonant properties which are actually present in the vowel. For example, in the case of Northern Chinese, 'a listener can be sure that a syllable has an initial labial consonant, when he has perceived its coda as an [o]' (Traunmüller 1999: 143). These effects occur across both CV and VC sequences, but he gives no indication as to the symmetry or otherwise of their frequency of occurrence. Examples of the reattribution phenomenon abound in Arrernte. For example, in some environments the most important cue to the presence of (what linguists analyse as) a labialized consonant is that the following vowel is fully rounded (e.g. /kʷətəɻ/ → [kʷʊtəɻɐ]). Likewise the strongest audible signal of the presence of a retroflex consonant may be the diphthongization of the PRECEDING vowel (e.g. /aʈəp/ → [ɛɪʈəpɐ]).

In the present paper we compare data from Arrernte with data from English. We also examine data from two other Aboriginal languages, Yanyuwa and Yindjibarndi, in order to see if they behave more like Arrernte, which has been argued to have a VC syllable structure, or more like English, which has been argued to have a CV syllable structure. Although there is no explicit hypothesis that Yanyuwa or Yindjibarndi have underlying VC syllables, there is a tendency in the description of Australian Aboriginal languages to use the phonological word, rather than the syllable, as the basis for the description of many phonological and morphological processes (cf. Dixon 2002). Since the phonological word often contains more than one vowel, it is possible that neither CV nor VC is an appropriate description for the underlying suprasegmental structure of Aboriginal languages. An additional consideration for the Aboriginal languages studied here is Steriade's (2001) suggestion that retroflexes provide the greatest contrast with other coronals in postvocalic position; hence, languages which contain retroflex consonants may perhaps control their VC transitions more carefully than languages which do not. Like Arrernte, both Yanyuwa[6] and Yindjibarndi have six places of articulation in the stop series, and all three languages include a retroflex (with a total of four coronal places of articulation in each language). Hence, it may be a property of all languages with retroflexes that VC transitions must be carefully controlled. We will return to these points in the final discussion section.

2 Method

2.1 Languages, speakers, stimuli, recordings

Arrernte is spoken around Alice Springs in central Australia and is a major language in the area, with around 3000 speakers. Yanyuwa is spoken south-east of Arnhem Land in northern Australia, and is spoken by fewer than 50 speakers, all of whom are above 40 years of age. Yindjibarndi is spoken near Roebourne in north-western Australia by about 600 speakers. All of these languages are thought to belong to the Pama-Nyungan group of Aboriginal languages. English is the official language of Australia (population around 20 million) and is an Indo-European language.

A total of twelve speakers was examined for this study: seven Aboriginal language speakers (one Arrernte speaker, and three speakers each of Yanyuwa and Yindjibarndi), and five English speakers.

The Arrernte language data were recorded by the second author in the Central Australian Aboriginal Media Association (CAAMA) radio studios in Alice Springs in the 1980s. The speaker (RF) is female and was aged in her thirties at the time of recording. The purpose of the recording was to provide a set of words for teaching pronunciation to non-native speakers of Arrernte. The speech was recorded onto reel-to-reel tape at the rate of 7½ inches per second (i.p.s.) for the first half of the data, and at 3¾ i.p.s. for the second half of the data. These recordings were transferred onto CD-ROM as WAV files for use in the present study, using a sampling rate of 22.05 kHz. The data were labelled by the first author using EMU (Cassidy & Harrington 2001).

The Yanyuwa and Yindjibarndi data were recorded in the 1980s by the third author. The speech was recorded onto cassette tape in a quiet field situation using a Sony TCM-5000EV cassette recorder and Sony ECM-D8 microphones. 3 speakers were recorded of each language, two female and one male in each case (Yanyuwa: AI and JM are female, DM is male; Yindjibarndi: KM and TD are female, YW is male). Speakers were aged from their early-forties to late-sixties at the time of recording. The purpose of the speech data collection was research-based. The original data were transferred onto SUN workstations at a sampling rate of 16 kHz. The data were labelled by the first author using WAVES+ and are the same data as were used in the Tabain & Butcher (1999) study.

The English data were recorded by the first author at the Speech Hearing & Language Research Centre, Macquarie University, in order to provide a set of comparison data to the

[6] It may be noted that the 7th place-of-articulation proposed for Yanyuwa by Kirton & Charlie (1978), the palatalized velar, has been ignored in the present study based on phonetic results such as those presented in Tabain & Butcher (1999).

Aboriginal data already in existence (the transfer onto CD-ROM and onto SUN workstations of the Aboriginal data had also been carried out at Macquarie University). The speech was recorded directly onto a PC at a sampling rate of 44.1 kHz. Five speakers (three female and two male) were recorded. The two male speakers (CR and RB) and one female speaker (VW) were naïve speakers with no background in speech and language research; the other two female speakers (LS and MT [the first author]) were involved in speech research. The English data were labelled in EMU. Three of the speakers (LS, MT and RB) were labelled by the first author, and the other two speakers (CR and VW) were labelled by a paid labeller who was naïve to the purpose of the study.

The Arrernte data wordlist was compiled by the second author with the help of a native speaker (see Henderson & Dobson 1994 for a recently-published dictionary of Arrernte); the Yanyuwa and Yindjibarndi wordlists were compiled by the third author using Kirton's then unpublished dictionary of Yanyuwa (now available as Bradley & Kirton 1992) and Wordick's (1982) grammar of Yindjibarndi; and the English wordlist was compiled by the first author using native-speaker intuition and the *Macquarie Dictionary* of Australian English (Delbridge, Bernard, Blair, Butler, Peters & Yallop 1997). It should be noted that electronic resources were not used in the construction of the English wordlist, in order to mimic the construction of the Aboriginal language wordlists, which were compiled using only hard-copy dictionaries/wordlists and speaker knowledge.

Stimuli in each wordlist consisted of real words which contained target syllables in word-initial, -medial and –final position. For the purposes of this study, only consonants which were both preceded and followed by a vowel were selected; hence, no word-initial or word-final consonants were used.

The target syllables for the Aboriginal languages were the stop consonants /b d̪ d ɖ ɟ g/ paired with each phonotactically permissible vowel in CV sequences. The consonants are, respectively, the bilabial, the lamino-dental, the apico-alveolar, the apico-postalveolar (or retroflex), the lamino-palatal and the velar. Yanyuwa and Yindjibarndi each have three vowels /i a u/, with Yindjibarndi having an additional contrast of vowel length. Note that the lamino-dental /d̪/ does not occur before /i/ in Yanyuwa, and the apico-alveolar /d/ does not occur before /u/ in Yindjibarndi.

As noted in the introduction, the Arrernte vowel inventory includes two major vowels /a/ and /ə/. Two other vowels, /i/ and /u/, may also exist in the phoneme inventory of Arrernte, although they are much less frequent in the lexicon (especially /u/). Nevertheless, the orthography maintains an /i/ and an /u/, and this tradition was maintained in the labelling process. Another orthographic sequence, /wə/ (spelled <we>) posed a problem. This is not necessarily a sequence of phonemes as the spelling suggests, but a sequence of /ʷ/, representing a feature of the preceding consonant, and schwa (the vowel following the consonant resembles [u] in this case). However, <we> is a sequence proper when the /w/ is intervocalic, in which case it is considered phonemic. It is also a sequence in some other environments, notably when /j/ follows, as in *kweye* 'girl'. We chose to proceed as follows: where there was a clear phonetic boundary between /w/ and /ə/ (such as a steady state formant pattern for each, and/or a change in intensity in the time waveform), we isolated the /ə/; and where there was no clear boundary, we labelled the sequence as /wə/.

Table 2 gives the total number of consonants, collapsed across all vowel contexts, for each speaker of each of the three Aboriginal languages studied here. It also includes the number of tokens for the English speakers. For the English data, both the voiced and voiceless stops were included in the list of stimuli. However, for the purposes of this study, only the voiced stops /b d g/ were used, since it was thought that they were more comparable to the stop articulations between vowels in Aboriginal languages (most Aboriginal languages, including the ones studied here, do not contrast voiced and voiceless stops). Since English has some twenty vowel phonemes, only a subset were selected for the present study. The English vowels chosen were the short monophthongs /i e a o u æ/ (as found in the words *hid*, *head*, *Hudd*, *hod*, *hood*, *had*, respectively) and the schwa vowel. However, since short vowels cannot occur

Table 2 Number of tokens per consonant for each speaker of each language. Note that English contains only three stop places of articulation, whereas the Aboriginal languages listed at the top of the table contain six stop places of articulation.

language	speaker	b	d̪	d	ɖ	ɟ	g	total
Arrernte	RF	48	30	32	60	44	103	317
Yanyuwa	AI	26	16	19	32	24	36	153
	DM	25	20	24	31	25	28	153
	JM	30	16	24	36	29	34	169
Yindjibarndi	KM	14	23	6	34	22	14	113
	TD	15	23	9	31	22	17	117
	YW	9	26	9	39	20	11	114
English	CR	69	–	69	–	–	66	204
	MT	66	–	61	–	–	60	187
	LS	69	–	62	–	–	60	191
	RB	66	–	51	–	–	43	160
	VW	66	–	66	–	–	60	192

at the ends of words in English (with the exception of schwa), a long vowel was used in this environment (note that a comparable long vowel does not exist for /æ/, so this context was left blank). Long vowels or diphthongs were also used when a CV token containing a short monophthong could not be found. It was hoped that the strategy used here for English might imitate the natural allophonic variability found in languages with only two or three vowels. However, it should be kept in mind that there is a good deal of variability across languages in the numbers and types of vowels used in this study: for example, most of the vowels in the Arrernte data are schwas (which we believe reflects the lexical frequency of this vowel), whereas English has a smaller proportion of schwas and Yanyuwa and Yindjibarndi have no schwas.

For all the languages studied, speakers read through the list of words from beginning to end, uttering each word either two times in succession (in the case of Arrernte) or three times in succession (for the other languages). Some speakers adopted a normal speech rate, while others adopted a slower, more careful style of speech. No speaker spoke quickly. All speakers tended to produce a phrasal boundary at the end of each word (often adopting a list-like intonation) with a strong falling utterance-final contour at the end of each group.

Before we go on to outline the analysis procedure, it is important to discuss the nature of stress in our dataset. Unfortunately, prosodic organization in Australian languages is a much under-studied area. Although grammars of many of these languages describe the first syllable of the word as being stressed, instrumental phonetic analyses have shown that these languages may not all be lexical stress languages. Of the two sets of studies with which we are familiar, one has shown that Warlpiri (a central Australian language) is better characterized as an accent language, with pitch-marking of left prosodic edges (Pentland 2003, Pentland & Ingram (under revision)); while another set has shown that Bininj Gun-wok, a language of northern Australia, displays characteristics typical of lexical stress languages, namely duration and RMS energy effects in addition to pitch effects (Bishop 2002, Fletcher & Evans 2002). There is no reason to believe that the languages of Australia do not vary in the same way European languages do when it comes to stress, accent, and syllable- vs. stress-timing. We are therefore left in a position where we can only say that Yanyuwa and Yindjibarndi seem to have some sort of prominence on the first (CV) syllable of their words, whereas Arrernte, as mentioned above, has some sort of prominence on the second (underlying) VC syllable of the word. Consequently, in our English dataset, although many words have stress on the first syllable, many words also have stress on the second syllable (this was often necessary in order for a full vowel to be placed in word-medial position, as described above). Our inability to control more carefully for effects of stress or accent, due to the lack of prosodic description and phonetic data on Aboriginal languages, is a limitation to the present study, which the reader must keep in mind.

2.2 Duration and formant analyses

Three main sets of measures will be presented: (1) a duration measure; (2) F2 and F3 formant measures at a single point in time; and (3) locus equation measures. All of these analyses were carried out in EMU (Harrington, Cassidy, Fletcher & McVeigh 1993) interfaced with the R statistical package (R Development Core Team 2003).

(12) Duration: The duration of the consonant was measured as a proportion of the total syllable duration, both VC and CV. This measure was intended to capture any differences in temporal organization of the CV vs. VC syllable and was expressed as a value between zero and one. For instance, in a sequence /V_1 C V_2/, where $V_1 = 30$ ms, $C = 30$ ms, and $V_2 = 60$ ms, C is 0.5 of total VC duration (30/[30 + 30]) and 0.33 of total CV duration (30/[30 + 60]). Any aspiration following a stop consonant was included as part of the vowel for purposes of calculating duration. Unfortunately, the raw CV and VC syllable duration was not an appropriate measure in the present study, since we were unable to control carefully enough for position in the word, and since, as mentioned above, the ontological status of lexical stress in our Aboriginal language dataset is unclear.

(13) F2 and F3: F2 and F3 were sampled at the stop closure in the VC syllable and at stop release in the CV syllable. For technical reasons, F3 was unreliable in the Yanyuwa and Yindjibarndi data, and will therefore not be reported here for the six speakers involved.

(14) Locus equations: Locus equations (Lindblom 1963, Sussman, Hoemeke & Ahmed 1993) are an attempt to quantify the F2 transition between the consonant and the vowel across a variety of vowel contexts. Locus equations aim to capture the 'average' amount of coarticulation between a given consonant and multiple vowel contexts. F2 at stop closure or release (depending on whether it is a VC or a CV transition) is treated as a function of the F2 vowel target value. Across a variety of vowel contexts, a regression equation is calculated for a single consonant. The regression analysis returns a slope coefficient, usually between zero and one, which has been deemed to be indicative of the average amount of coarticulation exhibited by the consonant across the vowel contexts (Krull 1987). Indeed, using electropalatographic (EPG) data, Tabain (2000) has shown that, at least for the voiced stop consonants of English, there is an articulatory basis to this claim regarding coarticulation (note, however, that Löfqvist (1999) did not find such support using EMA data). Since a low slope value suggests that F2 for the consonant is relatively fixed regardless of the vowel context, a low slope value is deemed to indicate a high degree of resistance to coarticulation on the part of the consonant. By contrast, a high slope value suggests that F2 for the consonant is primarily determined by F2 for the vowel, which implies much coarticulation between the two.

We expected the Duration and formant (F2 and F3) measures to show much less variability in the CV context than in the VC context for English, and the opposite to be true for Arrernte. This would suggest that the CV sequence is more tightly controlled than the VC sequence for English, and vice versa for Arrernte.

As regards the locus equation data, we expected slope values to be lower in the CV context than in the VC context for English (cf. Sussman, Bessell, Dalston & Majors 1997). This would suggest less coarticulation in the CV context, and hence a sharper identity for the consonant at this point. For Arrernte, we expected VC slope values to be lower than CV slope values.

Given the brief discussion of suprasegmental structure in Australian languages and of cues to retroflexion at the end of the introduction section, we hypothesized that Yanyuwa and Yindjibarndi were more likely to pattern with Arrernte than with English.

2.3 Statistical analysis

For the duration, F2 and F3 data, the consonant in the VC context vs. the CV context was compared using a modified paired t-test. Since our main interest here is in variability rather than in means, we modified the paired t-test so that it resembled the standard Levene test for homogeneity of variances (which is used with an ANOVA). We did this by subtracting the mean value of each condition from all of the values in that condition, and using the absolute value which remained. The paired t-test was then conducted on these remaining absolute values. For the initial tests, we set alpha at 0.05 for each measure for each speaker, and reduced it to 0.01 for posthoc tests according to consonant place-of-articulation. We did not adopt a strict Bonferroni approach in this case, since English and Aboriginal languages have different numbers of places of articulation. We felt that the fact that English has fewer places of articulation was balanced out by the fact that Yanyuwa and Yindjibarndi (and to a lesser extent Arrernte) tended to have fewer tokens per consonant; setting posthoc alpha at 0.01 therefore seemed an appropriate compromise.

For the locus equation data, statistical comparison of slope values was carried out following Pedhazur (1973: 436–450). The slope values in the VC vs. CV context for a given consonant were considered significantly different if the regression sum-of-squares obtained for the separate regression coefficients (= slope values) combined was significantly different from the regression sum-of-squares obtained when a single regression coefficient was used. Since locus equations are calculated for each individual consonant separately and not across consonants, alpha was set at 0.01 for all the locus equation tests.

3 Results

Figure 1 presents the VC and CV duration data for each consonant separately. English language data and Aboriginal language data are presented in different plots. Table 3 presents the statistical significance results for these data.

It appears that there is a trend in the English data for greater variability in VC duration than in CV duration, whereas for the Aboriginal language data there appear to be greater differences between speakers across VC vs. CV contexts and across different consonants. The statistical tests, however, show that with one or two exceptions, there are no significant differences between the CV and VC contexts for either the English speakers or for the Aboriginal language speakers. It might be noted that (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 (though speaker RF, the sole Arrernte speaker, shows no significant results whatsoever).

Figure 2 presents the F2 data sampled at the VC and CV boundaries, and figure 3 presents the F3 data sampled at these boundaries. Data are presented separately for each consonant and for each speaker, with the English language and Aboriginal languages data presented in different plots. Table 4 presents the statistical significance results for the F2 data, and table 5 presents the statistical significance results for the F3 data. (The reader is reminded that, for technical reasons, F3 was not measured for Yanyuwa and Yindjibarndi.)

For the English speakers, there is a clear pattern of greater variability in the VC context than in the CV context. This is in accord with our hypotheses. The results are particularly strong for /d/ where both F2 and F3 show highly significant results, and reasonably strong for /b/ where F3 shows a stronger effect than does F2. However, there is almost no effect for /g/.

For the Aboriginal language data, there is no clear pattern across speakers and almost no significant effects of CV vs. VC context. This is in stark contrast to the English language data for F2 and F3.

Table 6 presents the locus equation data, with English language and Aboriginal language data presented separately. The slope values in the VC and CV contexts and the significance

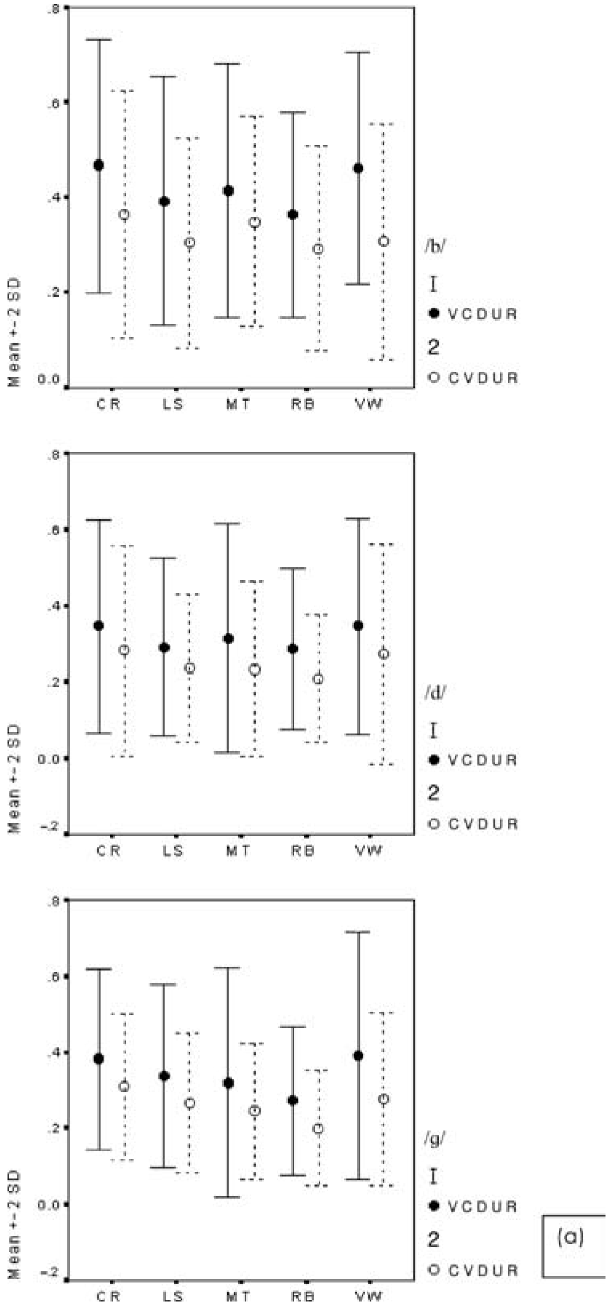

Figure 1 Plots showing mean and twice the standard deviation for DURATION in VC context and CV context for each consonant separately. For details of the measure used, see text. (**a**) English speakers. (**b**) Aboriginal language speakers. Speakers AI, DM and JM are Yanyuwa speakers; speakers KM, TD and YW are Yindjibarndi speakers; and speaker RF is the Arrernte speaker. Note that in this and all subsequent figures, '/dh/' = /d̪/, '/rd/' = /ɖ/ and '/j/' = /ɟ/; and that VC context is represented by a solid line, and CV by a broken line.

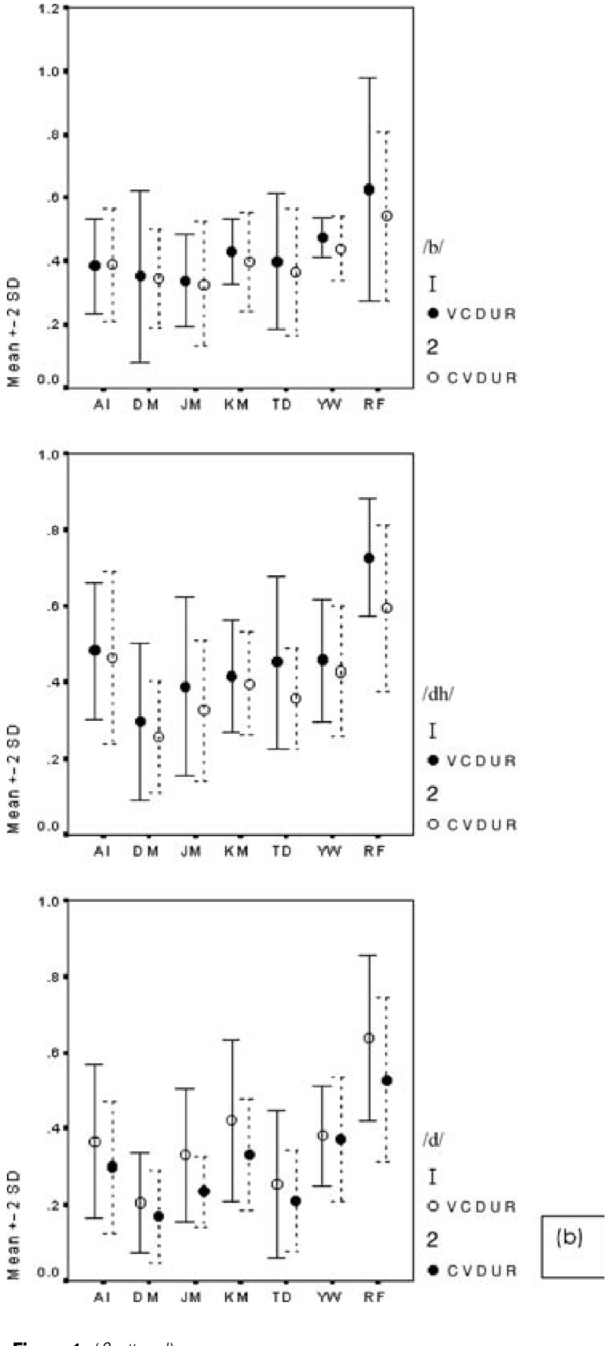

Figure 1 (*Continued*)

results are presented in the same table for ease of interpretation. It can immediately be seen that for English, the VC slope value is significantly higher than the CV slope value for 14 of the 15 statistical tests presented. By contrast, for the Aboriginal language data, of the 42 statistical tests presented, only six of these are significant (i.e. one out of seven). Of these

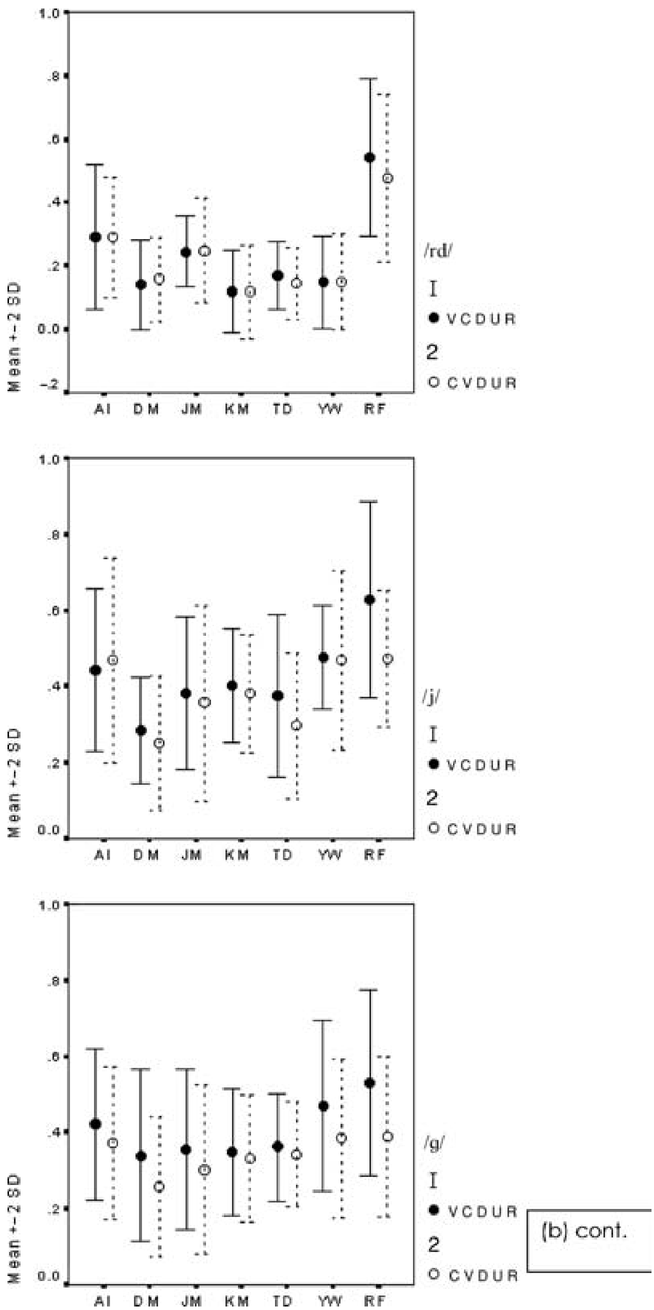

Figure 1 (*Continued*)

six, three show a higher slope value for the CV context, and three show a higher slope value for the VC context. These results strongly suggest that F2 consonant locus is more stable for the CV context in English, whereas for the Aboriginal languages examined here, there is no difference between the CV and VC contexts. It should also be noted that the general

Table 3 (**a**) Significance results for DURATION. Significance level set at 0.05 based on a modified paired t-test (see text). '>' indicates that VC was significantly greater than CV, '<' indicates that CV was significantly greater than VC, and '=' indicates that there was no significant difference. (**b**) & (**c**) Posthoc significance results according to consonant for Duration.

(**a**)

language	speaker	d.f.	t-value	p
English	CR	203	0.85	=
	LS	186	1.80	=
	MT	190	3.15	>
	RB	159	1.20	=
	VW	191	1.21	=
Arrernte	RF	316	1.29	=
Yanyuwa	AI	152	−0.63	=
	DM	152	1.69	=
	JM	168	−0.35	=
Yindjibarndi	KM	112	1.95	=
	TD	116	4.03	>
	YW	113	1.84	=

(**b**)

language	speaker	b	d	g
English	CR	–	–	–
	LS	–	–	–
	MT	–	*	**
	RB	–	–	–
	VW	–	–	–

(**c**)

language	speaker	b	d̪	d	ɖ	ɟ	g
Arrernte	RF	* #	–	–	–	–	–
Yanyuwa	AI	–	–	–	–	–	–
	DM	–	–	–	–	–	–
	JM	–	–	* #	–	–	–
Yindjibarndi	KM	–	–	–	–	–	–
	TD	–	* #	* #	–	–	–
	YW	–	–	–	–	–	–

* $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$
# significance is in the wrong direction (where VC > CV for English, and CV > VC for Aboriginal languages)

pattern of locus equation slope values is maintained in these data: i.e. /g/ slope values tend to be highest, followed by /b/ which is in turn followed by the anterior coronals, and finally /ɟ/ having the lowest slope values.

4 Discussion

Our results strongly suggest important spectral cues to consonant identity are more tightly controlled in the CV sequence than in the VC sequence in English, whereas for the Aboriginal languages, the CV and VC sequences are controlled to the same extent. The locus equation analysis which quantifies the F2 transition supports this interpretation, as do the raw F2 and F3 values.

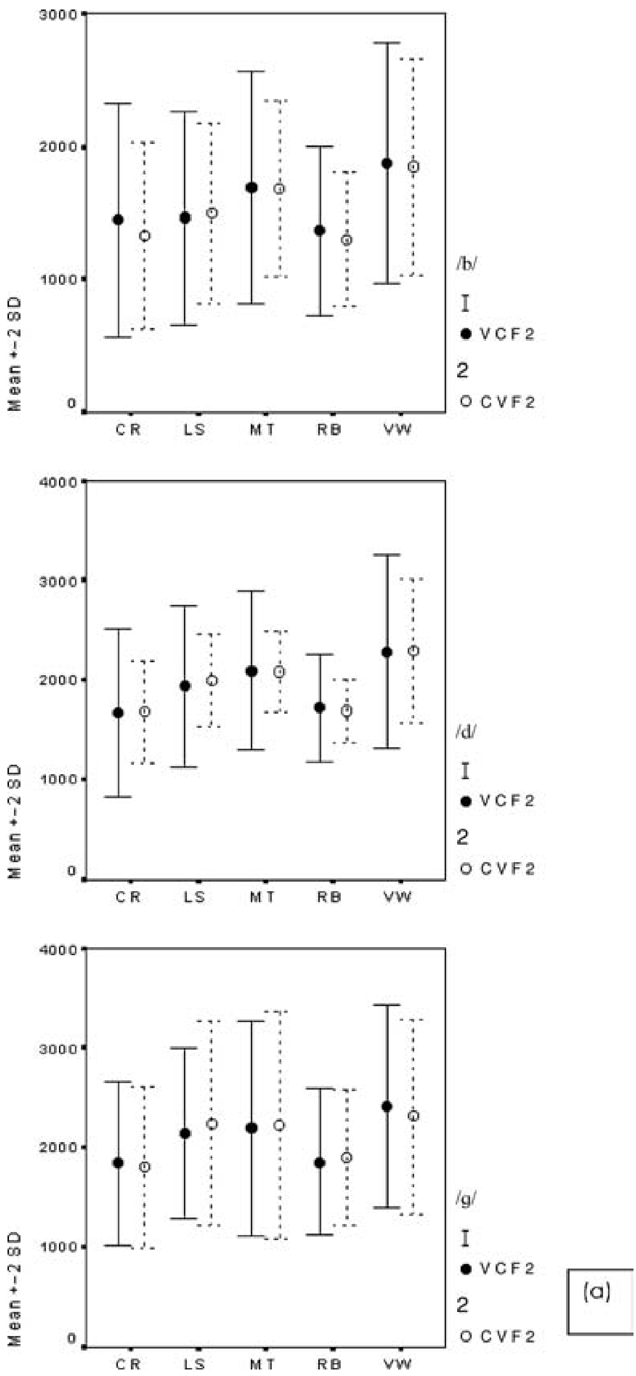

Figure 2 Plots showing mean and twice the standard deviation for F2 in VC context and CV context for each consonant separately. Values are in Hertz. (**a**) English speakers. (**b**) Aboriginal language speakers. Speakers AI, DM and JM are Yanyuwa speakers; speakers KM, TD and YW are Yindjibarndi speakers; and speaker RF is the Arrernte speaker.

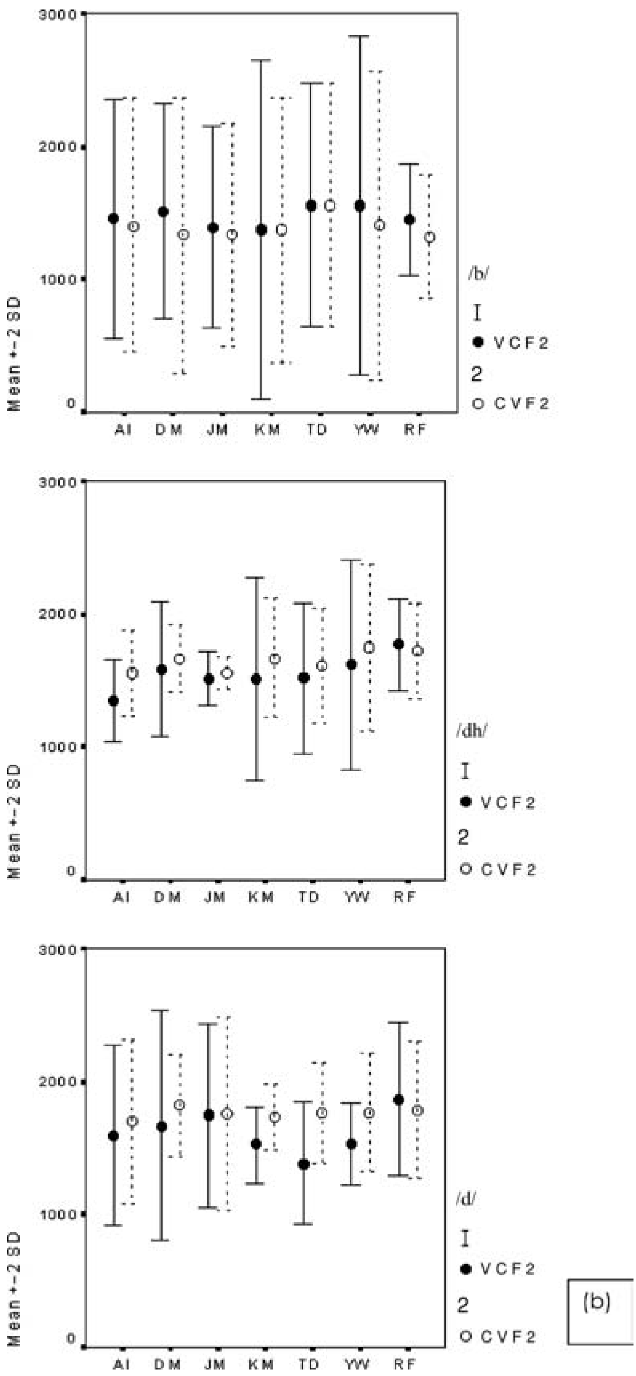

Figure 2 (*Continued*)

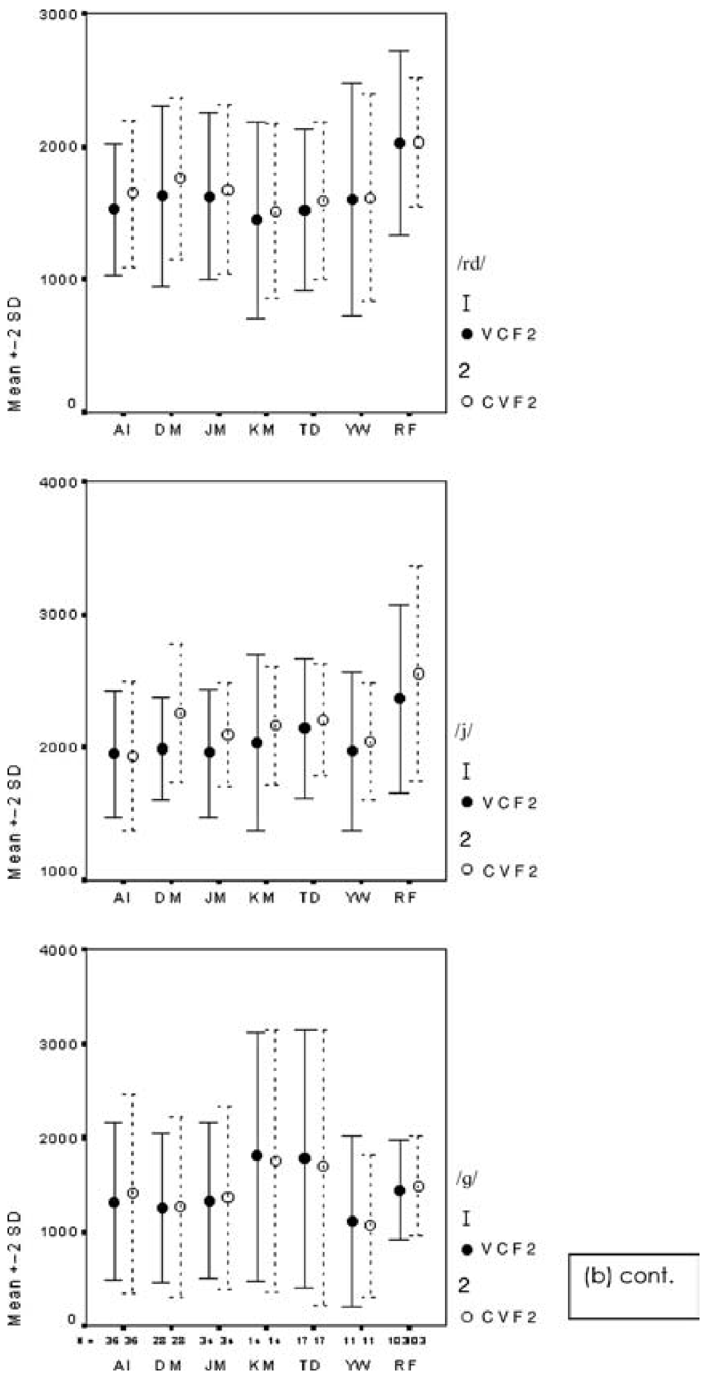

Figure 2 (*Continued*)

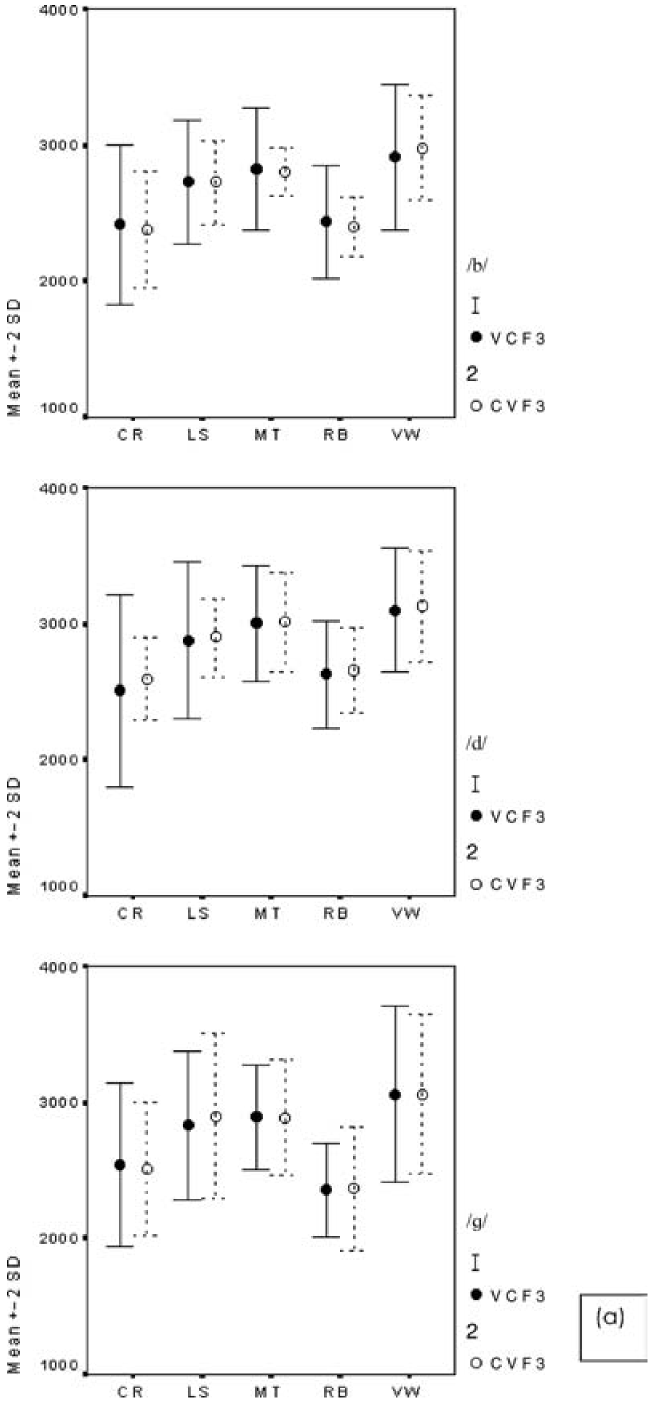

Figure 3 Plots showing mean and twice the standard deviation for F3 in VC context and CV context for each consonant separately. Values are in Hertz. (**a**) English speakers. (**b**) Speaker RF, the Arrernte speaker.

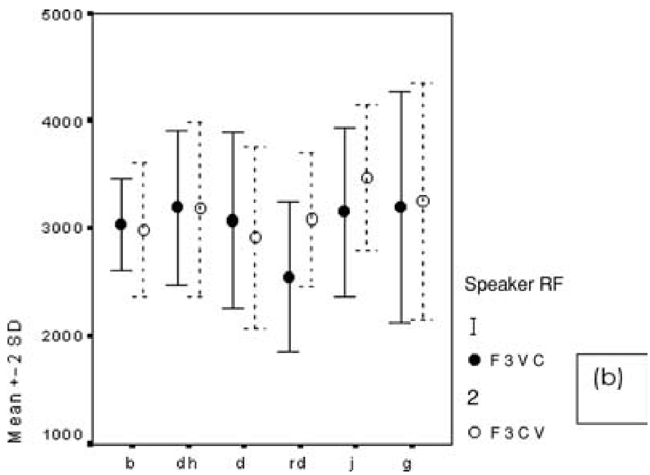

Figure 3 (*Continued*)

It is possible that tight control of both the CV and VC transitions is a necessary constraint on consonant production in languages which have multiple places of articulation. The Aboriginal languages studied here have six places of articulation in the stop series. This reaches the limits on stop contrasts based only on place (Ladefoged & Maddieson 1996). In Tabain & Butcher (1999), we speculated that languages with multiple places of articulation have recourse to multiple cues, both temporal and spectral, to aid the listener in identifying the consonant. We might now add 'VC transition' to the list which contains CV transition, stop burst, F2 & F3 at closure and release, closure duration, transition duration, etc. It is significant that this result of equal weight for CV and VC applies not just to Arrernte, for which there is phonological evidence of an underlying VC structure, but also to Yanyuwa and Yindjibarndi, for which no claim of VC dominance has been made. Following Steriade (2001), we may suggest that the presence of the retroflex consonant (= apico-postalveolar) in these languages motivates greater control of the VC sequence.

It is interesting that, historically, some Aboriginal languages have exhibited initial-consonant dropping (Blevins 2001), resulting in a large number of vowel-initial words. It is thought that the loss of the initial consonant may be due to the first syllable being unstressed in some of these languages, leading to weak spectral cues to word-initial consonant identity. (Significantly, some consonants are more affected than others, presumably due to the confusability of their spectral cues with features of the vowel; for instance, initial glides and nasals are more likely to be lost than initial stops, while some languages have lost only /g/). Arrernte words tend to be vowel-initial in their phonetic form, so that the cues to consonant place of articulation in this consonant-rich language are maximized by the presence of an initial vowel.

We may speculate also on the role of the initial phoneme in guiding lexical access during auditory word recognition in Arrernte. Although the initial phoneme of a word has been shown to play a crucial role in auditory word recognition (Cutler, Dahan & von Donselaar 1997), this is most likely not the case for a language such as Arrernte, with its limited vowel inventory. Instead, it is likely that the vowel is simply a means of increasing the number of cues to the 'real' initial phoneme, the consonant, in this language with so many places of articulation. The presence of a vowel adds cues such as VC transition and stop closure duration to cues such as stop burst duration and CV transition.

Until now we have only discussed the possibility that a language can have a phonological structure based on syllables, and that the underlying preference must be for either CV or VC.

Table 4 (**a**) Significance results for F2. Significance level set at 0.05 based on a modified paired t-test (see text). '>' indicates that VC was significantly greater than CV, '<' indicates that CV was significantly greater than VC, and '=' indicates that there was no significant difference. (**b**) & (**c**) Posthoc significance results according to consonant for F2.

(**a**)

language	speaker	d.f.	t-value	p
English	CR	203	5.02	>
	LS	186	1.95	=
	MT	190	4.10	>
	RB	159	2.95	>
	VW	191	3.58	>
Arrernte	RF	316	−2.33	<
Yanyuwa	AI	152	−1.31	=
	DM	152	−1.72	=
	JM	168	−3.58	<
Yindjibarndi	KM	112	1.67	=
	TD	116	1.40	=
	YW	113	0.42	=

(**b**)

language	speaker	b	d	g
English	CR	***	***	–
	LS	–	***	–
	MT	*	***	–
	RB	–	***	–
	VW	–	***	–

(**c**)

language	speaker	b	d̪	d	ɖ	ɟ	g
Arrernte	RF	–	–	–	–	–	–
Yanyuwa	AI	–	–	–	–	–	–
	DM	**	–	*** #	–	–	–
	JM	–	–	–	–	–	–
Yindjibarndi	KM	–	–	–	–	–	–
	TD	–	–	–	–	–	–
	YW	–	–	–	–	–	–

* $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$

\# significance is in the wrong direction (where VC > CV for English, and CV > VC for Aboriginal languages)

However, another possibility worth considering is that the unit of organization may be best represented by some other combination, such as VCV. In fact, as mentioned above, there is a long tradition in the description of Australian languages (Dixon 2002), of describing the sound structures in terms of a 'phonological word', whose basic form might be summarized as (C)VC(C)V(C). Most phonotactic rules can best be explained in terms of such a unit (cf. Hamilton 1996) and some historical processes appear to ignore the syllable boundary in the centre of this structure. Thus in several languages, for example, prosodic features such as rounding, palatalization, and retroflexion appear to have 'migrated' (from left to right) across this boundary and in others they are synchronically variable in position. Nowhere is this more true than in the Arandic languages. If such VCV languages exist, perhaps they do not form a category in their own right. Perhaps one should rather think in terms of a cline from CV to VC, and perhaps languages migrate between different points on this cline. There is evidence

Table 5 (**a**) Significance results for F3. Significance level set at 0.05 based on a modified paired t-test (see text). '>' indicates that VC was significantly greater than CV, '<' indicates that CV was significantly greater than VC, and '=' indicates that there was no significant difference. Note that the F3 measure was unreliable for Yanyuwa and Yindjibarndi (**b**) & (**c**) Posthoc significance results according to consonant for F3.

(**a**)

language	speaker	d.f.	t-value	p
English	CR	203	5.58	>
	LS	186	3.23	>
	MT	190	1.11	=
	RB	159	0.94	=
	VW	191	2.93	>
Arrernte	RF	316	1.61	=

(**b**)

language	speaker	b	d	g
English	CR	**	***	*
	LS	**	***	–
	MT	*	–	–
	RB	**	*	–
	VW	*	–	–

(**c**)

language	speaker	b	d̪	d	ɖ	ɟ	g
Arrernte	RF	–	–	–	–	–	–

* $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$
# significance is in the wrong direction (where VC > CV for English, and CV > VC for Aboriginal languages)

(from initial consonant dropping and stress shifting) that some Australian languages at least have undergone such a shift, and that Arrernte is amongst those which are furthest along the road from CV to VC.

We should point out that we have as yet not carefully teased out the temporal as opposed to the spectral aspects of CV vs. VC structure. Our most conclusive results are for the spectral data, whereas our temporal data are somewhat inconclusive. Although there is a slight hint that there is less variability in the CV duration ratio than in the VC duration ratio, our data are not controlled well enough for us to draw any conclusions. We assume that a database in which position in the word was more tightly controlled would yield clearer results (of course, number of tokens for each consonant and vowel would also be controlled for in an ideal setting). It would, however, be unusual if temporal and spectral aspects of syllable organization showed differing patterns, since it is quite clear that effects of coarticulation on the formant transition are highly dependent on duration (Moon & Lindblom 1994).

Returning finally to the initial question posed by our study, we did indeed find that Arrernte behaved differently to English as regards CV vs. VC dominance in its phonetic structure. However, we should point out that whilst our hypothesis was that VC would exhibit more stability in Arrernte than CV, we instead found that the two were equal according to our measures. Just what the implications for syllable theory are from our findings are not clear. To what extent our findings reflect suprasegmental organization rather than a need to maintain contrast in languages which have many places of articulation is also unclear.

It is also interesting that our phonetic data for English showed that the three stop consonants behaved differently, at least in terms of F2 and F3 measures, with the consonant

Table 6 Paired locus equation results for (**a**) English, and (**b**) Aboriginal languages. '>' indicates that VC was significantly greater than CV, '<' indicates that CV was significantly greater than VC, and '=' indicates that there was no significant difference following Pedhazur (1973).

(**a**)

language	speaker		VC		CV	d.f.	F-ratio	p
English	CR	b	1.00	>	.71	1,134	17.90	***
		d	.75	>	.37	1,134	71.91	***
		g	.97	>	.71	1,128	13.80	**
	LS	b	.82	>	.58	1,128	16.22	***
		d	.68	>	.19	1,118	87.70	***
		g	.94	=	.73	1,116	3.31	n.s.
	MT	b	.84	>	.50	1,134	29.93	***
		d	.55	>	.09	1,120	84.17	***
		g	1.00	>	.68	1,116	9.30	*
	RB	b	.88	>	.62	1,128	15.77	**
		d	.74	>	.16	1,98	68.16	***
		g	1.15	>	.42	1,82	19.99	***
	VW	b	.81	>	.61	1,128	12.16	**
		d	.67	>	.42	1,128	35.40	***
		g	1.02	>	.65	1,116	23.10	***

(**b**)

language	speaker		VC		CV	d.f.	F-ratio	p
Arrernte	RF	b	.35	=	.28	1,92	0.66	n.s.
		d̪	.50	=	.38	1,56	1.29	n.s.
		d	.41	<	.70	1,60	8.84	*
		ɖ	.51	=	.59	1,116	0.34	n.s.
		ɟ	.60	=	.20	1,84	3.08	n.s.
		g	.48	<	.81	1,202	12.76	**
Yanyuwa	AI	b	.89	=	.80	1,48	1.83	n.s.
		d̪	.42	=	.65	1,28	1.50	n.s.
		d	.80	=	.67	1,34	2.57	n.s.
		ɖ	.66	=	.58	1,60	1.29	n.s.
		ɟ	.15	<	.60	1,44	9.05	*
		g	1.00	=	1.03	1,68	0.15	n.s.
	DM	b	.79	=	.80	1,16	0.02	n.s.
		d̪	.37	=	.32	1,36	0.04	n.s.
		d	.79	>	.36	1,44	37.95	***
		ɖ	.68	=	.58	1,58	1.40	n.s.
		ɟ	.25	=	.28	1,46	0.08	n.s.
		g	.88	=	1.00	1,52	2.20	n.s.
	JM	b	.76	=	.73	1,56	0.34	n.s.
		d̪	.43	>	.02	1,28	11.03	*
		d	.83	=	.79	1,44	0.78	n.s.
		ɖ	.71	=	.71	1,68	0.01	n.s.
		ɟ	.26	=	.28	1,54	0.02	n.s.
		g	1.09	=	.84	1,64	6.61	n.s.
Yindjibarndi	KM	b	.86	=	.79	1,24	0.89	n.s.
		d	.17	=	.31	1,8	0.28	n.s.
		d̪	.61	=	.44	1,42	3.22	n.s.
		ɖ	.68	=	.63	1,64	0.88	n.s.
		ɟ	.50	=	.41	1,40	0.55	n.s.
		g	1.16	=	1.23	1,24	0.18	n.s.

Table 6 (*Continued*)

language	speaker		VC		CV	d.f.	F-ratio	p
	TD	b	.71	=	.82	1,26	2.59	n.s.
		d̪	.52	=	.38	1,42	3.06	n.s.
		d	1.09	>	.43	1,14	9.72	*
		ɖ	.60	=	.54	1,58	0.46	n.s.
		ɟ	.58	=	.41	1,40	4.09	n.s.
		g	1.08	=	1.09	1,30	0.01	n.s.
	YW	b	.86	=	.77	1,14	1.14	n.s.
		d̪	.68	=	.58	1,48	1.00	n.s.
		d	.73	=	.71	1,14	0.01	n.s.
		ɖ	.73	=	.68	1,74	0.70	n.s.
		ɟ	.52	=	.44	1,36	0.45	n.s.
		g	.86	=	1.10	1,18	4.23	n.s.

* $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$

which normally shows greatest resistance to coarticulation, /d/, showing the largest spectral effect between CV and VC, and the consonant which normally shows the greatest amount of coarticulation, /g/, being least affected by CV vs. VC.[7] However, the Aboriginal language data made no such distinctions between consonants: for example, /g/ had high locus equation slope values in both contexts, and /ɟ/ had low slope values in both contexts. If CV is indeed the underlying syllable structure in English, why should our measures be differentially affected by consonant type? Can a language ever exhibit greater stability in VC than in CV measures, rather than showing a parity between them as we have observed here? What, if any, are the best acoustic or articulatory phonetic measures to use in order to determine syllable structure? Or should all our arguments on syllable structure be based on the grammar or on psycholinguistic measures? We hope to have motivated a little deeper thought to these issues with the results presented here.

Acknowledgements

This work was supported by an Australian Research Council Fellowship to the first author. We would like to thank Chris Callaghan, Zoe Evans and Jonathan Harrington for help with various stages of the project, and Jaye Padgett and two anonymous reviewers for helpful comments. We would also like to thank all of our speakers for their time.

We are very sad to note that speaker RF passed away recently. She was pleased that the recording she had made some time ago was being used for further studies of Arrernte, and we dedicate this paper to her memory.

References

BENKÍ, J. (2003). Analysis of English nonsense syllable recognition in noise. *Phonetica* **60**, 129–157.

BISHOP, J. (2002). Aspects of intonation and prosody in Bininj Gun-wok. Ph.D. thesis, Department of Linguistics, University of Melbourne.

BLEVINS, J. (1995). The syllable in phonological theory. In Goldsmith, J. (ed.), *The Handbook of Phonological Theory*, 206–244. Oxford: Blackwell.

BLEVINS, J. (2001). Where have all the onsets gone? Initial consonant loss in Australian Aboriginal languages. In Simpson et al. (eds.), 481–492.

[7] We might note that there appeared to be a trend towards less variability in CV duration than VC duration for /g/; however, results did not reach significance.

BRADLEY, J. & KIRTON, J. (1992). Yanyuwa Wuka: language from Yanyuwa Country – a Yanyuwa dictionary and cultural resource (unpublished manuscript). http://eprint.uq.edu.au/archive/00000072
BREEN, G. (2001). The wonders of Arandic phonology. In Simpson et al. (eds.), 45–69.
BREEN, G. & PENSALFINI, R. (1999). Arrernte: a language with no syllable onsets. *Linguistic Inquiry* **30**, 1–25.
BUTCHER, A. & WEIHER, E. (1976). An electropalatographic investigation of coarticulation in VCV sequences. *Journal of Phonetics* **4**, 59–74.
BYRD, D. (1994). Articulatory timing in English consonant sequences. *UCLA Working Paper in Phonetics* **86**.
CASSIDY, S. & HARRINGTON, J. (2001). Multi-level annotation in the EMU speech database management system. *Speech Communication* **33**, 61–77.
CHOI, J. (1992). Phonetic underspecification and target interpolation: an acoustic study of Marshallese vowel allophony. *UCLA Working Paper in Phonetics* **82**.
CUTLER, A., DAHAN. D. & VON DONSELLAR, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech* **40**, 141–201.
DELBRIDGE, A., BERNARD, J., BLAIR, D., BUTLER, S., PETERS, P. & YALLOP, C. (1997). *The Macquarie Dictionary* (3rd edn.). Sydney: Macquarie Library.
DIXON, R. (2002). *Australian Languages.* Cambridge: Cambridge University Press.
FARNETANI, E. (1990). V-C-V lingual coarticulation and its spatiotemporal domain. In Hardcastle & Marchal (eds.), 93–130.
FARNETANI, E. (1997). Coarticulation and connected speech processes. In Hardcastle, W. & Laver, J. (eds.), *Handbook of the Phonetic Sciences*, 371–404. Oxford: Blackwell.
FARNETANI, E. & RECASENS, D. (1993). Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech* **36**, 279–302.
FLETCHER, J. & EVANS, N. (2002). An acoustic phonetic analysis of intonational prominence in two Australian languages. *Journal of the International Phonetic Association* **32**, 123–140.
HARDCASTLE, W. & MARCHAL, A. (eds.) (1990). *Speech Production and Speech Modelling*. Dordrecht: Kluwer.
HAMILTON, P. (1996). Phonetic constraints and markedness in the phonotactics of Australian Aboriginal languages. Ph.D. thesis, University of Toronto.
HARRINGTON, J., CASSIDY, S., FLETCHER, J. & MCVEIGH, A. (1993). The mu + system for corpus-based speech research. *Computer Speech and Language* **7**, 305–331.
HENDERSON, J. (1990). Number marking in Arrernte verbs. Ms., IAD, Alice Springs.
HENDERSON, J. & DOBSON, V. (1994). *Eastern and Central Arrernte to English Dictionary*. Alice Springs: IAD Press.
JAKOBSON, R. (1962). *Selected Writings 1: Phonological Studies*. The Hague: Mouton.
KIRTON, J. & CHARLIE, B. (1978). Seven articulatory positions in Yanyuwa consonants. *Papers in Australian Linguistics* **11**, 179–199.
KOCHETOV, A. (in press). Syllable position effects and gestural organization: articulatory evidence from Russian. In Goldstein, L., Whalen, D. & Best, C. (eds.), *Papers in Laboratory Phonology VIII: Varieties of Phonological Competence*. Berlin: Mouton de Gruyter.
KOZHEVNIKOV, V. & CHISTOVICH, L. (1965). *Speech: Articulation and Perception*. Springfield: US Department of Commerce.
KRAKOW, R. (1999). Physiological organization of syllables: a review. *Journal of Phonetics* **27**, 23–54.
KRULL, D. (1987). Second formant locus patterns as a measure of consonant-vowel coarticulation. *Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm (PERILUS)* **5,** 43–61.
LADEFOGED, P. & MADDIESON, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.
LINDBLOM, B. (1963). On vowel reduction. *Report 29, Royal Institute of Technology, Stockholm.*
LINDBLOM, B. (1990). Explaining phonetic variation: a sketch of the H & H theory. In Hardcastle & Marchal (eds.), 403–439.
LÖFQVIST, A. (1999). Interarticulator phasing, locus equations and degree of coarticulation. *Journal of the Acoustical Society of America* **106**, 2022–2030.
MCCARTHY, J. & PRINCE, A. (1986). *Prosodic Morphology*. Waltham, MA: Brandeis University.

MOON, S.-J. & LINDBLOM, B. (1994). Interaction between duration, context and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* **96**, 40–55.

PEDHAZUR, E. (1973). *Multiple Regression in Behavioural Research*. New York, NY: Holt, Rinehart & Winston.

PENTLAND, C. (2003). Stress in Warlpiri: stress domains and word-level prosody. Masters thesis, Department of Linguistics, University of Queensland.

PENTLAND, C. & INGRAM, J. (under revision). Fortition-lenition processes and prosodic constituency in Warlpiri. *Australian Journal of Linguistics*.

R DEVELOPMENT CORE TEAM (2003). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. [http://www.R-project.org]

RECASENS, D. (1984). V-to-C coarticulation in Catalan VCV sequences: an articulatory and acoustical study. *Journal of Phonetics* **12**, 61–73.

RECASENS, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech* **28**, 97–114.

RECASENS, D. (1989). Long range coarticulatory effects for tongue dorsum contact in VCVCV sequences. *Speech Communication* **8**, 293–307.

RECASENS, D. (2002). An EMA study of VCV coarticulatory direction. *Journal of the Acoustical Society of America* **111**, 2828–2841.

REDFORD, M. & DIEHL, R. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America* **106**, 1555–1565.

SIMPSON, J., NASH, D., LAUGHREN, M., AUSTIN, P. & ALPHER, B. (eds.) (2001). *Forty Years On: Ken Hale and Australian Languages*. Canberra: Pacific Linguistics.

STERIADE, D. (2001). Directional asymmetries in place assimilation: a perceptual account. In Hume, E. & Johnson, K. (eds.), *The Role of Speech Perception in Phonology*, 219–250. New York, NY: Academic.

SUSSMAN, H., BESSELL, N., DALSTON, E. & MAJORS, T. (1997). An investigation of stop place of articulation as a function of syllable position: a locus equation perspective. *Journal of the Acoustical Society of America* **101**, 2826–2838.

SUSSMAN, H., HOEMEKE, K. & AHMED, F. (1993). A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *Journal of the Acoustical Society of America* **94**, 1256–1268.

TABAIN, M. (2000). Coarticulation in CV syllables: a comparison of EPG and Locus Equation data. *Journal of Phonetics* **28**, 137–159.

TABAIN, M. & BUTCHER, A. (1999). Stop consonants in Yanyuwa and Yindjibarndi: a locus equation perspective. *Journal of Phonetics* **27**, 333–357.

TRAUNMÜLLER, H. (1999). Coarticulatory effects of consonants on vowels and their reflection in perception. *Proceedings of the XIIth Swedish Phonetics Conference, Göteborg University*, 141–144.

TULLER, B. & KELSO, J. A. S. (1990). Phase transitions in speech production and their perceptual consequences. In Jeannerod, M. (ed.), *Attention and Performance XIII*, 429–452. Hillside, NJ: Lawrence Erlbaum.

TURK, A. (1994). Articulatory phonetic clues to syllable affiliation: gestural characteristics of bilabial stops. In Keating, P. (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, 107–135. Cambridge: Cambridge University Press.

WORDICK, F. (1982). *The Yindjibarndi Language*. Canberra: Pacific Linguistics (C:71).

YI, K. (1999). The internal structure of Korean Syllables. *Second International Conference on Cognitive Science and 16th Annual Meeting of the Japanese Cognitive Science Society*. Tokyo, Japan.