CrossMark

Phonology **31** (2014) 1–49. © Julian Bradfield 2014. Published by Cambridge University Press
doi:10.1017/S0952675714000025

*Clicks, concurrency and Khoisan**

Julian Bradfield
University of Edinburgh

I propose that the notions of segment and phoneme be enriched to allow some concurrent clustering, even in classical theories. My main application is the Khoisan language !Xóõ, where treating clicks as phonemes concurrent with phonemic accompaniments allows the inventory size to be radically reduced, so solving the problems of many unsupported contrasts. I show also how phonological processes of !Xóõ can be described more elegantly in this setting, with support from metalinguistic evidence and production-task experiments. I describe a new allophony in !Xóõ. I go on to discuss other applications, some rather speculative, of the concept of concurrent phoneme. The article also provides a comprehensive review of the segmental phonetics and phonology of !Xóõ, together with previous analyses.

1 Introduction

Phonology can be said to have emerged as a discipline with the invention, or discovery, of the notion of 'phoneme' as a contrastive 'unit of sound'. Contrast is a much discussed topic, but in this article I concentrate on the term 'unit of sound', usually now called 'segment'.

When in 2006 the editors of the *Oxford English Dictionary* (2011) revised their entry for 'phoneme', they wrote 'a unit of sound in a language that cannot be analysed into smaller linear units and that can distinguish one word from another'. These words, although they reflect an early 20th-century view of the subject, neatly encapsulate both an old problem and the related problem I wish to discuss.

* E-mail: JCB@INF.ED.AC.UK.

I thank colleagues at Edinburgh: Bob Ladd for many discussions over the years, and comments on drafts, Geoff Pullum and Patrick Honeybone for comments on a draft, and other members of the P-Group, particularly Alice Turk, Bert Remijsen and Satsuki Nakai, for comments following presentations. I am very grateful to Christfried Naumann and Linda Gerlach at the Max Planck Institute for Evolutionary Anthropology in Leipzig for great generosity with unpublished DoBeS data, discussions and comments. Following talks, I received a number of helpful comments, from Dan Silverman, John Goldsmith, Krisztina Polgárdi, Amanda Miller, Tom Güldemann and others. I thank particularly Elan Dresher and Daniel Currie Hall for metatheoretical and theoretical suggestions. Finally I thank the reviewers and editors for their detailed and insightful comments.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

The old problem is what it means to say 'can be analysed into smaller linear units'. The best-known realisation of the problem is the question of affricates *vs.* clusters: the majority view is that /tʃ/ is a single segment in English, but two in German, and conversely for /ts/, but seventy years after Trubetzkoy (1939) discussed it, there is still no unanimity among phonologists. Phonologists studying German range from those who admit no affricates at all to those admit every phonetic affricate as a phonological affricate – see Wiese (1996) for a brief review.

This article, on the other hand, is concerned with 'linear', which is part of the usual understanding of 'segment'. I claim that the restriction to linearity is an undue restriction on the definitions of segment (and hence phoneme), and that in some languages, entities traditionally viewed as single segments should be viewed as clusters. The difference is that the clusters are concurrent, rather than sequential. To put the thesis in a sentence, sometimes a 'coarticulated segment' is better seen as two articulated co-segments.

The notion of concurrent units is already commonplace in certain situations; languages with lexical tone are viewed as placing tones atop segmental units, whether vowels, syllables or words, and sign languages often compose articulations from each hand – though there one can argue about whether the composition belongs in the 'phonology'. Here I extend it to sounds that are in the segmental layer. My main application is the Khoisan language !Xóõ, where treating clicks as phonemes concurrent with phonemic accompaniments radically reduces the inventory size, so solving the problems of many unsupported contrasts. I show also how phonological processes of !Xóõ can be described more elegantly in this setting, with support from metalinguistic evidence and production-task experiments.

I start with a brief discussion of theoretical assumptions and terms. I then discuss the data and previous analyses for the languages that provide the most compelling example of the thesis, present the new analysis, discuss theoretical and empirical evaluation and consider some other examples where the thesis might be applied.

1.1 Preliminaries

1.1.1 *Theoretical assumptions.* My view in this article is representational; adapting a computational process to deal with the new representations is a straightforward task, if it already deals with traditional phonemic representations. Thus, I assume informal notions of segment and phoneme as usually conceived.

Beyond that, I make no commitments in principle to any particular theory. I do not even need to assume the existence of features, though I shall use them descriptively. I do in general assume a mostly linear phonology; the relation to highly non-linear representations such as full-blown autosegmental phonology or gestural phonology is addressed briefly in §4.2.3. For the sake of illustration, I will present formalisations

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

in the framework of *SPE*; similar illustrations could be given for most currently popular frameworks.

1.1.2 *Click basics*. A CLICK is a sound made by creating a 'vacuum' within the oral cavity, bounded by the back of the tongue against the soft palate, with either the sides and front part of the tongue against the hard palate, alveolar ridge or teeth, or the lips. The contact of the tongue back against the soft palate is the 'posterior closure', and the other contact is the 'anterior closure'. The sound is made by releasing the anterior closure, causing an inrush of air to the cavity. If the anterior closure is released sharply, this causes a distinctive 'pop', which is mainly responsible for the very high perceptual salience of clicks. If it is released slowly, the 'pop' is softer, and overlaid with affricated noise. Usually, the posterior closure is released with or very shortly after the anterior, but it can be maintained.

Traditionally, clicks are assigned a velaric airstream mechanism, and placed in a separate section of the International Phonetic Alphabet chart (*IPA* 1999). As Miller *et al.* (2009) point out, the term 'velaric' is a little odd, since the velum is purely passive, and I enthusiastically adopt their suggestion of describing clicks as having a 'lingual' airstream.

The IPA has notations for five clicks, all of which are widely used across the world paralinguistically.

[ʘ] is a BILABIAL click: the anterior closure is made with the lips, and the cavity is made by closing the tongue body against the front of the soft palate, and then drawing it back. The release is affricated, giving a 'kiss' sound.

[ǀ] is a DENTAL click: the anterior closure is made with the blade of the tongue against the top teeth and alveolum. The release is affricated, giving the 'tsk! tsk!' sound.

[ǁ] is the LATERAL (ALVEOLAR) click: the cavity is formed by the sides and tip of the tongue against the alveopalatal region, and released along one side of the tongue. The affricated release gives the 'gee-up!' sound.

[ǃ] is the loudest click: it is ALVEOLAR, with tip and sides of the tongue against the alveopalatal area, and then the tongue sharply hollowed and released at the tip abruptly, giving a 'pop'.

Finally, [ǂ] is the PALATAL click: the closure is made with the blade of the tongue (not the tip) against the alveopalatal area, and the cavity is made by hollowing the centre part of the tongue, and then abruptly released at the front, giving a sharper 'pop'.

A sixth click, which has not received an IPA symbol, but is sometimes notated [ǃǃ] or [ǃ̢], is the true retroflex click. This is similar to [ǃ], but the tongue tip is placed a little further back, and the contact may be apical or sublaminal. The impression is slightly softer and higher than [ǃ], and in the Khoisan languages and dialects in which it appears, it corresponds to [ǂ] in the other languages.

A distinctive variation on [ǃ] which is sometimes heard allophonically or idiosyncratically is the ALVEOLAR-SUBLAMINAL PERCUSSIVE or PALATO-ALVEOLAR FLAPPED CLICK, which has the Extended IPA symbol [ǃ̡]. It is

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

made by pronouncing [!], but keeping the front of the tongue relaxed, so that after release the front flies downward and the underside of the blade strikes the floor of the mouth, which can generate a very audible 'thud' after the pop.

Linguistically, clicks are usually combined with various manners of articulation such as voicing or aspiration applied to the posterior release; this is the topic of this article, and will be discussed in detail. Traditionally, the term INFLUX is used to refer to the actual click sound created by the release of the anterior closure, and ACCOMPANIMENT (or 'efflux', in older work) to the accompanying pulmonic-initiated sounds from the release of the posterior closure.

1.1.3 *Notation.* This article deals primarily with Khoisan languages and their click consonants. This topic is particularly bedevilled by notational issues: the 'correct' phonological analysis is something on which almost every researcher has their own, different, opinion (this article is not an exception), and therefore their own notation; but it is even harder than usual to write a neutral 'phonetic' transcription without implicitly subscribing to one or other phonological analysis. In addition, scholars of the languages have used their own practical transcriptions when recording data; for example, Traill (1985, 1994), my main source here, uses a system that is IPA-like, but not quite IPA. I shall therefore be particularly careful to distinguish notations. In running text, I shall write sounds and words within guillemets ‹ ›, using an IPA-based notation, which tries to give a non-committal but phonological description of the sounds. I use standard IPA diacritics to indicate modification of the click's posterior release: for example, ‹ǃ̬› is a voiced alveolar click, and ‹ǁ̥̃› is a voiceless nasal lateral click (the redundant ‹◌̥› is added for clarity). An important point is that the writing of a velar or uvular stop next to a click (e.g. ‹!q›) indicates a phonologically significant prolongation of the posterior closure; it is *not* part of the notation for the click itself, unlike the notation of Ladefoged & Maddieson (1996). I use / / to make explicitly phonemic assertions, and [ ] when discussing non-phonological detail. Generally, I normalise data to this phonemic notation; when I quote literally from a data source, I shall use conventional orthographic brackets, i.e. ⟨ ⟩.

It is convenient to have a symbol for a generic click – I shall use ‹ʞ›. This metasymbol will be promoted to a phonological symbol during the course of the article.

1.2 Khoisan and clicks

1.2.1 *Khoisan languages and language names.* KHOISAN, first coined in the form 'Koïsan' by Schultze Jena (1928) as an ethnographic term to encompass the Khoekhoe and San 'races', is a Greenbergian (Greenberg 1950) classification of those languages of southern Africa which make extensive use of clicks, other than the Bantu languages (which are generally thought to have borrowed the clicks from Khoisan). The relatedness of all

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

the Khoisan languages is no longer accepted, but the term remains as one of convenience in linguistic use, although it is politically sensitive as an ethnographic term.

There are two Tanzanian languages, Hadza (about 800 speakers) and Sandawe (about 40,000 speakers), which are conventionally included under Khoisan. Hadza is not known to be related to other languages; Güldemann & Elderkin (2010) argue that Sandawe is related to Khoe-Kwadi.

The Khoe-Kwadi family includes several living languages, of which by far the largest is Khoekhoe, with around 270,000 speakers, mainly in Namibia. The Khoekhoe are the groups known as 'Hottentots' in colonial times.

The Tuu family has now only one living example: Taa or !Xóõ, with around 4000 speakers in Namibia and Botswana, the main object of my study here. There are also a few remaining elderly speakers of N|u. It is not generally accepted that Tuu is related to Khoe-Kwadi. Current researchers prefer the name Taa for the dialect cluster which includes !Xóõ (now spelt !Xoon); however, following my main source, and Lewis *et al.* (2013), I shall continue to use !Xóõ.

Finally, the !Kung or Ju family has around 45,000 speakers in Namibia, Botswana and Angola, and includes Ju|'hoansi; recently Ju has been related with the previously isolated language ǂHoã to form a larger Kx'a family (Heine & Honken 2010).

The term 'San' is used as an ethnographic term for the (largely hunter-gatherer) Tuu and Ju peoples, as opposed to the (largely pastoralist) Khoe-speaking groups. Some authors also use San to include the Khoe speakers, but this is resisted by some non-Khoe speakers, who also sometimes object to the 'Khoe-San' compound nomenclature. As *san* is itself a rather derogatory Khoekhoe word, literally 'gatherer, forager', but by extension 'a person who does not own cattle, poor person, outsider' (Haacke & Eiseb 2002), some 'San' prefer to be called by the colonial term 'Bushmen' (Besten 2006).

1.2.2 *Khoisan complexity*. The Khoisan languages are famous for their sometimes huge inventories of consonants. Of the living languages, !Xóõ is usually said to have the largest inventory. In Ladefoged & Maddieson (1996), the inventory for click consonants alone is given as 85 distinct segments (or rather 83, since two are unattested), and this increases to 115 in Naumann (forthcoming). The relatively modest Khoe-Kwadi language Khoekhoe has 20 click consonants, and most of the other languages fall between. (Using the same counting, Zulu has 15, and Xhosa 18.)

The typical Khoisan language has clicks at four places of articulation, of which three are borrowed by Bantu languages such as Zulu. These are alveolar ! (Zulu *q*), dental | (Zulu *c*), lateral ‖ (Zulu *x*) and palatal ǂ. A few surviving languages also have bilabial ʘ. The enormous inventories come from the many accompaniments with which these four or five basic clicks

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

can be varied. These languages, and !Xóõ in particular, provide the primary impetus for the thesis of this article.

2 !Xóõ phonetics and phonology: data

In this section, I review the data that I will use throughout this article. The data is complex, both inherently and because of changes in researchers' understanding, so I aim to provide not just the information necessary for this article, but also a comprehensive overview in a more accessible form than is found in the Khoisanist literature. The major omission is the tonology, which is complex and not perfectly understood; it is not relevant for the purpose of this article, so I give only a sketch.

2.1 The sounds of !Xóõ: overview

Until recently, our knowledge of !Xóõ came mainly from Traill's thirty-year study of the language, the major publications being Traill (1985) and Traill (1994). Traill chiefly studied an eastern dialect. Recently, a project team at the Max-Planck Institute in Leipzig, has, as part of a larger language-documentation project (DoBeS),[1] made a segment inventory of a western dialect (Naumann forthcoming). There are some differences in the analyses (Naumann finds even more distinctions than Traill), but these differences are not essential for the purposes of this paper. I will adopt the DoBeS inventory, but use mainly Traill's data, supplementing it from the DoBeS inventory as appropriate, as the full DoBeS data is not yet publicly available.

2.1.1 *Morphophonological structure.* Although the morphology of !Xóõ is not fully worked out, analyses by Traill (1994), Naumann (2008) and Kießling (2008) can be somewhat crudely summarised as follows.

!Xóõ has a very simple word structure. Phonologically, a content word (noun, verb, adjective) has the form C*V{V/CV/C}, i.e. a first mora, which starts with a possibly empty consonant cluster, followed by a vowel (which carries tone and may have several voice qualities), and a second mora, which is either a vowel (again with tone and perhaps nasalised), or a consonant (from a small set) and a vowel, or just a consonant (a nasal, which appears to carry tone in some cases). Function words are typically but not invariably monomoraic; and loanwords and onomatopoeic words may vary from this structure. For the content words, the first mora is the root, and the second mora carries grammatical information, such as concord class. Most words in a sentence have their second mora determined by that of the 'head noun'; the concord system is fairly complex. In citing

[1] Volkswagen Foundation programme *Dokumentation Bedrohter Sprachen* ('Documentation of Endangered Languages'), available at http://www.mpi.nl/dobes.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

words that inflect for concord, Traill uses the notations ‹V JV BV LV› as morphophonological representations of the second mora, as in (1).

(1) a. The noun ‹ǁχáũ› 'the point between the shoulder-blades' is ‹ǁχá-ű›, where ‹ǁχá-› is the root, and ‹ű› the suffix (it is a class 2 noun, but the ‹u› appears to be arbitrary, with the nasality the only observable association with the class 2 forms).

b. The verb ‹ǂqʰa̰JV› 'squash between the nails' has the nominal form ‹ǂqʰá̰i› and may appear concordially as ‹ǂqʰa̰ji›, ‹ǂqʰa̰ŋa›, ‹ǂqʰa̰je›, ‹ǂqʰa̰ju› or ‹ǂqʰa̰n›, with surface tones also determined by concord.

These words may then be extended with (usually monomoraic) affixes to form longer phonological words; such affixes do not contain clicks. Compound words are also possible, and (at least in the dialect studied by Traill) reduplication of the entire word is a common phenomenon.

2.1.2 *Tone*. Traill marks four surface tones, which apply to the (bimoraic) word: high ‹á›, mid-level ‹ā›, mid-falling ‹â› and low ‹à›. Naumann (2008) analyses this as two monomoraic tones, high and low, so that Traill's surface tones are represented as HH, LH, HL and LL. This analysis is not completely without problems (Naumann forthcoming), but is mostly successful. There remain some monomoraic words which appear to bear a compound tone. The tones are strongly affected by voice quality, and are extensively modified by the concord system. In this article, I shall use Traill's representations for surface tone when citing forms.

2.1.3 *Consonant overview*. Table I presents the consonant inventory of !Xóõ in chart form. The columns are labelled by place of articulation; the rows will be referred to by number. This chart presents the largest inventory: firstly, it includes the DoBeS western dialect analysis; secondly, it presents, in the lower half, a large number of 'consonants' which are notated as phonetic clusters. I discuss in § 3.3 whether these are phonological clusters. In the following sections, I describe the consonants in detail.

2.2 Non-click consonants

A striking feature of !Xóõ (and Khoisan more generally) is that all the consonantal complexity occurs word-initially – only a few consonants occur medially or finally. It is therefore natural to consider the positions separately, and I first describe initial consonants without clicks.

2.2.1 *Initial non-clicks*. This part of the inventory is already quite rich. At the top left of Table I, we have a set of stops with five or six places and five to eight manners, depending on count. Apart from the glottal stop, there are five places of articulation: labial, dental, alveolar, velar, uvular. A typologically unusual feature of !Xóõ is that oral labial stops

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

	labial	dental	alve-olar	pala-tal	velar	uvu-lar	glot-tal	lab click	dent click	alv click	lat click	pal click
1	p	t	ʦ		k	q	ʔ	ʘ	ǀ	ǃ	ǁ	ǂ
2	b	d	ǂ		g	ɢ		ʘ̬	ǀ̬	ǃ̬	ǁ̬	ǂ̬
3	p^h	t^h	$ʦ^h$		k^h	q^h		$ʘ^h$	$ǀ^h$	$ǃ^h$	$ǁ^h$	$ǂ^h$
4	b^h	d^h	$ǂ^h$		g^h	$ɢ^h$		$ʘ̬^h$	$ǀ̬^h$	$ǃ̬^h$	$ǁ̬^h$	$ǂ̬^h$
5	p’	t’	ʦ’		k’	q’		ʘ’	ǀ’	ǃ’	ǁ’	ǂ’
6			ǂ’		g’	ɢ’		ʘ̬’	ǀ̬’	ǃ̬’	ǁ̬’	ǂ̬’
7						qχ’						
8						ɢχ’						
9								ʘ̃	ǀ̃	ǃ̃	ǁ̃	ǂ̃
10	m	n		ɲ	ŋ			ʘ̬̃	ǀ̬̃	ǃ̬̃	ǁ̬̃	ǂ̬̃
11	ʔm	ʔn						ʔʘ̃	ʔǀ̃	ʔǃ̃	ʔǁ̃	ʔǂ̃
12	f		s			χ	h					
13	w	r,l		j								
14								ʘq	ǀq	ǃq	ǁq	ǂq
15								ʘɢ	ǀɢ	ǃɢ	ǁɢ	ǂɢ
16								$ʘq^h$	$ǀq^h$	$ǃq^h$	$ǁq^h$	$ǂq^h$
17								$ʘ̬q^h$	$ǀ̬q^h$	$ǃ̬q^h$	$ǁ̬q^h$	$ǂ̬q^h$
18								ʘq’	ǀq’	ǃq’	ǁq’	ǂq’
19								ʘ̬q’	ǀ̬q’	ǃ̬q’	ǁ̬q’	ǂ̬q’
20	pqχ’	tqχ’	ʦqχ’					ʘqχ’	ǀqχ’	ǃqχ’	ǁqχ’	ǂqχ’
21		dqχ’	ǂ qχ’					ʘ̬qχ’	ǀ̬qχ’	ǃ̬qχ’	ǁ̬qχ’	ǂ̬qχ’
22		tχ	ʦχ					ʘχ	ǀχ	ǃχ	ǁχ	ǂχ
23		dχ	ǂ χ					ʘ̬χ	ǀ̬χ	ǃ̬χ	ǁ̬χ	ǂ̬χ
24								ʘh	ǀh	ǃh	ǁh	ǂh
25								ʘ̬h	ǀ̬h	ǃ̬h	ǁ̬h	ǂ̬h
26								ʘʔ	ǀʔ	ǃʔ	ǁʔ	ǂʔ
27								ʘ̬ʔ	ǀ̬ʔ	ǃ̬ʔ	ǁ̬ʔ	ǂ̬ʔ

Table I
The consonant inventory of !Xóõ.

are marginal: in Traill (1994), almost all the few words starting with labial stops, and all words starting with ‹p›, are loans.

The manners are more or less as written: the voiceless, voiced and aspirated stops (rows 1–3) are familiar from languages with this distinction: voiceless stops have about zero VOT, whereas voiced stops have voice lead, and aspirated stops voice lag. The voiced aspirated stops (row 4) are, however, not like the familiar breathy voiced stops of Indic languages: they have voice lead, which persists into the [z] of ‹ʤ›; voicing ceases at release. Ejectives (row 5) are also familiar; the voiced ejectives (row 6) have voice lead, followed by an ejective release (so ‹ʤ’› is rather [ds’]).

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

The uvular ejective affricates ‹qχ' ɢχ'› (rows 7–8) might be considered another place or manner; because of their occurrence in clusters, it is convenient to arrange them as manners. They are pronounced as notated, although there is some room for argument about whether they are really velar or uvular – see the discussion in §3.1.

Of the plain nasals (row 10), only ‹m n› occur initially. The glottalised nasals (row 11) are initials, and are nasal stops with an initial glottal check. Amongst the continuants (rows 12–13), ‹s χ› and marginally ‹h› occur initially in native words; the others may occur in loanwords.

Finally, at the bottom left of Table I, there is a group of initials written as phonetic clusters. The pulmonic clusters (rows 22–23) are pronounced as written, with a strong uvular fricative. The ejective clusters (rows 20–21) vary according to dialect and register. Again, the exact place is arguable, and in careful eastern speech, Traill records pronunciations such as [t'q'], although with no instrumental confirmation of a true double ejective. These clusters are rare in the DoBeS data, but reasonably supported by Traill (1994), apart from ‹pqχ'›, which occurs only in the superbly onomatopoeic word ‹pqχ'áli› 'the sound of a rapid evacuation of the bowels'.

2.2.2 *Medial consonants*. As remarked in §2.1.1, the bimoraic word may be bisyllabic, with the second syllable starting with one of a very small set of consonants. These are ‹b m n ɲ j l r›. ‹j› in Traill's data varies from [j] to [ɟ]. In Traill, ‹r› occurs only in loanwords; in DoBeS, ‹l› occurs only in loanwords, and ‹r› corresponds to Traill's ‹l› in native words.

2.2.3 *Final consonants*. The final consonants are ‹m n ŋ p b r›. All but ‹m n› are marginal, occurring in loanwords or onomatopoeic words. According to DoBeS, final ‹m n› are more vocalic than consonantal, carrying a mora and a tone. Traill does not mention this, although it is very obviously the case in his recordings.

2.3 Click consonants

All click consonants are initial. I describe the clicks in the order laid out in Table I.

2.3.1 *Simplex click consonants*. The clicks in rows 1–11 of Table I are notated as phonetically simplex consonants. The anterior articulation of these clicks matches their non-click counterparts: for example, ‹ǂ^{h}› (row 4) is a palatal click, with voice lead up to the posterior (velar) closure, and aspiration following the posterior release. The voiceless nasal clicks (row 9) such as ‹$\tilde{ǂ}_{̥}$› have no non-click counterparts. They are pronounced as written: a voiceless ‹ǂ› together with velar lowering around the closure period. This accompaniment will be discussed further in §5.2.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

2.3.2 *Complex click consonants with long closure.* The clicks in rows 14–21 are written with a following [q], which, as noted above, is intended to indicate a significant prolongation of the posterior closure. Thus in ‹ʞ›, the click burst is more or less simultaneous with, and so drowns, the posterior release, whereas in ‹ʞq› the posterior release can be heard after the click burst (and seen on spectrograms).

The various modifications – aspiration, ejection, ejective affrication – of the posterior release are pronounced as written.

The voiced consonants, in the odd-numbered rows, are pronounced with voice lead into the posterior closure period, and it is not unusual to hear nasalisation as well, which is probably simply phonetic enhancement of the prevoicing. Voicing stops before the posterior release.

2.3.3 *Other complex click consonants.* Rows 22–27 contains clicks where the click appears to be (phonetically) followed by another sound. I discuss below whether these are phonological as well as phonetic clusters; here I consider just their phonetics.

The ‹ʞχ› fricative clicks in rows 22–23 are so notated because the fricative is fairly long and prominent, making [ʞχ] more descriptive than the possible alternative [ʞˣ], which suggests an affricated posterior release. As I discuss below, there are also systematic reasons for treating them as a click followed by a fricative.

The ‹ʞh› clicks in rows 24–25 have received special attention in the phonetic literature. This, or a similar, ‹ʞh› accompaniment is found in other languages, including Khoekhoe. It has a distinctive auditory impression, as a long crescendo aspiration (around 200 ms, sometimes even 400 ms) can be heard after the click; but the posterior release is not audible. For Khoekhoe (Nama), Ladefoged & Traill (1984) used airflow measurements to establish that the silent start is achieved by nasal venting during the click [ʞ̃h]; for !Xóõ, Traill (1991) showed that this is supplemented by breathing in during the click (so [ʞ̃ˈh]), making it the only established example of ingressive pulmonic airflow in normal language. There is a question about whether the nasalisation is phonetic or phonological, which will be touched on below. I treat it as phonetic, and do not represent it.

The clicks ‹ʞʔ› in rows 26–27, with glottal stop, also tend to have nasalisation, at least in the voiced version, and this may or may not be phonological – here I have assumed not. They are auditorily distinguished from the ejectives ‹ʞ'› in rows 5–6 mainly by the lack of an audible posterior release – similar to the difference between [ak'a] and [akˀˀa].

2.4 Vowels

The vowel system is also rich. Its basis is a simple five-vowel system, ‹a e i o u›. The front vowels ‹i e› are fairly well localised around approximately cardinal values; ‹o u› tend to spread out a little more, centralising in some contexts, sometimes to the extent of neutralising with each other;

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

‹a› is more variable, spreading over most of the lower half of the IPA chart, between [ɑ a ɜ]. I shall discuss the behaviour of ‹a› in some detail later, in §5.1. As most words are bimoraic, long vowels and diphthongs occur; there seems no reason to treat these as anything other than a sequence of two vowels. The following combinations are *not* attested in Traill: ‹ea eo eu ie io iu uo›, and are also not found in the DoBeS data.

The complexity of the vowel system arises from the addition of voice qualities and nasalisation to the basic vowels. Phonetically, one hears breathy vowels [V̤], where breathiness may extend over the entire stem; creaky vowels [V̰], where the creak usually occurs in the middle of the first vowel (as in, say, Vietnamese), and may vary from light creaky voice (or even be omitted in fast speech) to a full glottal stop; pharyngealised back vowels [Vˤ] in the first vowel; and the 'strident' back vowels [Vˢ], which have strong epiglottal friction and are often voiceless. Although Ladefoged & Maddieson (1996) treated stridency phonetically as a distinct phonation type (and notated it [V̺] to emphasise this), Traill considered (with good reason) that phonologically strident vowels are the realisation of breathy pharyngealised vowels /V̤ˤ/. This has also been adopted in the DoBeS orthography, and I adopt it here also.

Traill also reports breathy creaky vowels [Ṽ̰], which start breathy and then glottalise, creaky pharyngealised back vowels [Ṽ̰ˤ] and even strident creaky vowels [Ṽ̰ˢ], which start strident and become glottalised, and are phonemically creaky breathy pharyngealised /Ṽ̰ˤ/.

Furthermore, all of these also occur nasalised, where the nasalisation is usually heard over both vowels in the stem. However, there are good reasons to believe that nasalisation belongs on the second vowel of a word, whereas the voice qualities belong on the first vowel. Thus phonemically we take the above to be phonemes, and add /Ṽ/.

2.5 Phonotactics and phonological processes

There are several phonetic rules given in Traill (1994) which modify the phonetic realisation of the inventory given above, and also some phonotactic constraints (from Traill 1985) which limit the number of possible words. Here I will describe a few which will form part of my argument later.

(2) a. *Single Aspirate Constraint*
A word contains at most one segment that is aspirated, breathy or strident.

b. *Single Glottal Constraint*
A word contains at most one segment that is glottalised or creaky.

c. *Pharyngeal Constraint*
A pharyngealised or strident vowel may not follow an aspirated, ejected or fricated click (i.e. it may follow only ‹ʁ ǆ ʔǆ ʁq› and their voiced versions).

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

These constraints are strong, but apparently not quite inviolable. Traill (1994) contains four or five lexemes violating (2a), and DoBeS has two. In every case, non-violating alternatives appear to exist, so they may be instances of phonetic spreading. The appearance of 'strident' in (2a) forms part of the evidence for 'strident = breathy pharyngealised'. (2b) has two (related) violating lexemes in Traill (1994), and none in DoBeS, while (2c) applies for the most part to non-click stops as well, but there are a couple of violations there, and in particular, as I shall consider later, Traill (1994) gives half a dozen words in ‹h-› containing pharyngealised vowels.

(3) *Phonetic Back Vowel Constraint*

A back consonant may not be followed by a (phonetic) front vowel (‹i e›), where the back consonants are the velar and uvular non-clicks, or by a click that involves ‹ʘ ! ǁ›.

The Back Vowel Constraint (see e.g. Miller 2011) applies in some form across the Khoisan languages, with varying notions of 'back' (sometimes excluding velars, for example) and different strengths. In the Khoe languages such as Khoekhoe, the Back Vowel Constraint is far from absolute, but is a clear statistical tendency; in the Tuu and Ju languages it is stronger. The form in (3) addresses surface representations; Traill (1985: 90) in fact proposes the stronger (4).

(4) *Phonological Back Vowel Constraint*

A back consonant, including any click, may not be followed by a (phonological) front vowel.

He accounts for (most of) the exceptions to (4) by a phonetic rule which creates the exceptional front vowels from underlying ‹a› in the presence of 'front' clicks. I shall discuss this somewhat counterintuitive approach at length in §5.1; for the moment, I just state the rule as (5) (cf. Traill 1985: 70).

(5) *A-Raising Rule*

a. A first-mora plain, breathy or creaky ‹a› is raised to [ɜ] when the second mora contains ‹i›, or is a nasal, and the word starts with a dental non-click or ‹ǀ ǂ›.

b. It is further raised to [i] when the second mora is just ‹i›.

3 Click consonants: questions and analyses

3.1 Posterior place distinctions

Before turning to the question of clustering, I discuss one small controversy which interacts with it. I said above that the salient difference

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

between ‹ʞ› and ‹ʞq› is the prolongation of the posterior closure. However, Ladefoged & Maddieson (1996) describe the difference as one of velar *vs.* uvular place for the posterior closure. This description comes ultimately from Traill, who describes ‹ʞ› as velar and ‹ʞq› as uvular. He describes some of the other complex clicks as having velar articulation, and also considers the non-click ejective affricates to be phonetically and phonologically ‹kˣ'› rather the DoBeS ‹qᵡ'› that I have adopted. However, in Traill (1994) he is a little more cautious about this, and it is unclear what his final view was.

DoBeS, on the other hand, does not need to commit to the exact place of the posterior closure of clicks, and considers the complex prolonged closure clicks to be clusters with members of the uvular non-click series.

The ‹ʞ/ʞq› distinction is widespread in Khoisan, and so has been considered by other researchers. In particular, Miller *et al.* (2009) raise the question of whether it is even possible to maintain a velar/uvular distinction, and conclude that it is not. They adduce direct articulatory measurements for this – ultrasound imaging shows that clicks have a posterior constriction in the uvular to pharyngeal region, depending on the click type (see also Miller *et al.* 2007).

I have also carried out some informal experiments deliberately trying to make a velar/uvular posterior contrast (using ultrasound to check the actual articulations), and cannot convince myself that I can make such a distinction in a plain click, although with a prolonged closure it seems feasible to advance or retract the closure before release.[2] I therefore assume that no velar/uvular posterior place distinction exists in clicks, and refer to Miller (2011) for further discussion.

3.2 Features for clicks

Given their typological rarity, it is not surprising that there is no commonly agreed set of features, or even any several commonly agreed sets of features, for click consonants. Here I briefly review some of the proposals. All authors recognise the separation of click and accompaniment, so all proposals split into one set of features to distinguish the anterior closure/release, and another for the posterior release.

Both Jakobson (1968) and Chomsky & Halle (1968) considered click features. The former tried to re-use existing features, while the latter introduced a number of features into *SPE*, particularly to deal with accompaniments. I refer to Traill (1985: ch. 5) for a full description and convincing critique of these proposals. Snyman (1970) nominally uses distinctive features, but simply invents a feature for each articulatory characteristic.

[2] I am grateful to Jim Scobbie of Queen Margaret University for kindly allowing the use of their ultrasound equipment, and to Stephen Cowen for generous training and support.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Traill (1985) develops a system rather similar to Snyman's, but cleaner and better justified; however, he goes beyond standard feature theory by using contoured values for some features, such as his [friction]. He also discusses proposals to give segments internal structure, following e.g. Campbell (1974), so that the cluster phonemes can be internally split into click and accompaniment while remaining as single phonemes. In Traill (1993) he followed up on this by putting these thoughts into a formal feature-geometry setting, though he was not fully satisfied with this, and did not adopt it.

Güldemann (2001), as I discuss further in §3.3.3, is a cross-Khoisan study. His analysis emphasises hierarchical structure: he uses features that are ordered. For example, he has three distinct [stop] features: the first, high in the hierarchy, captures the difference between the nasal clicks and the rest; the second, below an [elaboration] feature, describes whether the elaboration (meaning most accompaniments) contains a separate stop in addition to the click. A subordinate [elaboration] feature describes the ejective accompaniments; and below that, the third [stop] distinguishes ‹ʞqʰ› from ‹ʞqχ'›. This is essentially feature geometry, but with added structure.

Miller-Ockhuizen (2003) works mainly at a phonetic rather than formal phonological level; she uses generally articulatory features, but in particular introduces [pharyngeal], characterising certain clicks, and the acoustic feature [spectral slope], capturing stridency and glottalisation. As I discuss in §3.3.5 below, Miller *et al*. (2009) go beyond Traill's tentative use of contoured features by introducing contoured airstream features.

In this article, the choice of features for clicks is not a primary concern. Indeed, I am not even committed to the use of features in any particular formal theory; here, it suffices to have some notion of classifying sounds. In the formal development, I will assume *SPE*-like features, and avoid discussion of the details that have vexed previous researchers.

3.3 Clusters or not?

3.3.1 *Unitary analyses*. Until the 1970s, linguistic descriptions of Khoisan languages recognised the different series of clicks, but did not analyse the accompaniments, which were then called 'effluxes' (Beach 1938). That work itself is a very thorough (and still useful) study of Khoekhoe, but Beach does not classify or analyse the accompaniments (of which Khoekhoe has only five – ‹ʞ ʞʰ ʞh ʞ̃ ʞʔ›).

Snyman (1970) takes the same approach in his study of the Ju language Ju|'hoansi, also called !Xũ. This language has the usual four ‹! | ǁ ǂ› click types, with, according to Snyman, some fourteen accompaniments.[3]

[3] Miller-Ockhuizen (2003) differs, giving twelve accompaniments. Whether this marks a difference in dialect or analysis, I do not know. Generally, Miller-Ockhuizen's analysis is substantially more complex than Snyman's.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

ʘ	9	ǀ	62	ǁ	103	ǃ	107	ǂ	69
ʘ̬	6	ǀ̬	57	ǁ̬	74	ǃ̬	70	ǂ̬	58
ʘʰ	6	ǀʰ	28	ǁʰ	30	ǃʰ	21	ǂʰ	39
ʘ̬ʰ	2	ǀ̬ʰ	5	ǁ̬ʰ	7	ǃ̬ʰ	7	ǂ̬ʰ	11
ʘχ	2	ǀχ	15	ǁχ	28	ǃχ	19	ǂχ	19
ʘ̬χ	1	ǀ̬χ	10	ǁ̬χ	13	ǃ̬χ	9	ǂ̬χ	10
ʘqχ’	1	ǀqχ’	7	ǁqχ’	17	ǃqχ’	15	ǂqχ’	8
ʘ̬qχ’	3	ǀ̬qχ’	9	ǁ̬qχ’	13	ǃ̬qχ’	4	ǂ̬qχ’	10
ʘq	5	ǀq	16	ǁq	28	ǃq	15	ǂq	35
ʘ̬ɢ	3	ǀ̬ɢ	29	ǁ̬ɢ	29	ǃ̬ɢ	35	ǂ̬ɢ	42
ʘ̬qʰ	0	ǀ̬qʰ	4	ǁ̬qʰ	7	ǃ̬qʰ	3	ǂ̬qʰ	0
ʘq’	3	ǀq’	16	ǁq’	25	ǃq’	12	ǂq’	5
ʘh	7	ǀh	13	ǁh	18	ǃh	24	ǂh	15
ʘ̰̃	1	ǀ̰̃	5	ǁ̰̃	8	ǃ̰̃	11	ǂ̰̃	3
ʘ̬̰̃	9	ǀ̬̰̃	60	ǁ̬̰̃	82	ǃ̬̰̃	104	ǂ̬̰̃	48
ʔʘ̬̰̃	2	ʔǀ̬̰̃	33	ʔǁ̬̰̃	41	ʔǃ̬̰̃	37	ʔǂ̬̰̃	15
ʘʔ	5	ǀʔ	23	ǁʔ	32	ǃʔ	28	ǂʔ	35

Table II
Click frequencies in the lexicon (Traill 1994).

Snyman explicitly presents each such consonant as a phoneme, ascribing *SPE*-style features to each.

This unitary click analysis has obvious drawbacks, which become more pressing as the number of accompaniments increases. In the case of !Xóõ, it leads to the statement that the language has 83 (attested) distinct click phonemes (per Traill), or 115 (per DoBeS), as in Table I. While few things can be said to be impossible, many people find this to be beyond the limits of what human language might be expected to maintain. There are several reasons for this. For one thing, it poses a considerable challenge to the language acquirer. This is especially so when one considers the rarity of many of the 'phonemes'. The size of the !Xóõ vocabulary is not known, but Traill (1994) lists about 3000 native words (or rather stems), of which about 2000 contain clicks. Though the true native vocabulary may be rather larger (or may have been before the enforced sedentarisation and migration in the 1980s and 90s), Traill was specifically looking for phonologically illustrative material. Nonetheless, there are three 'phonemes' that occur in only one word each – for example, the sound ‹ʘ̰̃› is supported only by ‹ʘ̰̃ã̰a› 'sit or stand close together' – and thirty that occur in fewer than ten words each, including every member of the ʘ series. Table II lists the number of words for each click sound recorded in Traill (1994).

Another indication of the functional load of each phoneme is the incidence of minimal pairs. While there is in general no reason to expect contrasts to be demonstrable between every pair of phonemes, counting

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

the total number of pairwise contrasts gives an indication of the global strength of contrasts. Taking English, for example, with its average-sized consonant inventory, more than 95% of the possible pairwise consonant contrasts are illustrated by minimal pairs, even when one only considers monosyllables.

In !Xóõ, the expected number of minimal pairs is decreased by its very large vowel inventory (as well as the non-click consonants), but increased by the very restricted shape of words: given the basically bimoraic word shape, and the various phonotactic restrictions, there are about 13,000 possible click-initial words in Trail's analysis, ignoring tone, compared to the 36,000 or so possible English monosyllables. It is perhaps remarkable that !Xóõ does have a little more than half of the 3403 unitary minimal pairs; and almost three quarters if one ignores tone.[4] Nonetheless, combined with the rarity of many unitary phonemes, one must wonder how so many distinctions survive.

If we take a more realistic approach, and only ask for each click to contrast with other clicks of the same anterior place (analogous to looking for contrasts among English /t d s θ ð n l r/), the picture is somewhat better, but still surprisingly rarefied: almost 30% of such contrasts are not supported by a minimal pair, even if we ignore tone. In English, all contrasts of manner at a given place are supported by multiple minimal pairs, even for such historically recent contrasts as /θ/ *vs.* /ð/.

3.3.2 *Cluster analysis*. Even a cursory glance at Table I invites the suspicion that at least the more complex accompaniments are really clusters. Consider, for example, the click ‹ʞqᵡ'› (row 20). Given that we see also the free-standing consonant ‹qᵡ'› (row 7), as well as the non-click ‹pqᵡ' tqᵡ' tsqᵡ'› combinations, also in row 20, the suspicion becomes practically unshakeable. Moreover, as I noted above, all these sounds vary similarly with dialect and register – [qᵡ'] (or velar [kˣ'] according to Traill) itself is a western dialect pronunciation, whereas the eastern dialect has [k'q] in citation form, with the western form in fast speech (Traill 1994: 36).

In Traill (1985), he assumes a unitary analysis, despite its 'implausibility', for most of the book, pleading reluctance to violate tradition. However, at the end of the book, he puts forward the above argument, and proposes what I shall call the CLUSTER ANALYSIS.

As can be seen from Table I, 'every one of the simple accompaniments *that forms a phonetic cluster* with a click (except possibly for delayed aspiration) exists as an independent consonant' (Traill 1985: 209; emphasis in the original). Traill therefore proposes a fairly extensive cluster analysis, in which the basic clicks are [ʞ ʞ̠ ʞ̟ ʞ̟̠], and all the others are viewed as

[4] The phonology of Traill (1994) has 83 attested click consonants, hence 3403 contrasts. The largest minimal set has 31 members, with the shape /C*àa/; ignoring tone, the set /C*aa/ has 49 members. The statements in this section about minimal pairs in !Xóõ were computed by scripts from a manually entered list of the headwords in Traill (1994).

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

clusters. This obviously simplifies the phoneme inventory dramatically: instead of $17 \times 5 = 85$ click phonemes, there are just $4 \times 5 = 20$, and all the others arise from combinations with phonemes already in the non-click inventory. It also (he asserts) has other nice effects on the phonological analysis, mostly by converting complex 'featural' rules into natural coarticulatory consequences of the components of the clusters.

This cluster analysis is not completely unproblematic. Traill mentions a couple of 'minor details', such as the awkward absence of free-standing /h/ other than in a couple of interjections; other problems arose later when in Traill (1993) he attempts to put !Xóõ in a feature-geometric framework: the durations of some clusters does not match very well with feature-geometric requirements on timing slots. Despite this, the analysis seems compelling to many.

In recent years, the cluster analysis has become quite widely accepted as the natural way to analyse Khoisan languages. I have already mentioned Güldemann's (2001) cross-Khoisan analysis, and will discuss it further below.

A recent substantial work discussing cluster analysis at some length is Nakagawa (2006). ǀGui is a Khoe language of fairly high click complexity, spoken in Botswana, with the usual four clicks, and thirteen accompaniments, which are a subset of the !Xóõ range. Nakagawa adopts a MODERATE CLUSTER ANALYSIS, based on Traill's proposals. Because, unlike Traill (1985), he recognises plain ejectives (‹ʞ'›) and aspirates (‹ʞʰ›), he includes these as basic clicks, so ending up with $4 \times 6 = 24$ click phonemes, plus $4 \times 7 = 28$ clusters.

Similarly, Naumann's (forthcoming) study of western !Xóõ also adopts a Traillian analysis, largely following and extending the moderate cluster analysis – my terms 'simplex' and 'complex' in §2.3 are chosen to match with the DoBeS view that rows 1–13 of Table I are phonemes, and rows 14–27 clusters. As well as the arguments involving parsimony and symmetry of systems, and on the grounds of the phonetic properties that I sketched in §2.3, Naumann also gives some informal observations of speaker behaviour that seem to support the cluster analysis: for example, his informants sometimes described ‹!qʰ-› words as starting with ‹!›. Under the moderate cluster analysis, the phonemes are those in (6).

(6) *!Xóõ click phonemes under the moderate cluster analysis*

ʘ	ʘ̬	ʘʰ	ʘ̬ʰ	ʘ'	ʘ̬'	ʘ̃	ʘ̬̃
ǀ	ǀ̬	ǀʰ	ǀ̬ʰ	ǀ'	ǀ̬'	ǀ̃	ǀ̬̃
ǃ	ǃ̬	ǃʰ	ǃ̬ʰ	ǃ'	ǃ̬'	ǃ̃	ǃ̬̃
ǁ	ǁ̬	ǁʰ	ǁ̬ʰ	ǁ'	ǁ̬'	ǁ̃	ǁ̬̃
ǂ	ǂ̬	ǂʰ	ǂ̬ʰ	ǂ'	ǂ̬'	ǂ̃	ǂ̬̃

3.3.3 *Güldemann's analysis.* The cross-Khoisan analysis of Güldemann (2001) is quite radical from the point of view of phonological theory.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Some of its roots lie in Traill's discussions of early notions of subsegmental structure, but Güldemann goes further. As sketched above, he uses a hierarchical structure, so that segments can combine to make bigger segments. One of his main aims is to integrate the click and non-click systems, so there is a top-level featural distinction [suction] (following *SPE*) distinguishing clicks, and then below that a hierarchy of features/subsegments. For him, 'simple' stops are the voiced and voiceless stops/clicks. Simple stops can be modulated by aspiration and glottalisation (ejectivity is treated as glottalisation for phonological reasons, such as the constraint in (2b)), to produce 'complex' stops. Either simple or complex stops can then be sequentially combined with other stops to form 'cluster' stops – which are both clusters and single segments with their own featural description.

Güldemann's discussion brings in a number of aspects of cross-Khoisan phonology, but a detailed review would take more space than is justified for the purposes of this article. Suffice it to make three observations. Firstly, he remains unable to settle firmly on the appropriate set of place features for clicks, because of some of the issues mentioned above in §3.2. Secondly, for him the !Xóõ alveolar affricate series (‹ts›, etc.) is indeed phonologically affricated, whereas Traill treats it (as I do implicitly) as an incidentally affricated series of alveolar stops. Finally, it is not entirely clear how this approach is to be integrated into formal phonological theories, whether rule- or constraint-based.

3.3.4 *The radical cluster analysis.* Nakagawa (2006: §5.3.1.3), a section in which he considers a RADICAL CLUSTER ANALYSIS, requires special mention in the context of this article. The radical cluster analysis is 'radical' in that it proposes that there is only one click phoneme at each place – which, as will be seen, is precisely the argument this article makes about Khoisan. However, Nakagawa sets up the radical cluster analysis as a straw man to justify his preferred analysis – it is germane, therefore, to explain why he argues that the radical cluster analysis fails. I will go on to argue, as the proposal of this paper, that it is in fact correct to propose such a radical analysis, but a conceptual change in the nature of phoneme and segment is required for it to work as desired.

The difficulty Nakagawa has is choosing which click is basic. An obvious first thought would surely be that the plain unvoiced click is the basic click. However, Nakagawa finds this untenable, because although |Gui has the voiced nasal click [ᵑ̃ǃ] (but not [ᵑ̃ǃ̬]), it does not have a plain velar nasal [ŋ] in its inventory with which ‹ʁ› could cluster. He concludes, therefore, that the only viable choice for the unit click in radical cluster analysis is the nasal click, with some phonetic rules to explain how it combines with other phonemes to form the other clicks – rules that have to be inelegantly restricted in their application, to avoid destroying the non-click inventory.

As a reviewer observes, it is questionable whether Nakagawa's reasons are sufficient; ‹ŋ› could simply have a defective distribution, or possibly

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

the nasal that combines with clicks is ‹n› (which is compatible with my formulation below in which click accompaniments are not specified as velar or back). However, I claim that while the radical cluster analysis is correct, a change to parallel clustering brings a number of improvements.

3.3.5 *Arguments against the cluster analysis.* Although Naumann (forthcoming) adopts the cluster analysis, he also finds some evidence weighing against it. Firstly, it is surprising that the A-Raising Rule (5) still operates following clusters with uvular stops – one would expect a uvular to block any raising effect of the previous click. Secondly, he conducted an informal onset-dropping experiment: two speakers were trained to drop the first sound of words in Afrikaans, and then asked to do the same with !Xóõ words. Neither speaker simply dropped the click from the cluster; either they dropped the entire cluster or they produced words starting with ‹h› or ‹ʔ›. My proposal will resolve both these issues (see §5.1).

Amanda Miller, whose study of Ju|'hoan was mentioned earlier (Miller-Ockhuizen 2003), has recently been working with a number of colleagues on the almost extinct language N|u. Although in that work she follows a cluster analysis, Miller *et al.* (2009) argue that cluster analyses are wrong. Instead, they propose to extend the range of features by which clicks are classified, and in particular to add contour values for the airstream feature. These are to simple airstream values as affricates are to stops and fricatives. N|u has a mid-sized range of accompaniments, which, adapting Miller *et al.*'s notation to ours, are ‹ʞ ʞʰ ʞ̧ ʞ̧ʔ ʞ̧ʰ ʞ̧ ʞq ʞqʰ ʞχ ʞχ'›. (N|u ‹ʞ̧ʰ› does not appear to have such a markedly long crescendo aspiration as !Xóõ [ʞh]. The ‹ʞχ'› is probably what I represent as ‹ʞqˠ'›.)

The way that Miller *et al.* classify these clicks by 'airstream mechanism' is as follows: (i) the simple and nasal clicks ‹ʞ ʞʰ ʞ̧ ʞ̧ʔ ʞ̧ʰ ʞ̧› are said to have just a lingual airstream; (ii) the clicks ‹ʞq ʞqʰ ʞχ› are said to have a 'linguo-pulmonic' airstream, reflecting their status (as in the similarly notated !Xóõ clicks) as moving from a click into a normal pulmonic release, with a clearly audible [q qʰ χ]; (iii) similarly, the click ‹ʞχ'› is said to have a 'linguo-glottalic' airstream.

From the phonetic point of view, this classification allows us to add the click consonants to the standard IPA chart by extending it with new sections for the different values of airstream feature. So we have a block of pulmonic consonants, followed by a block of lingual consonants, followed by a block of linguo-pulmonic, and so on. A concrete motivation for this concerns the difference between ‹ʞ› and ‹ʞq›, a distinction shared by !Xóõ and N|u. As discussed in §3.1, Miller *et al.* consider (and I agree) that there is no role for velar/uvular place in the contrast; therefore there is only a timing difference, and this is best seen as a contoured airstream.

From our perspective, this is still a unitary analysis, but with different feature values for the various accompaniments; it does not change the number or identity of phonemes in the unitary analysis.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Miller's more phonological arguments for this analysis are laid out in Miller (2011). Two of the major arguments are the difficulty of decomposing all clicks into segments that also appear independently (as noted by Nakagawa; see above) and the fact that typologically every language that allows obstruent–obstruent clusters also allows obstruent–sonorant clusters, whereas there are none of the latter in Khoisan languages. My proposal will address both these points (see §6.2).

4 Concurrent phonemes

4.1 The concurrent analysis

After this survey of the facts and current analyses, my proposal here can be very simply stated. Namely, *every* click is indeed a cluster. In the case of the basic clicks, the two component segments are the click influx and the accompaniment. Since there is no sequential order between these two components, they are clustered not serially, but concurrently. In IPA notation, this might be written, for example, ‹ʘ͡ɣ›; unfortunately, the tie-bar is widely used to denote a phonetic coarticulation that forms a single phonemic unit, which is exactly not my point. I shall borrow a computer science notation (one of many for the concept) and write ‹(!⊗ɣ)›, where it is stipulated that this is identical to ‹(ɣ⊗!)›.

Such an analysis brings the advantages of the radical cluster analysis, or even of Güldemann's structured cluster analysis, while retaining most of the simplicity of standard segmental and phonemic theories. Formally, it is straightforward enough to be easily incorporated into any theory that works with segments and phonemes.

4.1.1 *Concurrent clicks in !Xóõ.* If we apply this idea to the !Xóõ click inventory (I call this the CONCURRENT ANALYSIS), we obtain a dramatic simplification and reduction. The five clicks become phonemes in their own right; and we can now reinterpret our phonetic meta-notation for accompaniments, such as ‹ʁq›, in which the ‹ʁ› is really a variable ranging over the five click symbols, into a true phonetic and phonemic notation, in which ‹ʁ› is not a variable, but a novel phonetic symbol to indicate the point at which this sequence of segments synchronises with any concurrent click segment. The phonetic output now follows from common phonetic rules: ‹!q› is phonemically ‹(!⊗ʁq)›, and an unexceptional phonetic rule unifies the posterior closure required by the click with that required by [q], resulting in a long uvular stop with a click at the beginning.

Thus, even if we retain all 23 unitary accompaniments, the click-inventory size is now 5 + 23 instead of 5 × 23, as set out in (7). Instead of an exceptionally large array of consonants, we have a modest set, with the formerly apparent complexity being simply clustering. Apart from the fact that the clustering is concurrent rather than sequential, it is no more exceptional than, say, clusters in Russian.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

(7) *!Xóõ click phonemes under the concurrent unitary analysis*

ʞ	ʞ̠	ʞʰ	ʞ̠ʰ	ʞˀ	ʞ̠ˀ	ʞ̠̃	ʞ̠̃	ˀʞ̠̃
ʞq	ʞ̠G	ʞqʰ	ʞ̠qʰ	ʞqˀ	ʞ̠qˀ	ʞqχˀ	ʞ̠qχˀ	ʞχ	ʞ̠χ	ʞh	ʞ̠h	ʞʔ	ʞ̠ʔ
				ʘ	ǀ	ǃ	ǁ	ǂ					

Moreover, all the arguments for a sequential cluster analysis within accompaniments hold just as well in this setting as they do in the traditional setting. The moderate cluster analysis, for example, naturally becomes what might be called a concurrent moderate cluster analysis. Now there are five clicks and eight accompaniments, as in (8), and all the rest is clustering, both concurrent and sequential: for example, the click ‹ǃqˀ› can be analysed as /(ǃ⊗ʞ̠qˀ)/. In this analysis, !Xóõ has only thirteen click phonemes. For good measure, the arguments against clustering outlined at the start of §3.3.5 no longer obtain: since the 'onset' of a word is now a concurrent cluster, it is not surprising that speakers had difficulty deciding how to drop it; and we shall see soon how the failure of uvulars to block A-Raising emerges naturally.

(8) *!Xóõ click phonemes under the concurrent moderate cluster analysis*

ʞ	ʞ̠	ʞʰ	ʞ̠ʰ	ʞˀ	ʞ̠ˀ	ʞ̠̃	ʞ̠̃
	ʘ	ǀ	ǃ	ǁ	ǂ		

If one adopts Miller's proposal (§3.3.5), which is a unitary analysis, one can still adopt the concurrent analysis: at the phonological level, ‹ʞq› will be an accompaniment with a linguo-pulmonic airstream, which then combines with a phonological pure click to produce her phonetic 'linguo-pulmonic' consonant.

4.1.2 *A formal implementation.* I intend this proposal as one of basic linguistic theory (Dixon 1997), since it can be understood in any framework, formal or informal, that supports the notions of phoneme and segment. To demonstrate a precise implementation, I give a version in a variant of *SPE*. I shall use unspecified features in phonemes, rather than go down the formal route of *SPE* markedness theory – it is a routine but unenlightening exercise to recast everything in strict *SPE*. I use *SPE* notation for rules; unspecified features are written as e.g. [∅voice].

For theories such as Optimality Theory (Prince & Smolensky 1993) which also use a feature-based phonemic representation, it is similarly straightforward to add concurrency; and all the rules I consider can be readily translated to ranked constraints.

Recall that in *SPE* there is a set of binary features, underlying representations (URs) are strings of feature bundles, which may be unspecified for some features, and the output of the rewriting rules is a string of fully specified feature bundles. Despite Chomsky & Halle's express discouragement of such terminology, one can say that 'phoneme'

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

corresponds to a feature bundle in the UR, and 'segment' to a bundle in the output, and I will adopt this henceforth. I assume that features for clicks are as in (14)–(17) below, so that clicks share a feature [+lingual], and all the usual non-click phonemes are specified [−lingual]. The first step is to extend the strings in the URs, as in (9).

(9) *Cstring*

A phoneme is a one-element CSTRING ('concurrent string'). There is a commutative and associative binary combinator $\otimes$ on cstrings. Cstrings may be combined with $\otimes$ and concatenation. We let concatenation have higher precedence than $\otimes$ (i.e. $a \otimes bc$ means $a \otimes (bc)$, not $(a \otimes b)c$). The empty cstring ε is the identity for $\otimes$ (i.e. $a \otimes \varepsilon = a$). Every UR is a cstring.

Note that I will use parentheses with the usual mathematical meaning of grouping. This is potentially confusable with the *SPE* use of parentheses to indicate optional elements in rules, but in practice it will always be clear from context which meaning a given parenthesis has.

Definition (9) by itself allows arbitrary combinations; as concurrency is intended to reflect the physical possibility of combining different sounds, I impose (10).

(10) *Weak Concurrent Airstream Constraint*

In any UR containing a sub-cstring $a \otimes b$, the phonemes in a may not have contradictory values for [lingual]. (And by commutativity, the same holds for b.)

The effect of (10) is to forbid clicks and non-clicks to combine within one half of a concurrent composition. For the moment, I also stipulate (11).

(11) *Strong Concurrent Airstream Constraint*

In any UR containing a sub-cstring $a \otimes b$, if a contains a phoneme with a specified value for [lingual], then b may not contain a phoneme with that value.

(11) further restricts $\otimes$ to combining clicks on one side with non-clicks on the other. I now define the click phonemes.

(12) a. *Pure click phonemes*

The phonemes /ʘ ǀ ǃ ǁ ǂ/ are lingual obstruents with features as in (14).

b. *Accompaniment phonemes*

The accompaniment phonemes are specied for laryngeal and manner features (only), as in (15). They are notated by /ʁ/, together with diacritics for the positive features.

This is the definition that turns our accompaniment notation ‹ʁ› into a symbol for an actual phoneme. Now /ʁ ʁ̬ ʁʰ/, etc. are genuine phonemes

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

in the inventory, albeit with the unusual phonotactic constraint (which can be dispensed with, at least formally – see §6.1) that they occur only in concurrent clusters. This constraint is formulated as (13).

(13) *Click/accompaniment Constraint*
A UR may contain a [+lingual] phoneme *x* only if *x* is in a sub-cstring *a* of $a \otimes b$ such that *b* contains a [∅lingual] phoneme, and conversely.

This constraint forbids pure clicks and pure accompaniments from appearing by themselves in URs.

(14)–(17) sets out the featural specifications I assume in the discussion following, both for the click phonemes and for the other phonemes of !Xóõ. Some choices are of course a little arbitrary; others are justified in the following sections.

The pure clicks /ʘ ǀ ǃ ǁ ǂ/ are specified for [+consonantal, −vocalic, −continuant, +lingual], together with the features in (14).

(14)

	[ant]	[cor]	[high]	[back]	[del rel]
ʘ	+	−	−	+	+
ǀ	+	+	+	−	+
ǃ	+	+	−	+	−
ǁ	−	−	+	−	−
ǂ	+	+	−	+	+

The pure accompaniments are specified for the features in (15).

(15)

	[voice]	[nas]	[spr gl]	[glot cl]
ʁ	−	−	−	−
ʁ̬	+	−	−	−
ʁʰ	−	−	+	−
ʁ̬ʰ	+	−	+	−
ʁˀ	−	−	−	+
ʁ̬ˀ	+	−	−	+
ʁ̃	−	+	−	−
ʁ̬̃	+	+	−	−

Manner features for the pulmonic stops are specified as for the accompaniments, using [voice], [spread glottis] and [glottal closure], together with [+delayed release] for the alveolar affricated stops and the uvular ejective affricates /qᵡˀ ɢᵡˀ/. Place features are as in *SPE*, with one exception: we distinguish dentals /t d …/ from alveolars by [high] (motivated largely by the raising behaviour of dentals described in §5.1), as in (16).

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

(16)

	[ant]	[cor]	[high]	[low]	[back]
p	+	−	−	−	−
t	+	+	+	−	−
ʦ	+	+	−	−	−
ɲ	−	−	+	−	−
k	−	−	+	−	+
q	−	−	−	−	+
ʔ	−	−	−	+	−

Continuants, glides, liquids and nasals are as in *SPE*; I tentatively consider the glottalised nasals to be clusters: /ʔm ʔn/.

Vowels are standard, except that /a/ is unspecified for [back], as in (17).

(17)

	[high]	[low]	[back]	[round]
i	+	−	−	−
e	−	−	−	−
a	−	+	∅	−
o	−	−	+	+
u	+	−	+	+

Creaky vowels are [+glottal closure], breathy vowels are [+spread glottis], pharyngealised vowels are [+pharyngeal] and strident vowels are [+pharyngeal, +spread glottis].

Now the 'simplex' !Xóõ clicks have underlying representations such as /(!⊗ʁ̞ʰ)/. The question remains of the 'complex' clicks. As I discuss later, there is room for manoeuvre here. For the moment, I assert that ‹!qʰ›, for example, has the UR /(!⊗ʁqʰ)/: that is, it is a concurrent cluster, one half being the pure click, and the other being a sequence of ‹ʁ› and ‹qʰ›, as shown in (18) for the alveolar clicks.

(18) *The representations of clicks in the concurrent moderate cluster analysis*

!	(!⊗ʁ)	ʔ̰͡!	(!⊗ʔʁ̞)	!qχ’	(!⊗ʁqχ’)
!̰	(!⊗ʁ̞)	!q	(!⊗ʁq)	!̰qχ’	(!⊗ʁ̞qχ’)
!ʰ	(!⊗ʁʰ)	!̰ɢ	(!⊗ʁ̞ɢ)	!χ	(!⊗ʁχ)
!̰ʰ	(!⊗ʁ̞ʰ)	!qʰ	(!⊗ʁqʰ)	!̰χ	(!⊗ʁ̞χ)
!’	(!⊗ʁ’)	!̰qʰ	(!⊗ʁ̞qʰ)	!h	(!⊗ʁh)
!̰’	(!⊗ʁ̞’)	!q’	(!⊗ʁq’)	!̰h	(!⊗ʁ̞h)
!̰̃	(!⊗ʁ̞̃)	!̰q’	(!⊗ʁ̞q’)	!ʔ	(!⊗ʁʔ)
!̰̃	(!⊗ʁ̞̃)			!̰ʔ	(!⊗ʁ̞ʔ)

To complete the formalisation, I need to consider whether concurrency survives to the output stage of the *SPE* rewriting process. One may have

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

different views on this, according to where one prefers to draw the phonology/phonetics boundary. My preferred approach is to leave the click concurrency in the output, but to resolve the complex clustering, by adding rule (19) late in the *SPE* rule chain.

(19) *Lingual Synchronisation Rule*

$$[+\text{lingual}] \otimes [-\text{lingual}]_0 \; [\emptyset \text{lingual}] \; [-\text{lingual}]_0$$
$$1 \qquad\qquad 2 \qquad\qquad 3 \qquad\qquad 4 \qquad\qquad \rightarrow \quad 2 \; (1 \otimes 3) \; 4$$

This is *SPE* notation for 'a [+lingual] phoneme docks on to a lingually unspecified phoneme in the other concurrent half'. For example, $/(! \otimes ʁ q^h)/ \rightarrow [(! \otimes ʁ) q^h]$ by this rule, with 1 = /!/, 2 empty, 3 = /ʁ/, 4 = $/q^h/$.[5]

(19) is one of several variations on the technical devices that could be employed to achieve the effect of synchronising clicks with the pulmonic airstream sounds; this one is natural because of the intuition it gives for /ʁ/ being a manner-carrying placeholder waiting to receive a click.

One might wish to eliminate the idea of concurrent segments from the output. This can be done by adding a later rule, (20).

(20) *Concurrent Fusion Rule*

$a \otimes b \rightarrow a \sqcup b$

where $a \sqcup b$ is the phoneme whose specified features are the union of those of a and b – it is undefined, and the rule cannot apply, if a and b have inconsistent values for some feature.

The $\sqcup$ operation is not standard *SPE* notation, but has been recently suggested as a useful addition by Bale *et al*. (2013); the rule can of course be written out in standard notation, but is lengthy. The result of applying this rule to $/(! \otimes ʁ) q^h/$ is the purely sequential cluster $[\overset{!}{\underset{s}{}} q^h]$, where $[\overset{!}{\underset{s}{}}]$ has all its features specified.

4.2 Discussion

4.2.1 *Concurrent segments and phonemes: a natural concept.* The first question is whether, as I suggested in the introduction, the notion of concurrent segments and phonemes is consistent with the traditional informal understanding of segments and phonemes. In basic linguistic theory, the phoneme is still largely defined by structuralist considerations,

[5] The rule as formulated allows only one click to dock on a given /ʁ/. It could be formulated to allow a concatenation of clicks as component 1: it is perfectly possible to make an arbitrarily long sequence of clicks while maintaining the posterior closure. However, no language makes use of this possibility. It could also be formulated to allow a sequence of clicks to dock on to a sequence of accompaniments; but again, I know of no reason to do this.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

and the notion of segment is taken as something which we naturally extract from our representations – although, as I remarked, there is not necessarily agreement about what is or is not a single segment. If we look at clicks, and try to identify segments without preconceptions, I would argue:

(i) *Articulatorily*, the click influx is a clearly identifiable gesture, whose only necessary relation with the accompaniment is that it happens during a period of velar closure.

(ii) *Acoustically*, the anterior release is very obvious in its own right, both to any human listener and on a spectrogram. On the other hand, the accompaniment is easily recognised from a spectrogram, and, I would argue (not least from my own experience), easily heard in its own right by human listeners.

The latter claim is supported:

(i) *Perceptually*, the results of Best *et al.* (2003) suggest that click place is perceived independently of accompaniment: Zulu speakers discriminate !Xóõ click places they know, and assimilate !Xóõ click places they don't know, regardless of a non-Zulu accompaniment. It is also my own experience in learning to discriminate between !Xóõ clicks, at least once I had learned to hear clicks as speech. In addition, below I cite some evidence from the !Xóõ lexicon which also suggests perceptual orthogonality.

(ii) Moreover, it appears that *in production* speakers of click languages can immediately combine newly learned clicks-in-isolation with the accompaniments they already know. To my knowledge this has not been demonstrated before, and so I describe the relevant pilot experiment in the following subsection, and discuss this argument further.

Thus I claim that the notion of concurrent segment is well supported; and if the click and its accompaniment are both segments, they are certainly both phonemes by the usual contrast criterion.

4.2.2 *A click-production experiment.* If, as I claim, clicks are separate phonemes from accompaniments, then if one takes a speaker of a click language, and teaches them a new click by itself, it should be the case that if they can use the new click in words at all, they can, without further instruction, use it with all their native accompaniments. If, however, clicks are not so decomposable, then generalising to all accompaniments involves conscious featural manipulation, which is held by many to be outwith the competence of untrained speakers. There is a considerable debate about such claims, but it seems plausible that manipulating phonemic segments is at least easier than manipulating features, despite such examples as the difficulty of pronouncing clusters that are not in one's own language.

Here I report a pilot experiment which aims to test this prediction. Though there were only a couple of participants, the results are interesting and suggestive. I hope to seek support for a full version of this experiment in cooperation with colleagues elsewhere.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

The participants were young adult Nguni speakers, one Zulu and one Xhosa.[6] In my terminology, these languages have three clicks, ‹! ǀ ǁ›, written *q*, *c*, *x*. There are five accompaniments, ‹ʞ ɡ̈ ʞʰ ŋ̈ ɴ̈›, written *q*, *gq*, *qh*, *nq*, *ngq*, for example. (Xhosa also has a glottalised nasal ‹ŋ̈ˀ› (*nkq*), but my participant did not recognise my example words for it.) The two breathy accompaniments have several cues: there is breathy voice during the click, the following vowel is somewhat breathy, and perhaps most importantly, they depress the tone of the following syllable.

The first participant had no linguistic training at all. The second participant had had some exposure to introductory linguistics, mainly in semiology; in debriefing, he appeared to be unaware of standard phonological and phonetic descriptions of Nguni clicks.

The participants were first asked to demonstrate the fifteen unitary click phonemes, by reading single words presented in standard orthography (e.g. *ukugcoba*). By chance, one or two of the words were unfamiliar to each participant, and the first speaker had a little difficulty reading out an unknown word, whereas the second read easily from orthography in any case.

They were then taught, by demonstration, [ʘ] and [ǂ] in isolation, and then asked to read nonce words, presented in orthography with the IPA click symbols (e.g. *ukuʘhele*).

The first speaker had a little difficulty incorporating ‹ʘ› into words, and took several attempts at some, but produced (entirely without prompting) the accompanied versions as expected. For example, her rendition of *ingʘabha* showed prenasalisation, murmur and lowered tone. With ‹ǂ›, she read fairly smoothly, and apart from intrusive prenasalisation while hesitating on the first (plain click) word, the results were again as expected. (On subsequent review, I suspect that some of the renditions were the retroflex rather than palatal click; however, the accompaniments were not affected.) Recording quality was not very good, but illustrative spectrograms of some of her native and new clicks are shown in Fig. 1.

The second speaker found it very difficult to produce ‹ʘ› in words, and after several attempts, this part of the experiment was abandoned. With ‹ǂ›, he read fairly easily, and produced as expected. However, he informed me that ‹ǂ› was already known to him, as in his community it is used as a ‘softer’ version of ‹!› in play language and when talking affectionately to children, so all he had to do was read the nonce words as if talking to a child.

In summary: one speaker successfully produced two previously unfamiliar clicks in all of her native accompaniments; the other speaker did so with one click, but it was already familiar as a (previously unreported, to my knowledge) stylistic variation. However, the very fact that a conscious

[6] I thank Mabutho Shangase and an anonymous colleague for their kind participation.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

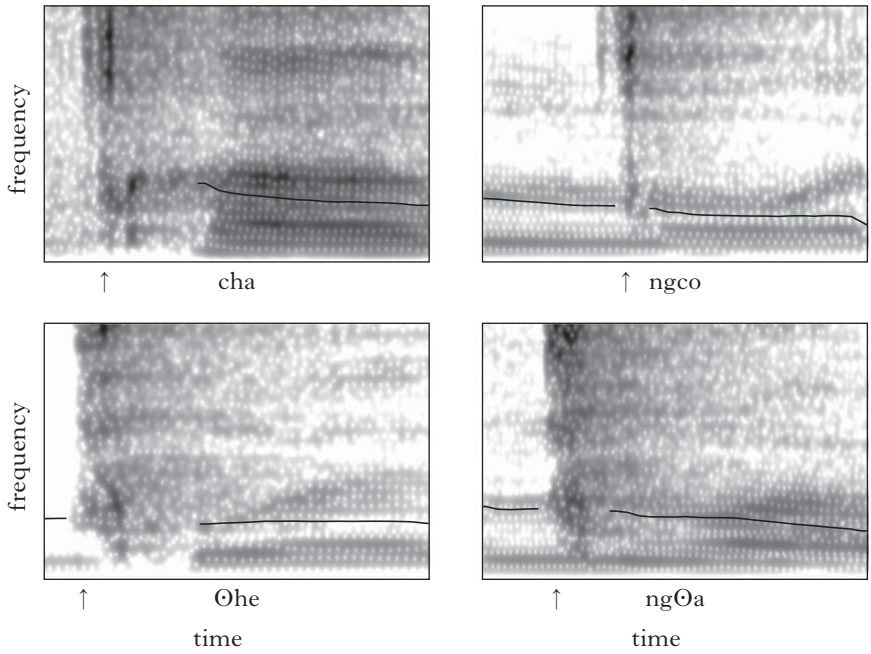


Figure 1

Productions of native (top row) and novel (bottom row) aspirated and breathy nasal clicks, from the orthographically presented words *ukuchaza, ukungcola, ukuΘhele, ingΘabha*. Pitch contour marked, with *y*-axis from 75 to 500 Hz. Samples are 250–400 ms wide; the location of the click burst is marked. The pitch contour interruption in the breathy nasals is probably the analysis being overwhelmed by the click burst. Analysis and rendering by Praat (Boersma & Weenink 2013).

stylistic variation consistently replaces one click with another across all accompaniments is itself supportive of the hypothesis.

It is also worth remarking that in debriefing, both participants were adamant that Zulu/Xhosa has three click consonants, and that, for example, *gq* is *q* combined with *g*. It would be interesting to see whether a speaker uninfluenced by orthography would say the same.

4.2.3 *Concurrent phonemes* vs. *autosegments*. In the original development of autosegmental theory, particularly as elaborated by Goldsmith (1976), it was conceived as having segments on different tiers, for example the usual phones/phonemes on one tier, and tones on another. Subsequent work looking at the melodic rather than prosodic content of speech moved towards identifying tiers with features (or with elements in the Government Phonology school), so giving a simple and natural account of, say, vowel harmony. Consequently, in such theories

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

both segments and phonemes are emergent, not stipulative concepts, arising from the associations between feature (or element) tiers and the skeletal tier: a (phonological) segment is the bundle of autosegments associated with a particular skeletal point, and the set of phonemes – insofar as the theory admits a notion of phoneme – is simply the set of such segments.

There are several differences between such an approach and my proposal here. In autosegmental theory, the tiers exist throughout the utterance, and are specified with features; the synchronisation between them is effected by association lines. Formally speaking, an autosegmental representation has the form of a parallel composition of a fixed number of sequential tiers, together with synchronisation information; multiple representations of this type may be concatenated sequentially, but then there have to be rules extending the synchronisation to the concatenation.

In my approach, however, concurrent and sequential composition act on the same entities, namely phonemes, and can (in principle) be composed with more complex nesting, although in the !Xóõ example I imposed constraints to restrict it. Because the entities being composed are phonemes, not features on tiers, they have to be justified as existing with contrastive power in the phoneme inventory of the language.

It is, of course, possible to do some formal encoding: we could analyse Finnish to have abstract phonemes /ɑ o u/ and /F/ (for Front), and assert that Finnish /y/ is really /(u⊗F)/, and then formulate harmony rules. However, to do that, we would have to argue that /F/ is a phoneme in the inventory according to the criteria above. Moreover there is no principled reason for choosing /F/ rather than /B/ (for Back) as the 'phoneme'. If we choose /F/, then we must argue either that /i/ and /e/ do not contain /F/, despite having all the same acoustic and articulatory signs of it as the other front vowels, or that they do contain it, but that there is a very specific phonotactic rule preventing /ɯ/ and /ɤ/ from occurring without it. (Note that we did claim above that click accompaniments do not occur on their own; but firstly they form a natural class, and secondly, it is at least formally possible to avoid this constraint – see §6.1 below.) If, instead, we choose /B/, we have to explain why /(i⊗B)/ does not appear – again requiring an ad hoc rule.

In summary, modern autosegmentalism deals with the structure inside segments, whereas the approach here deals with structures built out of segments. However, as I remarked at the beginning of this section, the earliest autosegmental phonology did allow for tiers to contain segments rather than features, and in that sense the proposal here can be seen as similar to it. Ladd (2014: ch. 1) contains a more substantial discussion on the historical and current relationships between concurrency, simultaneity and autosegmentalism.

It is possible to modify current autosegmental theories in such a way that my notion of concurrency here is added, above and beyond the built-in notion of tiers. However, a full development of this would occupy

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

some pages in a fairly detailed analysis, which is beyond the scope of this article.

4.2.4 *The combinatorial argument*. My claim that clicks and accompaniments are phonemes suggests that they should combine freely, modulo any phonotactic constraints, of which there appear to be none. This raises the question, which requires field investigation, of the gaps in the inventory. Traill heard no occurrence of the clicks ‹ʘqʰ ǂqʰ› over his thirty years of fieldwork. If, as seems to be the case, they do not exist in any word, then from a unitary viewpoint it is hard to argue that they exist as sounds in the language. One would therefore expect that if presented with a nonce word containing them, speakers would fail to recognise the sound correctly, and probably replace it by the nearest extant sound. On the other hand, if the clicks are independent of the accompaniment, one would expect the nonce word to be perceived and repeated with no difficulty. Christfried Naumann (personal communication) concurs that the expected result is the latter, but such an experiment has not yet been carried out. It would be even more compelling in the case of N|u: for !Xóõ, the non-concurrent cluster analysis would yield the same result, but N|u appears to be missing even some basic labial clicks, namely ‹ʘʰ ʘ̃ʰ ʘ› (Miller *et al*. 2009). Although I have not been able to test this hypothesis in the field, it is supported by the result of the experiment reported in §4.2.2.

Another combinatorial argument relates to the difficulty of learning. As I remarked in §3.3, the huge unitary analysis inventory makes it very hard to establish contrasts; but even the moderate cluster analysis leaves many contrasts without strong evidence. Obviously, a reanalysis such as the concurrent analysis that separates clicks and accompaniments solves these problems – an accompaniment contrast in the context of one click suffices to establish the contrast in the context of any click. For example, there is no support for the contrast between ‹ʘ̃› and ‹ʘ›; but if these are actually /(ʘ⊗x̰)/ and /(ʘ⊗ɣ̰)/, then the evidenced contrast between ‹ǃ̃› and ‹ǃ› also supports this contrast (and those in all the other click places).

It is no surprise that in the concurrent analysis, even without doing sequential clustering, most of the minimal pairs exist; the exceptions give rise to an interesting observation, to be discussed in §5.2.

4.2.5 *Metalinguistic evidence*. Some psycholinguistic evidence supporting my thesis comes from the !Xóõ lexicon. It turns out not only that clicks are very salient for non-speakers, but also for speakers: so much so that there are words for making the sound of the five basic clicks, and even a word for one variation particularly used in ritual incantations. So important are clicks that some of these words also mean simply ‘to talk about, converse’. The words are given in (21), in their full pseudo-reduplicated form.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

(21) a. ‹ʘ̃úũ-ʘ̃úũ› ‘to make the sound of the [ʘ] click’
b. ‹ǀhèẽ-ǀhèẽ› *or* ‹ǀʔèe-ǀʔèe› ‘to make the sound of the [ǀ] click’
c. ‹!hèẽ-!hèẽ› ‘to make the sound of the [!] click’
d. ‹ǃ̰àˤĩ-ǃ̰àˤĩ› ‘to make the sound of the [!¡] click, to talk about’[7]
e. ‹ǁhèẽ-ǁhèẽ› *or* ‹ǁ̃àa-ǁ̃àa› *or* ‹ǁʔàã-ǁʔàã› ‘to make the sound of the [ǁ] click’
f. ‹ǂhèẽ-ǂhèẽ› *or* ‹ǂʔèẽ-ǂʔèẽ› ‘to make the sound of the [ǂ] click, to talk about’

It is immediately striking that none of these words for clicks uses the plain unadorned click, at least in the unitary analysis. Even in the usual cluster analyses, the nasal clicks are viewed as primitive, and so some of these words do not contain plain clicks. In the concurrent analysis, of course, they all do. While this is not a topic on which there is extensive empirical evidence, it seems more plausible for a language to have iconic words for phonemes than for either a phonetic component of phonemes or a class of phonemes.[8]

5 !Xóõ phonology under the concurrent analysis

5.1 A-Raising and the Back Vowel Constraint

The formal development of the concurrent analysis above defined the representation of clicks, and showed some examples of rules involving concurrent clusters. Rules that do not involve concurrent clusters are unchanged, but the question arises of whether such rules need to be extended. For example, a rule might refer to properties of the first phoneme of a word – if a word starts with /(!⊗ɣ̞)/, what are those properties? The general form of such a rule in *SPE* is given in (22), where *x* specifies a class of phonemes and *y* specifies the modification to the phoneme matched against *x*.

(22) $x \rightarrow y \;/\; \#__$

In the concurrent analysis, this rule will not match a word-initial concurrent cluster – we must explicitly allow for this. For example, (23a) is

[7] Traill actually notes this as ‘the noisy [!!] click’; I am sure that by this he means [!¡], the flapped click.

[8] In this connection, it is interesting that early 20th-century researchers such as Beach and Doke used distinct letters for voiced, voiceless and nasal clicks, just as is done in the IPA for pulmonic sounds. For example, [!] was [ʗ], and [ǃ̰] was [ɜ]. The IPA adopted the plain symbols, but not the others; and then in 1989 it changed to the Africanist symbols (despite the great violence they do to the IPA’s typographic coherence). Possibly the resistance to distinct symbols was subconsciously reinforced by the reluctance to disguise the presence of the click phone itself.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

the same rule, modified to apply both to initial normal segments and to initial simplex accompaniments (assuming the concurrent airstream constraints in (10) and (11)), but not to initial clicks. A more economical notation exploiting $\varepsilon \otimes x = x$ and also allowing complex accompaniments is given in (23b).

(23) a. $x \to y \,/ \left\{ \begin{array}{c} \#__ \\ \#\,([+\text{lingual}] \otimes __) \end{array} \right\}$

b. $x \to y \,/\, \#([+\text{lingual}]_0^1 \otimes __\, C_0)$

Thus a rule may refer to the initial phoneme, or to the first phoneme of an initial concurrent cluster, as the evidence requires. The Back Vowel Constraint (3) and A-Raising Rule (5) provide good examples of this.

5.1.1 *Moderate A-Raising*. Recall that the first part of the A-Raising Rule (5a) raises ‹a› to [ɜ] before ‹i›, ‹Ci› or a nasal, *and* after a dental non-click or a dental or palatal click. This rule applies even in a word like ‹ǀq'án-tá› [ǀq'ɜ̀n-tá] 'small (PL)', showing that the rule targets the click rather than the accompaniment: the apparently intervening uvular, which one would normally expect to block a phonetic raising effect, does not do so. In the formal presentations that follow, I shall mostly omit the raising after dental non-clicks; this is merely to simplify the notation.

This rule provides the evidence for how we should distribute concurrent and sequential clustering. *A priori*, it is possible that ‹ǀq'àn› could start with /(ǀ⊗ʁq')/ or with /(ǀ⊗ʁ)q'/. Indeed, one could even analyse ‹ǀq'àn› as /(ǀ⊗ʁq'àn)/, and since Khoisan languages allow only one click per stem, this would make some sense from an autosegmental viewpoint. On the other hand, considerations of simplicity and economy suggest that (⊗) should be applied with the smallest scope, so that all of each half is genuinely concurrent with all of the other half, so favouring /(ǀ⊗ʁ)q'/. However, the behaviour of the A-Raising Rule suggests that /(ǀ⊗ʁq')/ is correct.

For the moment, I ignore the question of what it is that the triggering click types have in common, and just list them in rule (24).[9]

(24) *Formal moderate A-Raising Rule*

$$\begin{bmatrix} \text{V} \\ +\text{low} \\ -\text{phar} \end{bmatrix} \to [-\text{low}] \,/ \begin{pmatrix} \{\langle \text{ǀ} \mp \rangle\} \\ \otimes \\ C_0 \end{pmatrix} __ \begin{Bmatrix} C_0\,[\text{V}, +\text{high}, -\text{back}] \\ [+\text{cons}, +\text{nas}] \end{Bmatrix}$$

[9] *SPE* does not have a suitable feature for expressing pharyngealisation of vowels, so I use the ad hoc feature [pharyngeal]. I assume that /a/ is specified as [+low], and is unspecified for [back], so that raising it gives a mid vowel. This is purely for expressibility in the illustrative *SPE*-based framework; I would prefer a formalism with more gradience.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Formally, there is little difference between this and the equivalent rule in a standard cluster analysis, where the click context would be expressed as the class of dental and palatal simplex clicks followed by C_0, instead of a concurrent cluster of the two pure clicks with the accompaniments. Assuming all the constraints and rules in §4.1.1, it can be shown that any set of constraints and rules in this concurrent formalism can be translated into a standard set that will produce the same output; I am adding not expressive power, but naturalness. Here, we avoid the rather peculiar situation in sequential analyses of the raising power of the clicks passing through uvular stops (which one expects to be strongly lowering), because here the target vowel is immediately adjacent to both the click and the accompaniment.

The transparency of the /C/ in /-Ci/ requires a little comment – why is it transparent to the raising power of ‹-i›, while (I claim) a sequential uvular should block the licensing from the [+high] clicks? One could invoke theories that account for VV interactions as long-distance (e.g. Germanic umlaut), while requiring strict adjacency for CV interactions (e.g. English palatalisation). However, there is a simpler argument: the only permissible Cs are /b m n ɲ j l r/, all of which are either [+high] or do not involve the tongue at all – and the nasals cause raising in any case.[10]

5.1.2 *Full A-Raising*. The rules become more interesting when we consider Traill's account of the Back Vowel Constraint in eastern !Xóõ and the exceptions to it. Recall that his version of the Back Vowel Constraint (4) forbids front vowels after any back consonant, including all clicks – arguing that since clicks involve a velar/uvular closure, they are surely at least as back as ‹k›. One exception involves just ‹k›: there is a grammatical particle ‹kV›, which appears as ‹ke ki› in some concords. Traill notes that ‹ke ki› are often pronounced instead as ‹te ti›, so obeying the constraint phonetically. The other class of exceptions involves the clicks ‹ǀ ǂ›, where phonetic front vowels do appear, for example in ‹ǂìi› 'steenbok' and ‹ǀîi› 'to be'. Traill accounts for most of these by asserting that they are underlyingly, e.g. ‹ǂàì›, with both parts of the A-Raising Rule (5) applying to change ‹a› to ‹i›. The evidence for this is partly internal: the plural of ‹ǂíi› is ‹ǂàbatê›, with the morphology in (25).

(25) a. ‹ǂì› ‹-ì›
steenbok CL.1SG

b. ‹ǂà› ‹-bà› ‹-tê›
steenbok CL.1PL PL

[10] The reader may recall that final ‹ŋ› exists in the DoBeS inventory, and wonder whether it is included as a raising nasal. The phonemic status of ‹ŋ› is somewhat shaky – it may just be an allophone of ‹n› – but in the instances in the DoBeS dictionary in which it appears in A-Raising position with an audio recording (e.g. ‹ǃaŋ›), the vowel is indeed raised.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

where ‹-tê› is the current productive pluraliser. There is also cross-dialectal evidence: for example, in the DoBeS data, 'steenbok' is ‹ǂâi›, pronounced [ǂâi], with moderate A-Raising. Indeed, although Traill abandons ‹ǀîi› 'to be' as an unexplained exception, a referee points out that DoBeS has what may be the same verb ‹ǀâi› 'stay, be at a place', so even that can be accounted for.

I have not so far given a precise specification of the preceding context in full A-Raising. Traill (1985: 70, 1994: 40) is not explicit about whether it is triggered by any dental or palatal click, or just some of them, for example the plain clicks. However, in his dictionary (1994) he marks fully raising words: e.g. ‹ǂìi› is entered as ⟨ǂài (> [ǂìi])⟩. Thus one can see which posited underlying ‹-ai› words undergo full A-Raising. For example, ‹ǀχāi› 'bowstring hemp plant', which is also a class 1 noun, with plural ‹ǀχāba-tê›, is entered just as ⟨ǀxāi⟩. Indeed, a recording of it is available, and it is pronounced with moderate raising.[11] An examination of all the data shows (26).

(26) Words of the form ‹({ǀ, ǂ} ⊗ *x*)ai› undergo full A-Raising if *x* is ‹ʞʰ ʞʔ ʞ ʞh ʞ̃›; they do not (and therefore undergo only moderate A-Raising) if *x* is ‹ʞχ ʞχ ʞqˠʼ ʞq ʞqʼ›.

That is, although a uvular segment in the accompaniment does not block moderate A-Raising, it does block full A-Raising. In *SPE*, uvulars are contrastively specified for [+back] and [−high], so there is a choice of which feature to use in the rule. I will accept Traill's view that A-Raising is indeed raising rather than fronting, and employ [high]. So, using the fact that my pure accompaniment phonemes and the two glottal phonemes are unspecified for [high], we can formulate the full A-Raising rule as (27).

(27) *Formal full A-Raising Rule*

$$\begin{bmatrix} \text{V} \\ +\text{low} \\ -\text{phar} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{high} \\ -\text{low} \\ -\text{back} \end{bmatrix} / \begin{pmatrix} \{\text{ǀ}, \text{ǂ}\} \\ \otimes \\ [\emptyset\text{high}]_1 \end{pmatrix} __ \begin{bmatrix} \text{V} \\ +\text{high} \\ -\text{back} \end{bmatrix}$$

Now consider what distinguishes ‹ǀ ǂ› from the other clicks. There have been several suggestions for features that do so. I tend to prefer Traill's notion that the difference is that they leave the tongue blade in a high front position, whereas the others pull the tongue lower and backer, which suggests either [back] or [high], or perhaps both. The rules work nicely if both are specified, as I laid out without explanation in (17).[12] Miller uses [pharyngeal] – see below.

[11] Available at *The UCLA phonetics lab archive* (http://archive.phonetics.ucla.edu/language/nmn/nmn_word-list_0000_01.wav).

[12] Note that since we have separated the clicks from the accompaniments, there is no interference between specifying the [high] feature for clicks and for the accompaniment; without the concurrent clustering, it is necessary to use a different feature, such as [low] or Miller's [pharyngeal]. This use of [high] does involve a certain relaxed approach to the intrinsic content of *SPE* features, as does the

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Given this, the following rule suggests itself as a combined description of A-Raising before ‹i›.

(28) *Formal A-Raising Rule*

$$\begin{bmatrix} \text{V} \\ +\text{low} \\ -\text{phar} \end{bmatrix} \rightarrow \begin{bmatrix} (\alpha \wedge \beta)\text{high} \\ -\text{low} \end{bmatrix} / \begin{pmatrix} [\alpha\,\text{high}] \\ \otimes \\ ([\emptyset\,\text{high}][\beta\,\text{high}]_0)_0^1 \end{pmatrix} __ \begin{bmatrix} \text{V} \\ +\text{high} \\ -\text{back} \end{bmatrix}$$

where $\alpha \wedge \beta$ is − if either α or β is −, and is + otherwise, and β is $\emptyset$ if unmatched.

This rule does not explicitly describe the concomitant fronting that results in [i] rather than [ɨ] in the full case – as a referee suggests, it is probably simplest to assume that a later rule fills in [−back]. It is also of course possible to incorporate fronting in (28), as in (27), at the price of some additional inelegance.

This rule neatly shows the concept that the raising and fronting effect of the following ‹i› is moderated either by the click or by the accompaniment. Moreover, since in (16) I used [±high] to distinguish dentals from alveolars, this rule also captures A-Raising with initial dental non-clicks: a non-click initial matches the context by taking the optional lower half to be empty; α then matches against the initial.

Several similarly complex sets of interactions between different coronal consonants and vowel backness are studied by Flemming (2003), with similar arguments about the different behaviour of the tongue body. The above description can also, as I noted, be cast in terms of fronting rather than raising, and would mostly fit in to Flemming's (2003) framework.

As examples of the formal application, consider (29).

(29) a. ‹ǂai› = /(ǂ ⊗ ʁ)ai/ = $\begin{pmatrix} [+\text{ling}, +\text{high}, -\text{back}, \ldots] \\ \otimes \\ [\emptyset\,\text{ling}, \emptyset\,\text{high}, -\text{voice}, \ldots] \end{pmatrix}$ /ai/

and so $\alpha = +$ and $\beta = \emptyset$ (because unmatched), so $\alpha \wedge \beta = +$, so /a/ changes to [+high, −low, ∅back], and then the later rule fills in [−back] from [+high], so ‹ǂai› → [ǂii].

b. ‹χai› = /(ǂ ⊗ ʁχ)ai/ = $\begin{pmatrix} [+\text{ling}, +\text{high}, -\text{back}, \ldots] \\ \otimes \\ [\emptyset\,\text{ling}, \emptyset\,\text{high}, \ldots][+\text{ling}, -\text{high}, \ldots] \end{pmatrix}$ /ai/

and so $\alpha = +$ and $\beta = -$, so $\alpha \wedge \beta = -$, so /a/ changes to [−high, −low, ∅back], i.e. ‹|χai› → [|χɜi].

---

alternative use of [back]. See Traill (1985: 107–108) for an extended discussion, although he was additionally handicapped by the need to include accompaniment features with the clicks.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Note that (28) does *not* agree with Traill's A-Raising Rule (5), because (28) predicts that there should be moderate raising following a 'back' click without a uvular accompaniment (whereas in (5) only the 'front' clicks trigger any raising). Traill (1994) in fact states that in such contexts ‹a› undergoes mild raising, to [æ]. However, I have studied his available recordings, and in the readings, all ‹-ai› words with back clicks appear to show the same degree of raising as other cases of moderate raising. There is not enough data to make any statistically meaningful claim, but both auditory impression and acoustic measurements suggest this. For example, in one recording ‹!hai› appears to show considerable assimilation, varying from [əi] to [ɛi] in the same speaker.[13] (On phonetic grounds, one might expect raising to be particularly marked in ‹ʁh›, since the long [h] allows plenty of time for the tongue to move away from the position forced by the click. However, there is not enough data available to me to check this.)

It is of course simple to force (29) to match (5), but this requires removing the symmetry between click and accompaniment features, and since the symmetric version appears to be more accurate, there is no call to do so.

5.1.3 *The Back Vowel Constraint*. Though the underlying ‹a› in most full A-Raising words is adequately supported by other evidence, part of its motivation is to explain exceptions to Traill's phonological Back Vowel Constraint in (4), which prohibits front vowels after any back consonant. As I noted, there is an alternative formulation (3) of the general Khoisan Back Vowel Constraint, which recognises the distinction between the front and back clicks, and it is perhaps unclear why one should recognise this difference in the A-Raising Rule but not in the Back Vowel Constraint.

A similar situation with regard to the Back Vowel Constraint occurs in Ju|'hoan, where front vowels also occur after the front clicks. Unlike Traill, Miller-Ockhuizen (2003) does not try to explain this away by a phonetic rule operating after the constraint, but rather states the Back Vowel Constraint in the form in (3), which distinguishes the front ‹| ǂ› clicks from the back ‹! ǁ› clicks. Her technique is to assign the feature [+pharyngeal] to ‹! ǁ›, and use that in the statement of the Back Vowel Constraint. This use of [pharyngeal] is motivated by the other 'guttural' constraints she analyses, but many of these do not appear to apply in !Xóõ. The phonetic grounding of this feature is supported by ultrasound; impressionistically, it seems to me to be a consequence of the apical articulation of ‹! ǁ› rather than a primary feature. Miller-Ockhuizen discusses in detail both her own and others' work on the acoustic and

[13] Available at *The UCLA phonetics lab archive* (http://archive.phonetics.ucla.edu/language/nmn/nmn_word-list_1983_01.wav). Unfortunately, one of the potentially most useful recordings for this issue is truncated, and the original cannot be traced.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

articulatory properties of the various clicks, and there are a number of ways in which the front clicks can be seen to differ from the back clicks.

In my approach, the choice made above to specify back clicks as [+back, −high] can be exploited to state the Back Vowel Constraint in a more refined form.

(30) *Concurrent Back Vowel Constraint*

A [−back] vowel must be licensed by an immediately preceding [−back] consonant.

This makes fully A-Raised words licit at the phonological level, and so removes the notion that they are exceptions. It therefore also allows the few remaining unexplained exceptions, such as ‹ǀīi› ‘if’, and a dozen or so words in ‹-e-› following a dental or palatal click.

It also permits a front vowel to follow a click with an uvular accompaniment, because both the click and the accompaniment immediately precede the vowel; in a non-concurrent formulation the uvular would block the licensing from the front click. According to Traill (1994), there are indeed a couple of such words: ‹ǀq’ìi-sà ǂgɛ̄ɛ̃›.

5.2 ‘Delayed aspiration’ and the voiceless nasal

The so-called delayed aspiration accompaniment ‹ʞh›, which is widespread in Khoisan, has caused some confusion historically, particularly in terms of its relationship to ‹ʞʰ› – and as I described in §2.3, it seems that !Xóõ has ‹ʞqʰ› in addition, though Traill is unclear about this.

Moreover, as I also noted, ‹ʞh› involves nasality, in the form of a (possibly ingressive) voiceless nasal at the beginning. (Beach 1938 notes some nasality in Khoekhoe, though he describes occasional voiced nasality.) Given that most Khoisan languages have the voiced nasal accompaniment ‹ᵑǀ›, one might wonder whether they are related. However, there are good arguments for the nasality of ‹ʞh› being a phonetic detail; for example, both !Xóõ (per DoBeS) and Juǀ’hoan have a voiced version ‹ᵑǀh›, and all voiced accompaniments are prevoiced and often have phonetic nasality, since nasality is the easiest way to maintain the voicing; similarly, in ‹ʞh› the nasality allows for the ‘soft start’ to aspiration – and Naumann (forthcoming) reports that some of his speakers describe ‹!h› as ‘[!] with a pause’. In any case, !Xóõ has a distinct voiceless nasal accompaniment ‹ᵐǀ›.

However, the !Xóõ voiceless nasal is somewhat of a puzzle. With the possible exception of ǂHoã (Linda Gerlach, personal communication), !Xóõ is the only extant language to possess this accompaniment, and it is unclear how it emerged.

Güldemann (2001: 13) notes that it appears only before pharyngealised or creaky vowels, and suggests that it perhaps split off from the voiced ‹ᵑǀ› in reaction to ‘the specific phonetic character of the marked stem vowels’. It is, however, hard to see how this could have happened, as ‹ᵑǀ› still

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

occurs in this environment, and there are even minimal pairs, such as ‹l̪óˤli› '*Antizoma angustifolia*' and ‹l̪óˤli› 'wipe or rub the eyes, pick the nose'.

In §4.2.4, I remarked that almost all, but not all, concurrent analysis accompaniment contrasts are supported by minimal pairs. It is therefore striking, and not to my knowledge previously observed, that the contrast ‹x̪̊› *vs.* ‹ʁh› has no support. Not only is there no minimal pair, investigation shows that they are indeed in complementary distribution. As Güldemann observed, ‹x̪̊› occurs only before creaky or pharyngealised vowels. It follows from the Pharyngeal Constraint (2c) that a pharyngealised vowel cannot occur after ‹ʁh›, but checking through Traill (1994) shows the stronger fact that ‹ʁh› occurs only before plain vowels.

Thus ‹x̪̊› and ‹ʁh› are in complementary distribution, and given the phonetic link between them in terms of voiceless nasality, it is tempting to conjecture that ‹x̪̊› is an allophone of ‹ʁh›. If we unify ‹ʁh› and ‹x̪̊› in Traill's analysis, and adopt the unitary concurrent analysis (i.e. sequentially unclustered), then there are 120 minimal pairs of accompaniment phonemes to find, and 115 of these exist, with the remaining five also found if we ignore tone.[14] This is exemplified in (31), and in fact all the missing minimal pairs in the un-unified system are contrasts with ‹x̪̊›.

(31) a. ‹ʁqʰ› *vs.* ‹x̪̊› no minimal pairs
b. ‹ʁqʰ› *vs.* ‹ʁh› minimal pairs
e.g. ‹ǀqʰáa› *vs.* ‹ǀháa› (and many others)

At first sight, phonological arguments cut both ways when considering (31). On the one hand, it is striking that ‹x̪̊› does not occur before breathy or strident vowels, whereas ‹x̪̊› is attested before both. Given the general Single Aspirate Constraint (2a), this lends support to the idea that ‹x̪̊› represents a phonological aspirate. On the other hand, declaring ‹x̪̊› to be an aspirate then violates the Pharyngeal Constraint (2c).

However, as I noted, the Pharyngeal Constraint is violated by several words of the form ‹hVˤ-›, such as ‹hóˤlo› 'stand on tiptoe', so that ‹h› itself does not appear to trigger the constraint, and given that I treat ‹ʁh› as a sequential cluster with ‹h›, there is no reason to think that ‹ʁh› triggers it. I therefore suggest that the constraint indeed does not apply to ‹ʁh›, and that its apparent application is due to the formation of ‹x̪̊› as an allophone in the pharyngeal context.

My conjecture as to the emergence of this suggested allophony is that maintaining the long [spread glottis] aspiration characteristic of ‹ʁh› is awkward when it is followed by the glottal constriction of creaky vowels, and also when followed by pharyngeal constriction, because then it will tend to lead to stridency, and so the voiceless nasality took over as

[14] 120, because in the 1994 version of Traill's phonology, the only source for which there is extensive data, there are 17 accompaniments, so removing one gives $(16 \times 15)/2 = 120$ possible contrasts.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

the main cue. In the context of plain ‹h-›, however, there was no such alternative cue.

I should note that the dialect recently studied by DoBeS slightly muddies the water on this issue. Naumann (forthcoming) reports a word in which ‹ʖ̰› occurs before a plain vowel; the same word is reported with a creaky vowel by Traill. Moreover, in the DoBeS data, the 'delayed aspiration' seems to have considerably stronger aspiration than in the eastern dialect, decreasing the phonetic similarity. The extent of dialectal differences *vs.* differences in analysis requires further investigation, but we might very tentatively conjecture that the distinction is allophonic in eastern !Xóõ, but in the process of phonologisation in western !Xóõ.

6 Concurrent phonemes: variations and extensions

In this section, I will first discuss some possible alternative choices in the formulation above; and then I will go on to suggest that the notion of concurrent segment and phoneme might be useful beyond the world of clicks. With clicks, the justification of click and accompaniment segments, and hence phonemes, was quite strong. In this section, the justification will become increasingly open to attack, and so I use this part to explore the boundary between concurrent segments and autosegments or features, following on from §4.2.3.

6.1 The nature of ‹ʖ›

In definition (12b) above, the accompaniment phonemes are defined to be specified only for their values of voice, ejectivity, aspiration, and so on, but not for any other values, such as place. The click phonemes are specified for anterior place, height, backness and linguality, but for nothing else. Moreover, it is assumed that neither clicks nor accompaniments occur by themselves in URs, but only in conjunction – in what sense, therefore, are they like other phonemes?

In the case of the pure clicks, I would assert that it is a contingent, rather than necessary, fact about language that clicks do not occur alone. A pure click is a click unconnected to any other airstream – for example, English *tsk! tsk!* [ǀ ǀ] consists of pure clicks. A (not very human) language could be constructed out of pure clicks; but any language that combines clicks with vowels, for example, must synchronise them, and having done so, can take advantage of modifications of the posterior closure.

For the accompaniments, the question is more subtle. I have assumed that ‹ʖ› does not occur on its own in URs, but, as I have remarked in several places, one can reformulate the theory so that it can. It is debatable whether such reformulations are more or less natural than that of §4.1.1. I shall consider three, the last of which provides an opportunity to discuss the curious nature of !Xóõ clusters.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

6.1.1 *Accompaniments as pulmonic stops.* One might simply say that ‹ʁ› is just ‹q› (or ‹k›). This is essentially the radical cluster analysis, made concurrent instead of sequential: the accompaniments are the existing series of uvular (or velar if preferred) stops. It has the same distributional problem as the radical cluster analysis: there is no ‹ɴ›. (There is also no ‹ɴ̥›, but that is not a problem if we follow §5.2.) There is a rather marginal ‹ŋ›, but not in onset position. Either of the solutions suggested for the radical cluster analysis could be applied.

Although this proposal avoids the unusual phonotactic constraint that accompaniments must appear with clicks, it introduces others: for example, why is it that there are no initial clusters ‹qx qh qʔ›? One has to argue that the point of the click clusters ‹ʁx ʁh ʁʔ› is that the posterior release is inaudible, and that an initial ‹q› with no release is rather pointless, but then the phonotactics must distinguish ‹q› *qua* accompaniment from ‹q› *qua* independent consonant.

A major drawback to this approach is that the bare accompaniment now has values for height, backness and linguality, and so all the rules have to be recast in a less elegant form. In particular, if ‹ʁ› is just ‹q›, there's nothing to distinguish the two ‹q›s in ‹qq›, and so the synchronisation rule, which previously could identify ‹ʁ›, must instead be written to dock the click onto the first uvular segment in the accompaniment. This happens to work, because there is no ‹qʁ›, but it is not elegant.

6.1.2 *Accompaniments as clicks.* An alternative suggestion is that solitary ‹ʁ› is ‹!›. On this view, the accompaniment carries with it a 'default' click, which I have somewhat arbitrarily chosen to be ‹!›, but this can be changed by concurrent composition with a different pure click. In the implementation of §4.1.1, this would be done by leaving the ‹ʁ› phonemes as they are, and adding a rule that fills in the ‹!› features for an isolated accompaniment.

In such a setting, of course, the chosen default pure click becomes redundant, and can be omitted from the inventory. This solution solves some problems, but there are, to my knowledge, no phonetic or phonological grounds for treating one click as more fundamental than another; and more importantly, it complicates the statement of rules such as A-Raising and the Back Vowel Constraint, as they apply to the default click too.

6.1.3 *The place of voice.* Another possible, and more substantial, variation has been proposed by Daniel Currie Hall (personal communication). I have chosen to attach all laryngeal features to the accompaniment. Hall notes that [voice] varies orthogonally to all other features, whereas [spread glottis, glottal closure, nasal] are mutually exclusive (assuming the cluster analysis of ‹ʔɣ̆›). Why not, then, place [voice] with the clicks rather than with the accompaniment? This would give the basic phonemes /! ǃ …/ and accompaniments /ʁ ʁʰ ʁ' ʁ̆/.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Such an organisation is also used by Güldemann (2001) in his feature-geometric approach. Hall suggests the advantages in (32).

(32) a. The plain and voiced clicks no longer require an accompaniment;
b. and consequently there is no longer a need for sequential clusters within concurrent clusters (e.g. ‹ǃg› is just /(ǃ⊗q)/), which also explains why:
c. only plain and voiced clicks occur in clusters with other stops.

This has obvious merits, similar to those that motivated Güldemann's (2001) similar decision. The counterarguments invoke the conceptual basis of the proposal in this article. *Ad* (a), plain and voiced clicks require just as much synchronisation of separate airstreams as other clicks, and at least in my own experience, voicing clicks is no easier than aspirating them. A click on its own would demonstrate a failure of synchronisation. *Ad* (b), if there is no sequential clustering, then one must resort to phonetic rules to explain why the clusters with stops have a prolonged closure after the click rather than before or around it. *Ad* (c), the non-occurrence of ejective, aspirated or nasal clicks in clusters is discussed in the following section.

There is also a more drastic approach to voice, which deserves mention. As is clear from Table I, the voicing distinction pervades the stop system; and as discussed in §2.2–§2.3, it appears as distinct prevoicing in most cases, other than the simple voiced stops. It is therefore tempting to follow the orthographies, and replace the voiced accompaniments ‹ǧ ǧg …› by sequential clusters with a voiced stop: ‹gǧ gǧq …›. To the best of my knowledge, there is nothing in !Xóõ phonology to argue against this, although it goes against almost all phonological tradition.

6.2 The nature of !Xóõ clusters

Although the click clusters seem complex, they are not unreasonably so. The second element of each cluster in rows 14–27 of Table I is either uvular or glottal, and so forms either a geminate closure or a simple release when following the posterior closure of the click; and each such second element exists independently.

Formally, in my proposal, the fact that accompaniments do not have the feature [+consonantal] means that Miller's (2011) objection (see §3.3.5) to obstruent–obstruent clusters does not obtain: in my /(!⊗ʁq)/, there is a parallel cluster of obstruents, but not a sequential cluster. This reflects the conceptual status of ‹ʁ› as a synchronisation point, which may carry manner features, rather than an obstruent in its own right.

As for the question, raised in (32c), of why there are no click clusters of the form ‹!^{h}q› = /(!⊗ʁ^{h}q)/, for example, the answer is that realising the aspiration on ‹!^{h}› would require either releasing the posterior closure and then re-forming it for ‹q›, so creating a sequential cluster of released obstruents, or transferring the aspiration to the ‹q›, resulting in something

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

indistinguishable from ‹!qʰ›. One may note also that the nasal accompaniment does occur in clusters: I analyse ‹ˀɴ̃ɢ̃› as /ʔɴ̃ɢ̃/, and it may be that ‹ɴ̃h› is phonologically /ɴ̃h/.

The question remains of the pulmonic clusters in rows 20–23. There is no escaping the phonetic fact that these are sequential obstruent–obstruent clusters, which clearly violate any alleged constraint against such. It is, however, possible to suggest that they are licensed by an analogy with the click clusters, as follows. The click ‹ǀqᵡ'› is /(ǀ⊗ʁqᵡ')/. Suppose that the suction is weakened, so that the /ǀ/ switches from [+lingual] to [−lingual]. The result is the illicit parallel cluster /(t⊗ʁqᵡ')/, which can be legitimised by fusing the /t/ with the /ʁ/, resulting in /tqᵡ'/. Thus one can see the ‹p t ʦ› clusters as weakened versions of the ‹ʘ ǀ ǁ› clusters, for example. However, to quote Traill (1985: 161), 'it is not the intention of these observations to imply that non-clicks developed from clicks'. Rather, there are many interesting parallelisms between clicks and non-clicks, which, I think, neither Traill, Güldemann nor my proposal has yet fully explained.

6.3 Concurrency in the !Xóõ vowel space

As I described in §2.4, the phonetic vowel space of !Xóõ has five basic vowels, together with (in Traill's view) arbitrary combinations of pharyngealisation, creakiness, breathiness and nasalisation: so instead of the two-dimensional IPA vowel chart, there is a six-dimensional chart. The phonological analysis cuts things down somewhat, but even so there are 26 vowel phonemes for DoBeS and 37 for Traill.

From the point of view of acquisition and stability of the sound system, all the same arguments apply as with clicks. Thirty-seven is a lot of vowels, and as with clicks, some of them are rare, or even unattested. There is, for example, no attested occurrence of ‹õ̰›, but it would be strange indeed if a nonce word including it were not recognised as such.

As with the clicks, there is also morphological evidence that creakiness and nasalisation at least behave independently of basic vowel quality. I sketched the principles of the !Xóõ concord system in §2.1.1. For most dependent forms, the vocalic part of the concord is ‹-ã -e -i -u›, according to the class of the governing noun – for example, the function word described in the lexicon as ‹kV› will appear as ‹kã ke ki ku›, according to concord. The demonstrative 'this' is ‹tṼ̰V›, with the allomorphs ‹t̰ã t̰ḛ t̰ḭ t̰ṵ›̃ – thus the creakiness on the vowel, and indeed the length of the vowel, are part of the lexical specification, while the basic vowel quality and nasalisation vary with concord. So the qualities qualify as morphophonemes at least.

I have also noted that strident epiglottal vowels appear to be phonologically breathy pharyngealised, and that there are Single Aspirate and Glottal constraints (2a, b).

Given the free interplay of voice qualities and nasalisation, it is obviously tempting to treat them as phonemes rather than morphophonemes.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

One could do this by claiming that the first mora of a word may have coda consonants ‹ʕ ɦ ʔ›, and the second ‹ŋ›, as is written in the DoBeS orthography (as ⟨q h ’⟩ and ⟨n⟩), and that these consonants then spread their quality to the vowels. However, while both creakiness and pharyngealisation are often realised with a peak that sounds like a light stop (Traill 1985), this peak does not appear to occur between moras, but in the first: e.g. ‹aˤi› sounds more like [aʕai] than [aʕi].

Thus, if I wish to admit these qualities as phonemes, the obvious way to do so is to make them concurrent with the vowels, e.g. /(a⊗ʕ)/. In the formal setting, this requires relaxing the Strong Concurrent Airstream Constraint (11) to allow concurrent actions within the pulmonic airstream, and extending the synchronisation rules accordingly, but raises no other issues.

Following my discussion in §4.2.3, I also have to justify their existence as phonemes in the inventory. This requires a rather greater relaxation of the notion of segment than for click accompaniments, and leads into controversial issues.

(i) *Acoustically*, each of the four basic qualities have measurable correlates.[15]

(ii) *Articulatorily*, nasalisation and pharyngealisation are independent gestures. Breathiness and creakiness are not, as they require opposite laryngeal gestures; but the resolution of the conflict by sequencing permits them to be conceived of as such. Other languages, such as Chong (Theraphan 1991), have also been reported to have breathy–creaky vowels implemented by sequencing.

(iii) *Perceptually*, the four basic qualities are independently perceptible without training – even in English they are recognised paralinguistically, either as emotional indicators (breathiness and creakiness) or as stereotypes of other languages: the ‘nasal twang’ (Sweet 1877: 8, Mayo & Mayo 2011) of some accents of English, or the ‘guttural’ sound of Arabic, arising from the pharyngeal and uvular consonants. It is not always easy to distinguish breathiness and nasalisation, as these qualities share a number of acoustic cues (Arai 2006), but other languages (such as Mazatec languages and Hindi) use both breathiness and nasality.

(iv) In *production*, I predict that, for example, if one teaches a !Xóõ speaker [y], they will immediately be able to produce [y̰] and [ỹ̰].

6.4 Nasality in other languages

The suggestion of nasality as a phoneme immediately brings to mind other languages. Nasality occurs in many different language families, and its behaviour varies widely, from ‘featural’, through what I am arguing is ‘concurrent segmental’, to something that seems to be suprasegmental,

[15] This is not an entirely honest statement: nasality has a rather wide and complex set of acoustic cues (Raphael 2005).

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

even up to word level, and can be naturally characterised in autosegmental theory, as in the following examples.

In phonetic and purely phonological descriptions of French, the nasal vowels are standardly seen as having phonemic status. The qualities of some of the vowels have drifted far from the oral counterparts – e.g. the historical and orthographic *in* is not [ĩ] but [æ̃] – and although the connection between nasal and oral is live, in alternations such as masculine *gamin* /-æ̃/ *vs.* feminine *gamine* /-in/, this is usually seen as morphophonological, on a par with the English /ai/ *vs.* /ɪ/ in *divine*/*divinity*.

In Portuguese, the nasal vowels have essentially the same quality as their oral counterparts, and although the morphophonology is similar to French, some analyses of Portuguese phonology propose retaining the historical following nasal, e.g. as an archiphoneme /N/ (Barbosa & Albano 2004), and regarding the nasalisation as phonetic. One could argue that the situation is in fact neither of those: rather, nasalisation is a concurrent phoneme with the vowel.

For !Xóõ, I have argued that nasalisation appears to behave exactly like any other phoneme, save for that fact that it sits on top of a vowel rather than occurring after it, and so is a good example of a concurrent phoneme.

In many South American languages, nasality appears as a suprasegmental property, so that, for example, [m] may appear as an allophone of /b/ that occurs in nasal morphemes or syllables. There may be spreading rules which may propagate the nasality further in the word, subject to various blocking conditions (see e.g. Peng 2000 for illustrations). This extensive nasal harmony is naturally treated via autosegmental processes; for example, Botma (2004) treats such languages (and others) within the framework of Dependency Phonology. Of course, formally one could claim that Tuyuca (Barnes 1996) [mãr̃ĩ] and [tɨ̃ŋõ] are underlyingly /(˜⊗barí)/ and /(˜⊗tɨgó)/, but as Barnes' title suggests, there appear to be morphemes marked nasal, marked oral and unmarked. Asserting nasality as a quasi-segment is one thing, but asserting orality is quite another, and so I would not claim that concurrent phonemes are an appropriate way to analyse nasality in Tuyuca.

6.5 Concurrent phonemes in language change

Returning to the case of French, I would further suggest that the history of French may be understood more easily by the use of concurrent phonemes. A standard philological description of the development of the French nasal vowel in *quand* from Latin *quandō* would be, compressing irrelevant changes, as in (33).

(33) (/kwando:/ [kwando:] →)
/kant/ [kant] → /kant/ [kãnt] → /kãt/ [kãt]
(→ /kã/ [kɑ̃])

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

An equally standard criticism of such accounts is that there is an explanatory lacuna at the phonologisation stage: the trigger for the change disappears, and so the nasal vowel is phonologised – but if the trigger disappears, why doesn't the nasalisation? The most obvious answer is to invoke generational change: the parents have /kant/, realised as [kãnt], with spreading, while the children have reanalysed [kãnt] as /kãt/, with an excrescent [n], and so two grammars coexist with the same output. The phonologisation is Ohala's (1981) notion of hypocorrection, but in his account, it is not clear why the children should 'fail to hear' the [n], unless they do hear it and apply his hypercorrection to interpret it as zero. This simultaneous hypocorrection/hypercorrection seems a little contorted.

My preferred answer to this old puzzle is the one that says that 'phonologisation' can happen without contrast; or, more generally, that there is a continuum between allophony and phonemic contrast,[16] and an allophonic distinction can become gradually internalised in the mental representation, as suggested by, for example, Hooper (1981). (See also Peperkamp *et al*. 2003 for an experimental study of the allophone/phoneme distinction during acquisition, and Hall 2009 for a model of such systems.) In categorical terms, this amounts to promoting the phonetic intermediate to a non-contrastive but phonological intermediate, as in (34).

(34) /kant/ [kant] → /kant/ [kãnt] *phonetic spreading*
→ /kãnt/ [kãnt] *hypocorrection*
→ /kãt/ [kãt] *hypercorrection*

Such an account results in the simultaneous emergence of many unsupported 'phonemes', one for each oral vowel that undergoes nasalisation, existing without contrastive support for possibly several generations. If we cast the history in terms of concurrency, then the intermediate stage involves only one new phoneme to account for all the vowels that undergo nasalisation – and moreover, the use of concurrency avoids interference in existing phonotactics, as the sequential adjacency relation is unchanged. Only when nasalisation is completely fused (as perhaps in French but perhaps not in Portuguese) do we really have five new vowel phonemes. Thus we might have (35).

(35) /kant/ [kant] → /kant/ [kãnt] *phonetic spreading*
→ /k(˜⊗a)nt/ [kãnt] *hypocorrection*
→ /k(˜⊗a)t/ [kãt] *hypercorrection*
→ /kãt/ [kãt] *concurrent fusion*

[16] For example, in my fairly conservative RP speech, coda /l/ is dark but fully lateral, and until I read Tolkien (1966: 392) at the age of 12 or so, I had never considered coda and onset /l/ to be different. My 10-year-old son, however, has a fully vocalised coda /l/ [ɣ], and considers this to be clearly a 'different sound' from onset /l/ [l], although he has no evidence for a contrast between them, and otherwise shows no particular ability in phonetic discrimination.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

A similar story might be told about palatalisation changes. In Gaelic, for example, palatalisation emerged from adjacent front vowels in the usual way, but a standard synchronic phonemic analysis simply posits separate palatalised and plain (or velarised) versions of most consonants. However, speakers are (at least in the presence of elementary education) well aware of the distinction, and every Gaelic speaker knows that there is broad (*leathan*) /t/ and slender (*caol*) /t/. So one might even say that Gaelic has not yet fused the palatalisation, and /tʲ/ ([tʲ ~ ʧ]) is still /(t⊗ʲ)/ – whereas in English, there is no synchronic relationship at all between /k/ and /ʧ/, although the latter is historically a palatalisation of the former.

6.6 Tone

No discussion can be complete without mentioning tone, the concurrent quality *par excellence*. It has always been considered, in both the Western and Chinese linguistic traditions, that Chinese tone is a property of syllables, parallel to the segmental content. Other tone languages also do this, and indeed tone, despite being contrastive, is often not considered worth writing in everyday use, even when the official orthography supports it (e.g. Zulu and Xhosa – and Khoisan languages).

In the case of typical African language families, the tonology is rich and involves sometimes very long-range processes. Such complexity was one of the main motivations for Goldsmith's (1976) elaboration of autosegmental phonology, and for the same reason, it is too rich to be sensibly encompassed within my notion of concurrent phoneme.

With Chinese and similar languages, on the other hand, it seems plain that tone meets every test I have suggested for segmenthood rather than featurehood, and so I would certainly claim that a toneme is a concurrent phoneme. However, unlike the situation with clicks, such a statement is purely a rephrasing of what everybody already agrees, and gives no new insights.

7 Conclusion

In this article, I have proposed a modification of the traditional understanding of the terms 'segment' and 'phoneme' to include the notion of parallel as well as sequential clustering. In the case of Khoisan languages, such a modification dramatically reduces the inventory sizes, and thereby makes the languages appear much less exotic – and also much easier to acquire and maintain, if one accepts that maintaining a large number of phonemic contrasts is harder than using contrasts between clusters of phonemes. It also allows a better account of some phonological processes found in the languages. I may note that such a radical reduction in inventory sizes naturally challenges the methodology of some recent proposals (Atkinson 2011) about language dispersion.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

In addition, the use of concurrent analyses of clicks has exposed hitherto unobserved facts about phonological distributions in !Xóõ, thereby suggesting an allophonic relationship between two accompaniment phonemes, one of which is a long-standing puzzle because of its rarity.

I have also demonstrated a range of other uses for the concept of concurrent phoneme, where an audible character appears to behave more like a segment than a feature, and proposed that this gives a better motivated account of various diachronic processes.

REFERENCES

Arai, Takayuki (2006). Cue parsing between nasality and breathiness in speech perception. *Acoustical Science and Technology* **27**. 298–301.

Atkinson, Quentin D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* **332**. 346–349.

Bale, Alan, Maxime Papillon & Charles Reiss (2013). Sets of features *vs*. sets of sets of features: a notational innovation for phonological rules. Poster presented at the 21st Manchester Phonology Meeting.

Barbosa, Plínio A. & Eleonora C. Albano (2004). Brazilian Portuguese. *Journal of the International Phonetic Association* **34**. 227–232.

Barnes, Janet (1996). Autosegments with three-way lexical contrasts in Tuyuca. *IJAL* **62**. 31–58.

Beach, D. M. (1938). *The phonetics of the Hottentot language*. Cambridge: Heffer.

Best, Catherine T., Anthony Traill, Allyson Carter, K. David Harrison & Alice Faber (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. In M. J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 853–856.

Besten, Michael Paul (2006). *Transformation and reconstitution of Khoe-San identities*. PhD dissertation, University of Leiden.

Boersma, Paul & David Weenink (2013). *Praat : doing phonetics by computer* (version 5.3.18). http://www.praat.org.

Botma, Bert (2004). *Phonological aspects of nasality : an element-based dependency approach*. PhD dissertation, University of Leiden.

Campbell, Lyle (1974). Phonological features: problems and proposals. *Lg* **50**. 52–65.

Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.

Dixon, Robert M. W. (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press.

Ermisch, Sonja (ed.) (2008). *Khoisan languages and linguistics : Proceedings of the 2nd International Symposium, January 8–12, 2006, Riezlern/Kleinwalsertal*. Cologne: Rüdiger Köppe.

Flemming, Edward (2003). The relationship between coronal place and vowel backness. *Phonology* **20**. 335–373.

Goldsmith, John A. (1976). *Autosegmental phonology*. PhD dissertation, MIT.

Greenberg, Joseph H. (1950). Studies in African linguistic classification. VI: The click languages. *Southwestern Journal of Anthropology* **6**. 223–237.

Güldemann, Tom (2001). *Phonological regularities of consonant systems across Khoisan*. Leipzig: Institute of African Studies, University of Leipzig.

Güldemann, Tom & Edward D. Elderkin (2010). On external genealogical relationships of the Khoe family. In Matthias Brenzinger & Christa König (eds.) *Khoisan languages and linguistics : Proceedings of the 1st International Symposium, January 4–8, 2003, Riezlern/Kleinwalsertal*. Cologne: Rüdiger Köppe. 15–52.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Haacke, Wilfrid H. G. & Eliphas Eiseb (2002). *A Khoekhoegowab dictionary*. Windhoek: Gamsberg Macmillan.
Hall, Kathleen Currie (2009). *A probabilistic model of phonological relationships from contrast to allophony*. PhD dissertation, Ohio State University.
Heine, Bernd & Henry Honken (2010). The Kx'a family: a new Khoisan genealogy. *Journal of Asian and African Studies* **79**. 5–36.
Hooper, Joan Bybee (1981). The empirical determination of phonological representations. In Terry Myers, John Laver & John Anderson (eds.) *The cognitive representation of speech*. Amsterdam: North-Holland. 347–357.
IPA (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
Jakobson, Roman (1968). Extra-pulmonic consonants (ejectives, implosives, clicks). *Quarterly Progress Report of the Research Laboratory of Electronics* **90**. 221–227.
Kießling, Roland (2008). Noun classification in !Xoon. In Ermisch (2008). 225–248.
Ladd, D. Robert (2014). *Simultaneous structure in phonology*. Oxford: Oxford University Press.
Ladefoged, Peter & Ian Maddieson (1996). *The sounds of the world's languages*. Oxford & Malden, Mass.: Blackwell.
Ladefoged, Peter & Anthony Traill (1984). Linguistic phonetic description of clicks. *Lg* **60**. 1–20.
Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.) (2013). *Ethnologue: languages of the world*. 17th edn. Dallas: SIL International. Available at http://www.ethnologue.com.
Mayo, Carolyn M. & Robert Mayo (2011). Normative nasalance values across languages. *ECHO* **6:1**. 22–32.
Miller, Amanda L. (2011). The representation of clicks. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*. Malden, Mass.: Wiley-Blackwell. 416–439.
Miller, Amanda L., Johanna Brugmann, Bonny Sands, Levi Namaseb, Mats Exter & Chris Collins (2009). Differences in airstream and posterior place of articulation among N|uu clicks. *Journal of the International Phonetic Association* **39**. 129–161.
Miller, Amanda L., Levi Namaseb & Khalil Iskarous (2007). Tongue body constriction differences in click types. In Jennifer Cole & Jose Ignacio Hualde (eds.) *Laboratory phonology 9*. Berlin & New York: Mouton de Gruyter. 643–656.
Miller-Ockhuizen, Amanda L. (2003). *The phonetics and phonology of gutturals: a case study from Ju|'hoansi*. New York & London: Routledge.
Nakagawa, Hiroshi (2006). *Aspects of the phonetic and phonological structure of the G|ui language*. PhD thesis, University of the Witwatersrand, Johannesburg.
Naumann, Christfried (2008). High and low tone in Taa ǂaa (!Xóõ). In Ermisch (2008). 279–302.
Naumann, Christfried (forthcoming). The phoneme inventory of Taa (West !Xoon dialect). In Rainer Vossen & Wilfred H. G. Haacke (eds.) *Essays in memory of Anthony Traill*. Cologne: Rüdiger Köppe.
Ohala, John J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick & M. F. Miller (eds.) *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society. 178–203.
Oxford English Dictionary. Online edition. Oxford: Oxford University Press. http://www.oed.com.
Peng, Long (2000). Nasal harmony in three South American languages. *IJAL* **66**. 76–97.
Peperkamp, Sharon, Michèle Pettinato & Emmanuel Dupoux (2003). Allophonic variation and the acquisition of phoneme categories. In Barbara Beachley, Amanda Brown & Francis Conlin (eds.) *Proceedings of the 27th Annual Boston University Conference on Language Development*. Somerville: Cascadilla. 650–661.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press

Prince, Alan & Paul Smolensky (1993). *Optimality Theory : constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder.
Raphael, Lawrence J. (2005). Acoustic cues to the perception of segmental phonemes. In David B. Pisoni & Robert E. Remez (eds.) *The handbook of speech perception*. Malden, Mass. & Oxford: Blackwell. 182–206.
Schultze Jena, Leonhard (1928). *Zur Kenntnis des Körpers der Hottentotten und Buschmänner*. Jena: Fischer.
Snyman, J. W. (1970). *An introduction to the !Xũ (!Kung) language*. Cape Town: Balkema.
Sweet, Henry (1877). *A handbook of phonetics*. Oxford: Clarendon.
Theraphan, Thongkum L. (1991). An instrumental study of Chong registers. In Jeremy H. C. S. Davidson (ed.) *Austroasiatic languages : essays in honour of H. L. Shorto*. London: School of Oriental and African Studies, University of London. 141–160.
Tolkien, John Ronald Reuel (1966). *The return of the King*. 2nd edn. London: George Allen & Unwin.
Traill, Anthony (1985). *Phonetic and phonological studies of !Xóõ Bushman*. Hamburg: Buske.
Traill, Anthony (1991). Pulmonic control, nasal venting, and aspiration in Khoisan languages. *Journal of the International Phonetic Association* **21**. 13–18.
Traill, Anthony (1993). The feature geometry of clicks. In P. M. S. van Staden (ed.) *Linguistica : Festschrift E. B. van Wyk : 'n huldeblyk*. Pretoria: van Schaik. 134–140.
Traill, Anthony (1994). *A !Xóõ dictionary*. Cologne: Rüdiger Köppe.
Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
Wiese, Richard (1996). *The phonology of German*. Oxford: Clarendon.

https://doi.org/10.1017/S0952675714000025 Published online by Cambridge University Press