

Stabilising determinants in the transmission of phonotactic systems: Diachrony and acquisition of coda clusters in Dutch and Afrikaans

Andreas Baumann 

Department of English and American Studies, University of Vienna, Austria
E-mail: andreas.baumann@univie.ac.at

Daan Wissing 

Centre for Text Technology, North-West University, Potchefstroom Campus, South Africa
E-mail: daan.wissing@nwu.ac.za

Abstract

The phonotactic system of Afrikaans underwent multiple changes in its diachronic development. While some consonant clusters got lost, others still surface in contemporary Afrikaans. In this paper, we investigate to what extent articulatory difference between the segments of a cluster contribute to its successful transmission. We proceed in two steps. First, we analyse the respective effects of differences in manner of articulation, place of articulation and voicing on the age at which a cluster is acquired by analysing Dutch acquisition data. Second, we investigate the role that these articulatory differences play in the diachronic frequency development from Dutch to Afrikaans. We demonstrate that large differences in manner of articulation between segments contribute to a cluster's success in acquisition and diachrony. In contrast, large differences in place of articulation have impeding effects, while voicing difference shows a more complicated behaviour.

Keywords: Dutch/Afrikaans phonotactics, articulatory difference, first-language acquisition, diachronic change

1. Introduction

In its history, the sound system of Afrikaans underwent a number of changes. A considerable amount of these changes is related with the way in which sounds are combined to form strings in speech, i.e. with the language's phonotactics. Of course, not all combinations of sounds – or more precisely, phonemes – are permitted in speech. For instance, no Afrikaans word ends in /ɲn/, a sequence that is for instance perfectly fine in (Austrian) German, as in /d̥ɛ:tʃn/ (*Tetschn*, 'slap in the face'), nor is there an Afrikaans word ending in /rm/. The latter was not always the case. For instance, words like *arm* ('arm'), *skerm* ('screen') or *wurm* ('worm') were, and still are, articulated with a final /rm/ sequence in Dutch, while the two phonemes are separated in contemporary Afrikaans by a neutral vowel so that they end in /rəm/. Similarly, some consonant

sequences have been reduced in the history of Afrikaans so that words that originally ended in /xt/ like (historical and contemporary) Dutch *nacht* ('night') or *specht* ('woodpecker') are pronounced with a single word-final consonant in contemporary Afrikaans (e.g. /nax/, *nag*, and /spex/, *spēg*, although *houtkapper* is admittedly more widely used in Afrikaans; see Figure 1).

However, although obviously some phonotactic restructuring took place in the diachrony of the Afrikaans phonotactic system, not all sequences of consonants (or "consonant clusters") have been affected by deletion processes. Still a large number of consonant pairs surface in Afrikaans speech, /rx/ (in *berg*, 'mountain'), /nt/ (in *kind*, 'child'), /rt/ (in *boord*, 'edge') and /ls/ (in *dikwels*, 'often') being just a small selection of examples offered by the consonant-cluster inventory of Afrikaans. Some consonant sequences, although still present in Afrikaans, became less frequent while others are used more frequently.



Figure 1. Irrespective of whether or not these animals are fine with it, woodpeckers and flatworms underwent cluster-deletion processes in the history of Afrikaans: *specht* > *spēg* ('woodpecker') by consonant elimination from /xt/ to /x/ and *worm* > *wurm* ('worm') by schwa epenthesis from /rm/ to /rəm/. But why?¹

The question of why some clusters have been deleted in Afrikaans diachrony, while others are still abundantly used – and thus diachronically more stable – figures centrally in this paper. More precisely, we consider which articulatory factors might determine the diachronic stability of word-final consonant clusters in Afrikaans. Consonants are conventionally described by three primary articulatory features:

1. Place of articulation (PoA): where does the obstruction occur in the vocal tract?
2. Manner of articulation (MoA): how tight is the obstruction?
3. Voicing: do the vocal folds vibrate?

These features, naturally, play a crucial role when consonants are paired to form strings, i.e. phonotactic items. There are reasonable arguments for the assumption that clusters whose segments differ with respect to some of these articulatory features are more easily processed and, as a consequence over multiple production-and-perception cycles, diachronically more stable. Likewise, there are reasons for assuming on the contrary that phonotactic items are

¹ Photographs taken from Wikimedia Commons (2005) and Wikimedia Commons (2012), and modified to include phonological transcriptions.

particularly successful if the segments they are composed of mesh with each other with respect to manner, place or voicing (cf. discussion in 1.2).

The aim of this paper is to shed light on the differential effects that manner-of-articulation difference, place-of-articulation difference and difference in voicing exert on the diachronic stability of phonotactic items in Afrikaans. Thus, we seek to explore which of these factors represent relevant determinants in the emergence of the Afrikaans inventory of sequences of two consonants. We address this aim in two separate studies. First, the differential impact of articulatory distances in the acquisition of Dutch is assessed. In a second study, we investigate the effects of articulatory distances on the diachronic success of consonant clusters in the transition from Dutch to Afrikaans. The underlying argument goes like this: if a certain articulatory difference – say, difference in voicing – between the segments of a cluster has beneficial effects in phonotactic acquisition, then similar effects should apply diachronically so that differentially voiced clusters are on average diachronically more stable than those clusters in which the segments share the same voicing pattern.² Altogether, we show that clusters benefit from large intersegmental manner-of-articulation differences and are restrained by large place-of-articulation differences, while voicing places a more complicated role in the acquisition and change of word-final phonotactics in Afrikaans.

The paper is structured as follows: First, the history of Afrikaans phonotactics is recapitulated (1.1), followed by a more detailed discussion of articulatory differences in phonotactic production and perception (1.2). Subsequently, the link between language acquisition and change is discussed, thereby particularly focusing on the phonological domain. In the end of the first section, the research hypotheses to be investigated in this paper are brought forth (1.4). The introductory section is followed by a detailed description of the two studies which assess the above-sketches research questions. Section 2 describes the acquisition study of Dutch phonotactics, while section 3 presents the (historically) comparative study meant to evaluate the diachronic stability of phonotactic items. Finally, the respective results are compared to each other and discussed (4).

1.1 Processes in phonotactic change: From Dutch to Afrikaans

We first discuss the processes that shaped Afrikaans phonotactics in its history originating in Dutch. We focus on word-final phonotactics, which implies that we restrict ourselves to dynamics in the coda. As in all Germanic languages, Dutch features complex consonant phonotactics in the coda, mostly clusters consisting of two consonants such as /sp/ in *wesp* ‘wasp’ or /xt/ in *nacht* ‘night’, as will be discussed below. In its evolution from Dutch, starting in the 17th century, Afrikaans coda phonotactics underwent a number of processes that lead to the loss but also to the emergence of cluster types.

Leaving changes of the quality of a consonant aside, we can distinguish between two major types of processes that are relevant to coda phonotactics: (a) deletion and (b) insertion

² This is expected because the same cognitive mechanisms apply to language learners and proficient speakers, and because linguistic knowledge of the latter clearly depends on the process of language acquisition, if only by the differently strong entrenchment of early- and late-acquired linguistic items. We argue that the question of whether children or adults contribute more to linguistic change is secondary in this respect, but see 1.3 for some discussion.

processes. In word-final V(C)CC structures, both of them can apply either (i) in the end of the cluster or (ii) cluster-medially. In the diachronic development of Afrikaans, all of the four logically possible changes that have occurred are shown in Table 1:

Table 1. Coda processes in the diachrony of Afrikaans (Conradie 2017; Donaldson 1993; Roberge 2002)

	Process	Example
(ai)	final t-deletion	/xt/ > /x/ in <i>nacht</i> > <i>nag</i> ('night')
(aii)	medial t-deletion	/xts/ > /xs/ in <i>slechts</i> > <i>slegs</i> ('only')
(bi)	final t-insertion	/n/ > /nt/ in <i>oven</i> > <i>oond</i> ('oven')
(bii)	medial p-insertion; medial schwa- insertion	/mt/ > /mpt/ in <i>hemt</i> > <i>hempt</i> > <i>hemp</i> ('shirt'); /rm/ > /rəm/ in <i>worm</i> > <i>wurm</i> ('worm')

Let us go through (ai) to (bii) above in more detail. Final t-deletion (ai) – also t-apokope – is a process that, according to van Veen (1964), is known since the Middle Ages. It first occurred in and around Utrecht in the Netherlands and later also in the vernacular of Dutch cities in the west of the Netherlands. It is generally assumed to have been transferred to the Cape, where it developed further, while since then being reversed in Standard Dutch (Schouten 1982). In Afrikaans, the process is context sensitive and only applies to word-final VCC structures in which the first consonant is an obstruent (Conradie 2017). This is illustrated in Table 2, which lists all obstruent-obstruent clusters in Dutch and Afrikaans. The process does not apply when the first consonant is a sonorant, e.g. /lt, rt, nt/ in Dutch and Afrikaans *asfalt* ('asphalt'), *hart* ('heart') or *land* ('land' pronounced with final /nt/ due to final devoicing), respectively, although Ponelis (1989) argues that general stop-deletion after sonorants was productive in early Afrikaans due to Malay influence, e.g. /mp/ > /m/ in *lamp* 'lamp', /lt/ > /l/ in *wild* 'wild' or /rt/ > /t/ in *boord* 'board'. According to Kotzé (1984), however, this contact-induced process is restricted to clusters in which the first segment is either a nasal or /l/ (e.g. in /mɔnt/, *mond* 'mouth'; /xəstamp/, *gestamp* 'bumped'; /xelt/, *geld* 'money'). The latter process still occurs in the Cape region. Beyond that, it is seen sporadically in unstressed syllables, e.g. /rt/ > /t/ in *mosterd* 'mustard' (Ponelis 1989: 4). Note that the dental stop still surfaces in plural forms like *nagte* (/naxtə/ 'nights') when the /t/ occurs in the onset of the final syllable (Watermeyer 1996).

Table 2. Word-final obstruent-obstruent clusters in Dutch and Afrikaans. Dutch examples are taken from Linke (2017). Sporadically unreduced types are indicated with a dagger (†).

	Dutch		Afrikaans	
	example	translation	example	translation
/pt/	<i>concept</i>	'concept'	—	—
/ps/	<i>rups</i>	'caterpillar'	<i>raps</i>	'slightly'
/ts/	<i>muts</i>	'hat'	<i>flits</i>	'torch'
/tʃ/	<i>kitsch</i>	'kitsch, junk'	<i>kitsch</i>	'kitsch, junk'
/kt/	<i>pact</i>	'pact'	<i>pakt</i> †	'pact'
/ks/	<i>heks</i>	'witch'	<i>heks</i>	'witch'
/ft/	<i>kaft</i>	'cover'	—	—
/fs/	<i>vergeefs</i>	'vainly'	<i>vergeefs</i>	'vainly'
/sp/	<i>wesp</i>	'wasp'	<i>wesp</i>	'wasp'

/st/	<i>beest</i>	‘beast’	—	—
/sk/	<i>kiosk</i>	‘stall’	<i>grotesk</i>	‘grotesque’
/xt/	<i>macht</i>	‘power’	<i>agt†</i>	‘eight’
/xs/	—	—	<i>slegs</i>	‘only’

In contrast to final position, t-deletion within a coda cluster (a_{ii}), necessarily VCCC, to our knowledge only occurs in /xts/ > /xs/ (*slechts* > *slegs*, ‘only’) and it is debatable if this process merely represents final t-deletion together with -s suffixation (the latter likely being an adverbial derivational suffix). This boils down to the question of whether or not *sleg+s* (derived from *sleg*, ‘bad’, with final /x/; originally in Dutch *slecht* ‘bad’ but formerly also ‘simple’, with final /xt/) is still morphologically transparent.

Final t-insertion (b_i), or t-paragoge, applies sporadically in forms like *oond* from *oven* (‘oven’), *reent* from *regen* (‘rain’) or *behoort* from *behooren* (‘belong’) (Conradie 2017). It seems to apply only after sonorants, which is consistent with the context-sensitive t-deletion process discussed before.

Finally, there are some interesting processes of epenthesis (b_{ii}) involving consonants and schwa. On the one hand, there is p-insertion in final /mt/ clusters to yield /mpt/, e.g. in *hemt* > *hempt* > *hemp* (‘shirt’), a process which is likely articulatorily motivated, the opening of the lips in the transition from bilabial /m/ to /t/ functioning as a release of the bilabial stop /p/ (Conradie 2017). Similar processes can be observed in (Austrian) German in *Hemd*, /hempt/, ‘shirt’ or *Samt*, /sampt/, ‘velvet’. More commonly, however, is schwa-epenthesis in clusters in which the first segment is a liquid /r, l/ and the second segment a nasal /n, m/, e.g. in *worm* /vɔrm/ > *wurm* /vʌrəm/ (‘worm’) or *psalm* /psalm/ > / pəsələm/ (‘psalm’, here also in the onset) (Donaldson 1993). Donaldson (1993) points out that schwa-epenthesis does not apply to liquid-stop clusters, a process which occurs sporadically in Dutch (e.g. /rk/ > /rək/ in *kerk*, ‘church’). We suspect that the latter represents a more recent development in Dutch (Kuijpers, van Donselaar and Cutler 1996), so a process of checked schwa deletion /rək/ > /rk/ should not be assumed for Afrikaans.

From what has been covered so far, it becomes clear that the phonotactics of Afrikaans underwent numerous changes. That is, some phonotactic items, or strings of consonants in our case, got lost while others have emerged during the past couple of centuries. Table 3 gives an overview of those consonant-cluster types that were, historically speaking, unsuccessful in that they got lost in the history of Afrikaans. Note that this table does not consider token frequency, in other words, a cluster type is considered present even if it only features a single token. In the remainder of this paper we will adopt a more fine-grained assessment of linguistic stability which also takes token frequencies into account.

Table 3. Word-final CC cluster types that got lost in the history of Afrikaans.

Cluster type	Process
/pt/	final t-deletion
/ft/	
/st/	
/mt/	medial p-insertion
/lm/	medial schwa-insertion
/rm/	

In the subsequent section, we discuss the articulatory and perceptual factors that influence the historical stability of a phonotactic item, in order to better understand the motivation behind the phonotactic processes (ai) to (bii) discussed above.

1.2 Articulatory and perceptual determinants in phonotactic stability

A phonotactic item, e.g. a consonant cluster, that remains within a language's phonotactic inventory through multiple generations of speakers is considered as diachronically stable. Diachronic stability has at least two different aspects. First, phonotactic items can be stable because they are frequently uttered and used. Trivially, a certain sound sequence that is not produced any more although speakers might have mental representations of that sequence stored in their memories, will cease to exist and drop out of a language's phonotactic inventory. Sound sequences that are, in contrast, produced frequently have a higher chance of being perceived and memorised by other speakers, which leads to enforced cognitive entrenchment of that sequence. This, in turn, entails higher utterance frequencies, as many cognitive linguists argue (Bybee 2007, Bybee 2010, Croft 2000, Langacker 2008).

The second aspect relevant to phonotactic stability is that of accuracy in production and perception. If it is extremely difficult to produce a certain sequence of sounds – for instance because the articulation of that sequence involves inconvenient and laborious movements of the tongue – it is likely that speakers will facilitate that sequence by, for example, switching places of articulation. Notably, this might be the case no matter how frequently the sequence should surface in speech, e.g. caused by very productive morphological operations (e.g. plural *-s* in Afrikaans as in *kinders* 'children'). Indeed, it has been argued that utterance frequency even decreases accuracy in sound production (Bybee 2007, Diessel 2007), a notion also referred to as “frequency-driven phonological (or phonetic) erosion” in diachronic change (Heine and Kuteva 2007).

We conclude that utterance frequency and accuracy [in more general terms, also referred to as “fecundity” and “copying fidelity” (Ritt 2004: 123)] are two orthogonal aspects of diachronic stability in phonotactics, which may or may not be linked to each other by a trade-off, where accuracy is diminished by high utterance frequency and vice versa. The focus of this paper, however, is not on phonotactic utterance frequency but rather on articulatory and perceptual determinants in phonotactics, that is, on factors that determine the accurate transmission of phonotactic items. Frequency may well interact with these determinants in one way or another, but this is not of primary concern to our endeavour (see Pagel, Atkinson and Meade 2007; Diessel 2007, and works cited above for more discussion). Nevertheless – and crucially independent thereof – we will, in this paper, operationalise diachronic stability of phonotactic sequences by means of diachronic increase or decrease in frequency. This is not contradictory, since inaccurate production and perception of a phonotactic sequence yield, everything else being equal, a diachronic decrease in frequency irrespective of whether or not that sequence was initially highly frequent. Synchronic frequency and diachronic growth (or decline) in frequency are clearly conceptually different from each other. So, we will consider a sound sequence relatively stable if it gains in frequency, and unstable if its frequency of use declines. We think that this notion provides a better proxy for stability than synchronic token frequency does because linguistic items can persist diachronically even if they are rarely used.

The question now is which articulatory factors determine accuracy and hence diachronic stability of phonotactic items. We focus on the three most prominent articulatory features in phonological

research throughout various theoretical approaches: manner of articulation, place of articulation, and phonation or voicing (Chomsky and Halle 1968, Dressler 1989, Dziubalska-Kołodziejczyk 2014, Hogg and McCully 1987, Prince and Smolensky 2002). Although the cognitive implementation of these articulatory features is by no means clear (Bybee 1994, Daland and Pierrehumbert 2011, Välimaa-Blum 2005, Wedel 2006), they have clear physiological interpretations and without doubt serve as valuable models in phonological theory.

Three phonotactic concepts related to these articulatory features shall be discussed exemplarily: ‘net auditory distance’, ‘voicing harmony’, and ‘sonority sequencing’. Although we focus on pairs of consonants in this paper, these principles can be extended to larger sequences of phonemes.

First, the framework of beats-and-binding phonology (Dziubalska-Kołodziejczyk 2014; Marecka and Dziubalska-Kołodziejczyk 2014) predicts that phonotactic sequences in general, and sequences of consonants in particular, are preferred if the articulatory difference between the sounds they are composed of is large. In order to operationalise articulatory difference, Dziubalska-Kołodziejczyk and colleagues derive the so-called “net auditory distance” (NAD) for a sequence of consonants. This metric combines differences between the respective manners of articulation, and places of articulation of the consonants and vowels involved. Via NAD, it is determined whether or not a consonant sequence is preferred (or “well-formed”); generally, the larger NAD between the consonantal segments (and the smaller NAD between neighbouring V and C), the better the cluster. For instance, NAD would predict that word-final /mj/ is more preferred than word final /mf/ because the latter cluster does not exhibit a sufficiently large difference between PoA of its segments, while the former does so (Dziubalska-Kołodziejczyk, Pietrala and Aperliński 2014). Although defined in terms of articulatory features, the NAD principle is motivated on perceptual grounds. It is argued that large articulatory differences facilitate the perception of a phonotactic sequence and its decomposition into segments. However, the dominance of assimilation as opposed to dissimilation processes in casual – and, crucially, speaker-friendly – speech suggests that speakers favour small articulatory differences in phonotactic sequences (Dziubalska-Kołodziejczyk 2014: 17).

A second important strand of phonotactic research involving articulatory differences is that of “voicing harmony” (Blevins 2004, Coetzee 2014, Hansson 2004). Regressive and progressive voicing agreement among pairs of consonants occurs in various languages across long distances (i.e. crossing intervening segments; Cho 1991, Hansson 2004) as well as within consonant clusters (Grijzenhout and Krämer 2000).

In general, voicing agreement has been argued to be motivated in multiple ways (see Coetzee 2014: 696–700 for an excellent discussion). On the one hand, this agreement serves the speaker since changing voicing in the transition from one consonant to another incurs increased articulatory effort to the speaker (although Dziubalska-Kołodziejczyk (2014: 17) challenges that notion by posing the question of whether actually retention or modulation is physiologically more costly). Similarly, regressive voicing agreement could be driven by “anticipatory activation” in production (Coetzee 2014, Hansson 2001) in which an articulatory feature of a consonant that is about to be produced is activated already before the consonant is actually produced and mapped onto a preceding consonant. On the other hand, progressive voicing agreement might be driven by perception errors, in which the listener maps formant values

characterising voicing in a consonant to a nearby and originally voiceless consonant (Coetzee 2014: 697).³

It is important to note that voicing agreement crucially depends on the position of the consonants involved or, more precisely, on other processes that apply to consonants in a certain position. For instance, voicing agreement among word-final consonant clusters coincides with the process of word-final devoicing which could – perhaps accidentally – produce voicing harmony (in the sense that both consonants share the same voicing feature) if the penultimate consonant is voiceless (as in, e.g., the German pronunciation of the acronym *OMFG* of *oh my fucking god*, /ɔmfk/). Likewise, devoicing could disturb voicing agreement if the penultimate consonant is voiced (e.g. Afrikaans *hond* /hont/, ‘dog’ vs. *honde* /hondə/, ‘dogs’). However, in some languages such as English, word-final devoicing is restricted to cases in which it assists voicing harmony, as illustrated by the differential voicing of consonantal suffixes attached to base forms ending in voiceless and voiced consonants (e.g. *picked* /pɪkt/ vs. *rigged* /rɪgd/).

Finally, let us consider the principle of ‘sonority sequencing’ (SSP) which, in a nutshell, asserts that consonant sequences in the onset position must rise, while consonant sequences in the coda position must successively fall in sonority (Clements 1990). For the present discussion, the SSP is relevant because sonority is tightly linked to manner of articulation and, to a lesser extent, to voicing. In general, consonants are more sonorous the less the airstream is obstructed by the articulators, so glides and liquids are more sonorous than nasals followed by fricatives, affricates, and stops (in decreasing order). Within these categories, voiced sounds are considered more sonorous than voiceless sounds (Burquest and Payne 1993: 101).

Crucially, apart from directionality, the SSP entails a strict ordering with respect to sonority and prefers large sonority differences to small ones. Cross-linguistically, large sonority rises (in onset position) or drops (in coda position) are more common than small rises or drops, which in turn are more widely attested than sonority plateaus (Berent et al. 2007: 294, Clements 1990). More specifically, Berent et al. (2007) have shown experimentally that onset-clusters with rising sonority are processed faster by speakers of English than onset-clusters remaining on the same sonority level. They do, however, point out that this behaviour is language specific and might depend on linguistic experience. Russian speakers did not show differential processing in their experiment, which is in turn reflected in the generosity (or laxness) of the Russian language with respect to the SSP.

Sonority sequencing interferes with the above described principles in an interesting way. Although in word-final position and under the assumption that the penultimate consonant is fixed, it is trivially consistent with the principle of maximising NAD. Since manner of articulation corresponds to the sonority scale, it counteracts the principle of voicing harmony in those cases in which progressive voicing assimilation applies to coda clusters, e.g. in English plural *bells* /belz/ or in the loan *Pils* /pɪlz/ imported from German into English. Sonority sequencing would prefer /ls/ to /lz/ while voicing harmony would prefer /lz/ to /ls/ (*mutatis mutandis* the same holds for regressive assimilation in onset clusters).

³ Coetzee (2014) mentions yet a third plausible mechanism namely that of “lexical accumulation” which led to voicing agreement in Afrikaans CVC sequences as a consequence of multiple unrelated sound changes.

What we can infer from this discussion is, at least, that articulatory (or, more generally, phonological) differences matter in phonotactic production and perception. More specifically, manner of articulation seems to play a particularly important role as it is a defining feature in the principle of sonority sequencing, which is arguably quite prominent in phonotactic research. Interestingly, the dominance of manner of articulation over place of articulation and voicing in the formation of phonotactic sequences is supported by recent neurological research. Mesgarani et al. (2014) have found that during processing in the superior temporal gyrus in the brains of speakers of English, consonants sharing the same manner of articulation are locally patched together, while in each of these patches consonants with different places of articulation and voicing patterns are mixed and processed closely together. This implies that on the neurological level, manner of articulation has higher discriminatory power than place of articulation, while voicing is shown to have discriminating effects within the subset of fricatives and plosives (Mesgarani et al. 2014: 1009). As a consequence, consonant clusters exhibiting large manner-of-articulation differences should be less confusable and hence more stable than clusters with large place-of-articulation differences. Whether this neurological setup reflects universal properties related to articulatory organs or is exclusively shaped by linguistic experience and thus language specific, is – as far as we know – a matter for future research.

1.3 The link between phonotactic acquisition and change

In this paper, observations from language acquisition data and diachronic data are contrasted with each other, the underlying hypothesis being that what is acquired early tends to be diachronically more successful. This deserves some discussion, since the exact relationship between language acquisition and language change is under debate.

While it is generally acknowledged in generative research that language acquisition and change are inherently linked to each other – since, in this framework, the grammatical output of the language-acquisition process ultimately determines the next generation's grammar (Roberts 2007, Yang 2000) – the respective roles that children and adults play in language change are much more contested in more functional and usage-based paradigms (Bybee 2010). Trivially, only what is acquired can be passed on to the next generation but the question remains as to whether it is really the acquisition process which constitutes the source of linguistic variation.

Indeed, arguments have been brought forth that adults are, to a larger extent, responsible for linguistic change (Bybee 2010: 114-119). First, changes do occur at the adult stage and the phonological domain particularly seems to be subject to changes in linguistic behaviour, as known from research on articulatory loss and phonological attrition (Ballard et al. 2001, Seliger and Vago 1991). Furthermore, it is argued that phonological errors, in particular, occurring during the process of language acquisition do not resemble attested diachronic developments (Diessel 2012) and that these errors do not persist at later ages (Bybee 2010).

Nevertheless, it has been shown empirically that age of acquisition (which we are focusing on in this paper) indeed correlates with diachronic stability. Words that are acquired early tend to be more resistant to phonological change than later acquired words (Monaghan 2014), and early used phonotactic strings have shown to be abundant in historically old word forms (MacNeilage and Davis 2000). Thus, a certain link between language acquisition and language change cannot be denied. Indeed, from a cognitive perspective, it makes sense that items which are acquired late are, to a larger extent, prone to change as a consequence of being cognitively less

entrenched (Bybee 2007, Croft 2000, Diessel 2012, Rosenbach 2008). This is so because items that are acquired late have less opportunities to be processed and produced – and consequently less entrenched – than words which are acquired early. Thus, it may well be the case that it is variation in adult speech which drives linguistic change but the underlying mechanism that enables this variation might, to some extent, be originally grounded in language acquisition.⁴

1.4 Summary and research hypotheses

We conclude that contrasting phonotactic acquisition with phonotactic change is worthy of investigation and, moreover, we expect the same articulatory and perceptual pressures to apply and be visible in both domains. More specifically, we suppose that if a certain articulatory distance in phonotactic items, say manner-of-articulation difference between consonants, is shown to be positively correlated with ease of acquisition so that consonant clusters featuring large manner-of-articulation differences are acquired early, then this articulatory difference is expected to positively correlate with diachronic stability. Hence, clusters featuring large differences should diachronically become more established and clusters featuring small – in this example, manner-of-articulation – differences are expected to become less frequently used on the diachronic time scale. The same reasoning applies *mutatis mutandis* to place of articulation and voicing.

By the same token, we expect that the relative ordering of the respective strength of effect of manner, place and voicing to language acquisition is preserved in diachrony. That is, if, for instance, voicing difference is shown to exert a larger facilitating effect on the acquisition of consonant clusters than place of articulation does, then place of articulation is expected to contribute less to the diachronic stability (or more to diachronic change) of a cluster than voicing does.

Thus, given the discussion in 1.2 on the relevance of differences in manner, place and voicing to processes and phenomena in phonotactics, we propose the following two research hypotheses:

- (1) **Articulatory differences in phonotactic acquisition.** Large manner-of-articulation difference between the segments of a cluster decrease its age of acquisition. Facilitating effects should be smaller for difference in voicing and smallest for place-of-articulation difference.
- (2) **Articulatory differences in phonotactic diachrony.** Large manner-of-articulation difference between the segments of a cluster increase its diachronic stability. These stabilising effects should be smaller for difference in voicing and smallest for place-of-articulation difference.

It is possible that these two hypotheses causally hang together if acquisition is indeed a driving force of diachronic change, as outlined in the previous section. If (1) is confirmed, then (2) is expected to hold as well.⁵ They will be approached by two separate studies drawing on Dutch

⁴ In that sense, the generative model of diachrony discussed above is an abstraction of what might happen in language change, in that it does not precisely show what happens in an individual through his or her lifetime. However, from a less fine-grained perspective, it fits as it predicts a similar outcome.

⁵ Clearly, it is logically impossible to verify an implicational relationship between (1) and (2) based on a single study, even if both hypotheses turn out to be correct. For instance, and quite plausibly, there might be factors

and Afrikaans data: Study 1 (section 2) tackles research hypothesis (1) by investigating phonotactic acquisition in Dutch while Study 2 (section 3) addresses hypothesis (2) by means of a (historic) comparison of Dutch and Afrikaans.

Since both studies are based on empirical data, it is clear that the causal relationship between (1) and (2) can only be partially accounted for, the reason for this being the obvious lack of acquisition data of historical Dutch or Afrikaans. In other words, the present project rests on the simplifying assumption that Dutch did not change much phonotactically (or at least less than Afrikaans) through the past two to three centuries (but see, for instance, Szemerényi (1996) for a justification of this assumption). Being fully aware of this, we assume for the sake of argument that the results from Study 1 about the acquisition of Dutch phonotactics can be transferred to historical Afrikaans (see 3.1 below for more comments). Both studies will be described in detail in the subsequent sections.

2. Study 1: Articulatory effects on phonotactic acquisition in Dutch

Here, we tackle the question of whether large differences are beneficial to the acquisition of Dutch consonant clusters. More specifically, hypothesis (1) in the previous section is addressed, which states that the facilitating effects of manner of articulation and voicing should be larger than those of place of articulation. This hypothesis can be corroborated. It can be shown that large place-of-articulation differences delay cluster acquisition as opposed to voicing and manner-of-articulation differences. However, it is voicing difference (and not MoA) which excels in enhancing phonotactic acquisition. Only in highly frequent consonant clusters do both voicing and manner show similarly strong facilitating effects (cf. 2.4). The data, variable operationalisation, statistical modelling procedure, analysis and results are discussed in more detail below.

2.1 Data

We combined two data sets in order to assess the effect of articulation on the acquisition of phonotactic sequences in Dutch. Age-of-acquisition (AoA) data were taken from Brysbaert et al. (2014) who provide AoA ratings for 30,000 Dutch lemmas. These ratings were collected in a large study in which participants were asked to estimate the age at which a given stimulus word was acquired, a methodology which has been shown to yield high correlations with estimates obtained under laboratory conditions (Kuperman, Stadthagen-Gonzalez and Brysbaert 2012). Phonological transcriptions were taken from the CELEX database (Baayen, Piepenbrock and Gulikers 1995) and added to the lemmas. Additionally, word-length in terms of the number of phonemes and token frequency was retrieved from CELEX. Only a core vocabulary of the 5,000 most frequent words was considered (based on INL token-frequency scores in CELEX) in order to exclude non-prototypical low-frequency items.⁶ Foreign words were excluded as well.

Since we are analysing consonant clusters in word-final position in this study, only words ending in a sequence of two consonants were considered. In total, 828 Dutch lemmas met these requirements, featuring a range of 33 word-final consonant-cluster types. Articulatory features of the two final consonants were assigned based on the IPA chart.

independently supporting both hypotheses. Diessel (2012) suggests a number of psychological and cognitive factors that would do so.

⁶ Inspecting CELEX (INL frequency), it can be easily shown that the 5,000 most frequent lemmas cover about three fourths of all Dutch tokens.

2.2 Variables

The aim of this study is to assess the effect of articulatory differences on the AoA of consonant clusters. Thus, a number of variables were considered in the statistical analysis. First and foremost, for each cluster type, age-of-acquisition (AoA) was determined by calculating the first decile of all AoA ratings given by Brysbaert et al. (2014) of the words showing that cluster type word-finally ($\mu_{\text{age}} = 6.7 \text{ yr}$; $sd_{\text{age}} = 1.5 \text{ yr}$; $rg_{\text{age}} = (4.3, 10.7)$). The first decile was chosen as it provides a more robust, albeit at the same time more conservative, estimate of the minimum of cluster-wise AoA ratings than the actual minimum does. This is so, since the actual minimum is obviously sensitive to outliers at the lower range of AoA. In other words, this procedure ensures that cluster AoA is not just determined by the knowledge of a single word. As AoA featured a slightly skewed distributional pattern, the data were Box-Cox transformed (Box and Cox 1964) in order to fit the statistical modelling requirements (after transformation: $\mu_{\text{AoA}} = 0.84$; $sd_{\text{AoA}} = 0.03$; $rg_{\text{AoA}} = (0.77, 0.91)$).⁷

The primary predictors we are interested in are difference in manner of articulation, difference in place of articulation and difference in voicing of the two final consonants. In order to parametrise these differences, manner of articulation as well as place of articulation were first translated into ordinal scores as in Dziubalska-Kořaczyk (2014) and Dziubalska-Kořaczyk et al. (2014). While manner-of-articulation scores (MoA1, MoA2) depend on sonority (from 1 to 7, higher sonority yielding a higher score), place-of-articulation scores (PoA1, PoA2) were defined by the phoneme's place of articulation in the vocal tract (from 1 to 7; the closer to the front, the lower the score). Voicing was simply considered as a binary variable (1/0; voiced/unvoiced). See Table A1 in the appendix.

Subsequently, difference in manner of articulation (ΔMoA) was operationalised as the absolute value of the difference between manner-of-articulation scores, normalised to the unit interval with respect to the maximal absolute difference, thus restricting values to scores going from 0 to 1, where 1 is the maximal difference of all cluster types considered and 0 denotes identity with respect to manner of articulation. Taking the absolute value ensures that the arbitrarily assigned directionality of the ordinal manner-of-articulation scores does not skew the results, while the normalisation procedure facilitates interpretation of and comparisons between the respective effects of manner, place and voicing without the need to compute normalised regression coefficients in the statistical analysis (Nakawaga and Cuthill 2007), and at the same time retains any relevant information. Difference in place of articulation (ΔPoA) was determined *mutatis mutandis* as above. Difference in voicing (ΔVoice) was simply operationalised as a binary variable ('different' if consonants are voiced differently and 'same' if they are not).

A side note is in order on the operationalisation of ΔMoA of consonant clusters in word-final position by means of sonority. One might wonder whether relying on absolute values as

⁷ This procedure of using AoA ratings corresponding to word types to estimate AoA of a cluster type might potentially strike one as odd, since it tends to suggest words to be acquired, say, late *because* they contain a particularly ill-formed cluster. However, although the latter suggestion may – all other things being equal – be correct, we will remain agnostic about it. Rather, the decision to parametrise cluster-wise AoA, as done above, is a primarily pragmatic one. Since phonotactic sequences hardly occur in isolation, there is simply no better way of estimating their AoA than via the linguistic items they are embedded in.

opposed to actual differences makes sense from a theoretical perspective, since the principle of sonority sequencing (see 1.2) would predict sonority to decline word-finally. However, it was found in our data that absolute and actual differences in $M_{\circ A}$ deviate from each other in only 3 out of 33 cluster types, which as a matter of fact illustrates that the sonority-sequencing principle is fulfilled anyway to a large extent in Dutch. Thus, taking actual differences would not substantially change the results reported below.

Finally, two additional controlling variables entered the analysis. For each cluster type, token frequency (`frequency`) was determined by computing the median INL token frequency (see section 2.1) of all words featuring that cluster. We preferred the median of all words to, e.g., the frequency of only the subset of initially acquired words featuring that cluster as it better represents the average overall exposure of the language learner to the cluster type. Due to expected distributional properties, `frequency` was Box-Cox transformed as well, before it entered the statistical analysis (after transformation: $\mu_{\text{frequency}} = 2.73$; $sd_{\text{frequency}} = 0.11$). The remaining controlling variable, phonological length (`length`), was simply operationalised as the median number of phonemes of all words featuring that cluster ($\mu_{\text{length}} = 5.42$; $sd_{\text{length}} = 1.29$). All scores are summarised in Table A2 in the Appendix.

2.3 Modelling procedure

The goal of the statistical analysis is to assess the differential effects of $\Delta M_{\circ A}$, $\Delta P_{\circ A}$ and ΔVoice on $A_{\circ A}$. We opted for linear models (LM; Baayen 2013; West, Welch and Gałecki 2015) as this model family provides a flexible way of combining numerical as well as categorical variables and simultaneously allows for the inclusion of controlling factors. Thus, $A_{\circ A}$ was implemented as a dependent variable into a LM in which $\Delta M_{\circ A}$, $\Delta P_{\circ A}$ and ΔVoice function as predictor variables. In addition, six interacting terms were included in which $\Delta M_{\circ A}$, $\Delta P_{\circ A}$ and ΔVoice are controlled by `length` and `frequency`, respectively. All distributional requirements were met (see previous subsection). Computations were done in R (R Core Team 2018).

This resulted in a model featuring nine regression terms which harbours the risk of being overspecified and hence insufficient fitting properties, rendering conclusions drawn from the estimated coefficients unreliable. In order to find the most informative and at the same time most parsimonious model of $A_{\circ A}$ with the best fit, AICc-driven model selection was employed. This requires some elaboration. AICc (“corrected Akaike Information Criterion”; Johnson and Omland 2004) is a measure of information – or more precisely, of information loss relative to the data – of a given model which balances goodness-of-fit and model complexity, and which is in addition corrected for applications to small samples. The smaller the AICc, the better the model. AICc is superior to plain goodness-of-fit measures such as (adjusted) R^2 in that the latter automatically increases when more predictors are added to a model. Thus, AICc accounts for model overspecification.

In the model-selection procedure, linear models for all theoretically interesting subsets of predictor regression terms together with their AICc are computed. In the present analysis, nine predictor terms were considered, three for the isolated variables and six controlling terms, as described above. We assumed that controlling interaction terms always co-occur with their corresponding controlled predictor in isolation. For instance, if $\Delta M_{\circ A}$ is controlled by

frequency in a model it includes the configuration $\Delta\text{MoA}+\Delta\text{MoA}:\text{frequency}$. This restriction ensured that the controlling variables (*frequency* and *length*), which we are actually not interested in in this study, do not occur in isolation in the analysis. This resulted in a set of 124 candidate models. The optimal – or “AICc-best” – model is then the one model with the lowest AICc score, i.e. the least loss of information.

This information-theoretic model selection procedure allows for yet deeper investigations, namely “multimodel inference” (Burnham and Anderson 2002; Burnham, Anderson and Huyvaert 2011). To begin with, an important observation is that although there is always a single best model, that model does not necessarily have to be much better than other candidate models. There might be some other relevant information contained in some of the remaining candidate models which would be lost if one only considers the single best model. By comparing a candidate model’s AICc score with that of the best model, the relative strength of evidence of that candidate model – the so-called “Akaike weight” w – can be computed. It can be interpreted as the probability of that model given the data and the set of all competing candidates. Thus, the Akaike weight measures how much evidence there is in the data for a candidate model (Johnson and Omland 2004: 104). Clearly, the best model has the largest Akaike weight.

In multimodel inference, one can exploit Akaike weights in order to combine all candidate models. A whole set of models obviously contains more information than a single best model. From the model set and the corresponding set of Akaike weights, average regression coefficients \bar{c}_i can be computed. These regression coefficients can then be used to calculate average predictor effects, under the assumption of average token frequency and phonological length. For instance, the average effect \bar{e} of ΔPoA can be computed as $\bar{e}_{\Delta\text{PoA}} = \bar{c}_{\Delta\text{PoA}} + \mu_{\text{frequency}} \cdot \bar{c}_{\Delta\text{PoA}:\text{frequency}} + \mu_{\text{length}} \cdot \bar{c}_{\Delta\text{PoA}:\text{length}}$, where μ_j denotes the mean.

Moreover, by using the Akaike weights, it is possible to compute “relative variable importance” (RVI), a measure of how often a predictor appears in the models contained in the candidate set (Burnham and Anderson 2002: 168). This measure is very informative: not only can one determine the average strength of the effect of some predictor variable, it is also possible to make assertions about how important that predictor is for obtaining information about the output variable, relative to the other predictors. Note that large importance of a variable does not necessarily imply a large effect and vice versa. Finally, predictor variables can be ranked by their RVI. This procedure allows for an in-depth analysis of the effects that articulatory differences have on acquisition as well as of their respective informational importance.

2.4 Results

The AICc-best model of AoA obtained by the procedure described above features six terms⁸, as can be seen in Table 4. Most notably, ΔMoA only contributes to AoA in isolation, having a slight enhancing effect on acquisition, so that no significant interactions surface in this model. Translating transformed AoA back into age of acquisition measured in years (based on the

⁸ Note that the Akaike weight of the best model equals about 0.123, while that of the maximal model containing all terms equals 0.001. Thus, evidence for the best model is roughly 123 times stronger than for the maximal and least parsimonious model. This illustrates the necessity of careful model building in quantitative research.

respective ranges), the effect on AoA corresponds to roughly 3 years if ΔMoA is maximal (note at this point that the intercept in the model conveniently coincides – approximately – with μ_{AoA}). A similar acquisition-enhancing effect can be seen in the averaged model. Taking average length and frequency into account, maximal ΔMoA reduces AoA by about $\bar{e}_{\Delta\text{MoA}} = -0.06$ (Table 4). Also note that of all three variables, manner-of-articulation difference scores highest on RVI, suggesting that ΔMoA plays an important role in phonotactic acquisition.

In the best model, ΔPoA shows three significant effects on acquisition (Figure 2a). In isolation, ΔPoA considerably increases AoA, an effect which becomes even significantly larger in long words but which is diminished significantly the more frequent clusters are in terms of tokens (Table 5). Inspecting the average model, and assuming average length and frequency, it can be seen that ΔPoA has an inhibiting effect on phonotactic acquisition of about $\bar{e}_{\Delta\text{PoA}} = 0.04$ (Table 5). From all three primary predictors, place-of-articulation difference seems to be least relevant to explaining age of acquisition and to be the most dependent on the controlling variables frequency and length, as can be judged from the respective RVI scores in Table 5 (see Figure 2b).

Table 4. AICc-best model: $R_{\text{Adj}}^2 = 0.42$; $\text{AICc} = -140$; $w = 0.12$.

Variables	Estimate (c_i)	SE	t	p
Intercept	0.86	0.01	82.9	<0.001
ΔMoA	-0.08	0.02	-3.29	<0.01
ΔPoA	0.43	0.15	2.83	<0.01
ΔVoice (different)	-0.38	0.19	-2.05	0.05
ΔPoA :length	0.01	0.00	3.41	<0.01
ΔPoA :frequency	-0.18	0.06	-3.32	<0.01
ΔVoice (different):frequency	0.15	0.07	2.21	0.04

Table 5. Predictor coefficients in the combined model resulting from model averaging procedure over 124 candidate models. Average coefficients together with standard errors, relative variable importance and averaged overall effects are shown.

Variables	Average estimate (\bar{c}_i)	SE	RVI	Average effect (\bar{e}_i)
ΔPoA	0.16	0.21	0.73	0.04
ΔPoA :length	0.01	0.01	0.46	
ΔPoA :frequency	-0.06	0.07	0.39	
ΔMoA	-0.01	0.37	0.96	-0.06
ΔMoA :length	0.01	0.01	0.45	
ΔMoA :frequency	-0.03	0.13	0.19	
ΔVoice (different)	-0.27	0.27	0.83	-0.11
ΔVoice (different):length	0.00	—	0.16	
ΔVoice (different):frequency	0.05	0.09	0.32	

Finally, ΔVoice shows significant enhancing effects on acquisition (Table 1). Crucially though, these effects are diminished considerably in frequent clusters, as shown by the

significant interaction term in the best model. This is confirmed, by the average effect computed from the averaged model, $\bar{e}_{\Delta\text{Voice}} = -0.11$. In the upper part of the frequency spectrum, however, this effect becomes weaker.

We conclude that, everything else being equal, differences in manner of articulation as well as in voicing have a facilitating effect on the acquisition of word-final consonant clusters in Dutch, while clusters are acquired later if the phonemes they are composed of differ with respect to their place of articulation (e.g. /pt/, /lp/ or /kt/). In the upper part of the frequency spectrum, and thus in cognitively more entrenched clusters, manner becomes more enhancing while voicing becomes less enhancing ($\bar{e}_{\Delta\text{Voice}} = -0.10$ vs. $\bar{e}_{\Delta\text{MoA}} = -0.07$ at maximum frequency and average length). For making predictions about acquisition, manner is most important, followed by voicing and finally place of articulation (cf. RVI in Table 5 and Figure 2b). This corroborates hypothesis (1) presented in section 1.4, albeit only partially, as voicing difference turns out to enhance acquisition more than expected.

Interestingly, this is in line with results from neurological and cognitive research on the organisation of phonemes, as discussed in section 1.2. The question is whether the differential behaviour of manner of articulation, place of articulation and voicing in phonotactic acquisition also yields diachronic long-term reflexes. This question will be dealt with in the subsequently presented study.

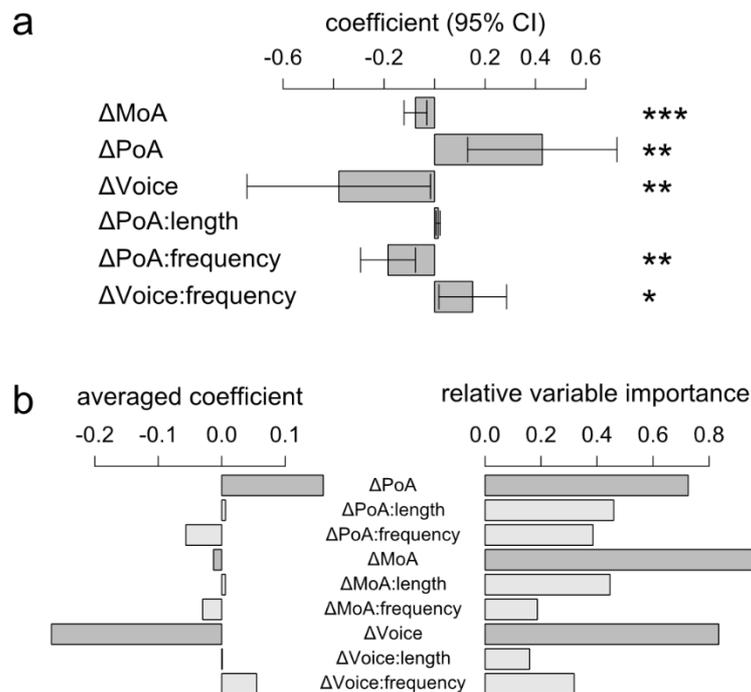


Figure 2. (a) Plot of the coefficients in the optimal model resulting from model-selection procedure. Vertical axis measures effect on A_oA (transformed age-of-acquisition). Error bars denote 95% confidence intervals. Significance code: ‘***’ $p < 0.001$, ‘**’ $p < 0.01$, ‘*’ $p < 0.05$ with respect to zero-effect hypothesis. (b) Averaged model resulting from multimodel inference. Dark gray and light gray bars correspond to isolated variables and interacting terms, respectively. On the left: averaged coefficients measuring the effect on A_oA. On the right: RVI scores corresponding to variables and interaction terms.

3. Study 2: Articulatory effects on stability in the history of Afrikaans

In this section, we assess the second research hypothesis – i.e. (2) in section 1.4 – in this study and expect clusters featuring large manner-of-articulation differences to be diachronically more stable than clusters with differentially voiced segments followed by clusters that show large differences with respect to place of articulation. Indeed, as presented in what follows this hypothesis is largely supported by data from (historical) Dutch and contemporary Afrikaans. While large place-of-articulation difference and difference in voicing diminishes diachronic stability, clusters showing large differences in terms of manner of articulation are more successful in contemporary Afrikaans than what would be expected under the null-hypothesis that no change occurred in the development of the Afrikaans phonotactic system. The hypothesis and data considered in this study require an analytic approach which differs from that in the first study. Our second study will be described in the following subsections together with its outcome.

3.1 Data

Two additional corpora were used to address the research question rephrased above. First, Afrikaans final clusters were retrieved from the NCHLT corpus (Eiselen and Puttkammer 2014). The corpus consists of 58,096 annotated word tokens (distributed among 6,464 word types) retrieved from written Afrikaans, mostly from government websites. The fact that its tokens are lemmatised as well as morphologically decomposed was crucial in order to extract all tokens ending in a cluster, excluding those tokens in which the final cluster involves a morphological operation and thus spans a morpheme boundary. This was necessary, as the Dutch data did not include word forms containing boundary-spanning clusters either. Phonological information was taken from Coetzee (1969). In total, 445 word types ending in a consonant cluster were retrieved, featuring 26 different cluster types. Subsequently, type frequencies were obtained for each cluster type, i.e. the number of word types a cluster type surfaces in word-finally.

Second, in order to obtain slightly more representative historical Dutch data, the 5,000 most frequent word types (cf. section 2.1) were retrieved from the pre-1900 subset of the De Gids corpus (henceforth DGC; see van de Velde 2009). The data were matched with the CELEX lemma list in order to obtain phonological transcriptions. After excluding those items not ending in a consonant cluster, 819 word types featuring 33 cluster types remained. Cluster-specific type frequencies were also based on DGC. It is worth mentioning that the 33 cluster types found in historical Dutch represent a proper superset of the 26 Afrikaans cluster types (only double clusters considered).

We emphasise that the procedure of retrieving frequency data from historical corpora, and phonological data from contemporary language sources implies that the present approach is, at best, pseudo-historical. Also, 19th century Dutch data can obviously not be equated with historical Afrikaans which – at least, phonologically – emerged from Early Modern Dutch vernacular spoken in the 17th or 18th century (Roberge 1993). Nevertheless, historical arguments are often based on comparative grounds solely considering contemporary data (cf. e.g. Szemerényi 1996). In that sense, the present approach finds itself somewhere halfway between a synchronic comparative study and a purely diachronic one.

Having clarified this, the reader might be relieved to see that the distributional difference between the coda phonotactics in contemporary and 19th-century Dutch is not substantially

large anyway (phi-coefficient based on chi-squared test of independence: $\varphi = 0.12$, 95% CI: (0.11,0.12); $\chi^2 = 67,662.94$; $N = 5,042,350$; $df = 32$; cluster-wise token frequencies in CELEX and DGC considered).

3.2 Variables

We now address the following question: do articulatory differences determine the historical stability of word-final consonant clusters in the diachrony of Afrikaans? We do so by measuring the productivity of word-final clusters in historical Dutch and contemporary Afrikaans, and subsequently testing whether articulatory differences have an effect on the relationship between both productivity scores. We straight-forwardly define the productivity of a cluster as the number of word types it occurs in. Note that these scores will differ substantially when considering raw figures in historical Dutch and Afrikaans due to the different corpus sizes of DGC and NCHLT, respectively. This does not pose any problem, since relationships among these scores rather than raw differences between them will be considered (but see 3.3 below). Cluster productivity in historical Dutch and contemporary Afrikaans shall be denoted as *DutProd* and *AfrProd*, respectively. Whenever a cluster only occurred in DGC but not in NCHLT, its *AfrProd* score was set to 0 in order to maintain as many data points as possible. We decided to leave both variables untransformed and to resort to nonparametric methods instead. *DutProd* and *AfrProd* scores for the set of 33 clusters are illustrated in Figure 3a. Articulatory differences (ΔMoA , ΔPoA and ΔVoice) were defined exactly as in 2.2. All scores are summarized in Table A3.

3.3 Modeling procedure

In a preliminary analysis, the strength of the relationship between productivity in historical Dutch (*DutProd*) and productivity in contemporary Afrikaans (*AfrProd*) was assessed. This was done by means of Spearman's ρ (rank correlation) due to the nonparametric nature of the productivity scores. The relationship between *DutProd* and *AfrProd* can be shown to be relatively strong at $\rho = 0.68$ (95% CI: (0.44,0.83); $t = 5.2$, $df = 31$). This was to be expected: on average, clusters which have been productive in historical Dutch are also productive in contemporary Afrikaans.

Subsequently, the influence that ΔMoA , ΔPoA and ΔVoice exert on the relationship between *DutProd* and *AfrProd* was analysed in three separate generalised additive models (GAMs, Wood 2006a). GAMs are statistical models which do not only involve linear terms (as – generalised – linear models do) but combine linear, quadratic and even more complicated components. Hence, in contrast to straight lines or piece-wise linear surfaces, they potentially yield curves and “wiggly” surfaces, of course depending on the underlying data. They provide an efficient way of detecting non-linear or, more generally, non-monotone interactions among variables (Baayen 2013). Indeed, the relationship between productivity in Dutch and Afrikaans might be for instance curved rather than linear, and this holds even more so for the relationship between productivity and articulatory difference. Most importantly, however, interactions among two predictor variables that influence a third (dependent) variable can be easily modelled using GAMs by implementing so-called “tensor-product terms” (Wood 2006b), especially if predictor variables are located on different scales (e.g. in this case: articulatory differences on the one hand, and type frequencies on the other). Details are not relevant at this

point; it is sufficient to note that this modelling toolkit provides a convenient way of analysing the interaction of articulatory difference on the relationship between productivity in Dutch and Afrikaans. Finally, GAMs have the advantage of being innately non-parametric, thus not imposing any particular distributional requirements on the data to be analysed. This is particularly convenient given the skewed distribution of the productivity scores (Figure 3a, see also previous section). We opted for three separate GAMs as opposed to a single GAM featuring interactions among all variables due to the relatively small number of cluster types. Moreover, by comparing R^2 scores of the three separate models, the relative explanatory power of each of the three articulatory differences can be assessed (somewhat similar to – albeit not identical with – the information theoretic measure RVI in Study 1).

The question to be asked now is the following: in which way does the relationship between `DutProd` and `AfrProd` change, if we consider articulatory differences at different degrees? For instance, it could be the case that the relationship between `DutProd` and `AfrProd` is positive and increasing for small articulatory differences, say, in manner of articulation, but decreasing for large articulatory differences. This would indicate that ΔMoA would have a negative effect on cluster stability since, in this scenario, clusters with large manner-of-articulation differences are less frequent in Afrikaans than what would be expected based on Dutch productivity scores. Thus, in the first GAM, `AfrProd` is implemented as an outcome variable depending on the interaction of `DutProd` and ΔMoA (integrated as a tensor-product term). In the second GAM, ΔMoA is simply replaced by ΔPoA , in order to assess the effect of place-of-articulation. In the third GAM, finally, `AfrProd` again functions as outcome variable predicted by `DutProd` which is controlled (or “smoothed”; Wood 2006a) by the binary variable ΔVoice . All computations were done in R, in particular using the `mgcv` package for computing GAMs (R Core Team 2018; Wood 2006a).

3.4 Results

The models reveal differential results about the impact that articulatory differences have on the relationship between cluster productivity in Dutch and Afrikaans, and hence on the diachronic stability of clusters. Let us begin with manner of articulation. The first GAM shows a significantly non-zero intercept at 13.49 ($SE = 1.59$; $t = 8.50$; $p < 0.001$) and a significant tensor-product term modelling the interaction of `DutProd` and ΔMoA ($edf = 7.66$; $F = 23.37$; $p < 0.001$) as well as remarkable fitting properties ($R^2 = 0.87$; $N = 33$). From the significant interaction term, we see that `DutProd` and ΔMoA indeed affect `AfrProd`. In order to better understand the interacting behaviour, however, the model has to be visualised.

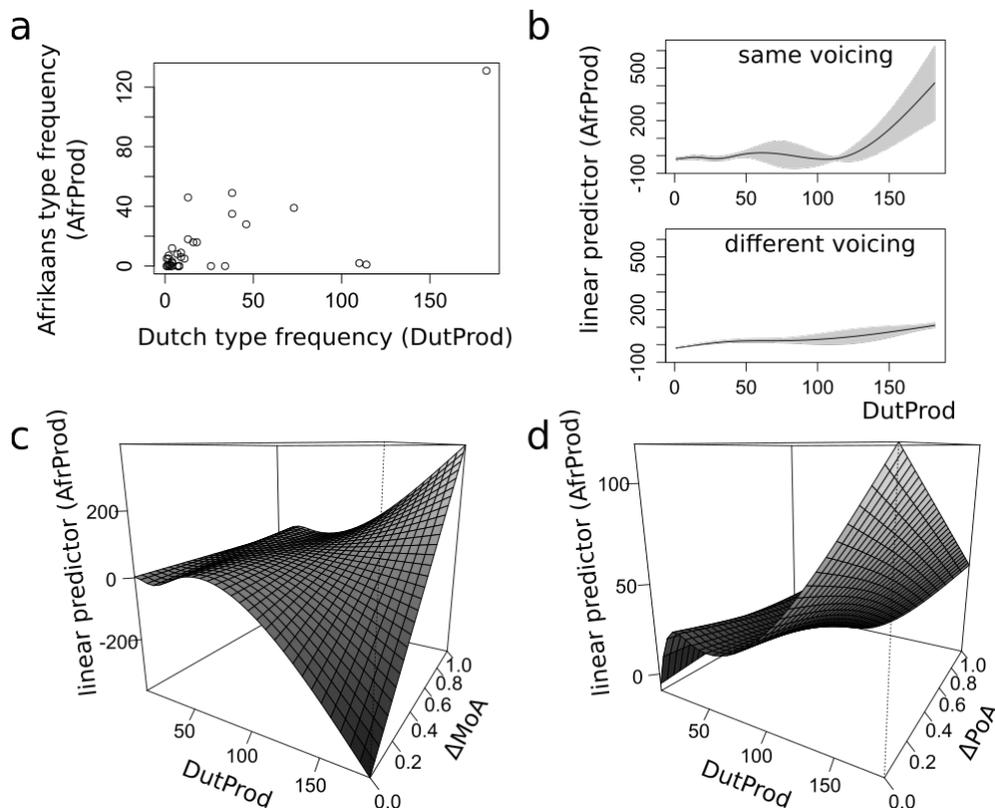


Figure 3. (a) Distribution of Dutch and Afrikaans type frequency in word-final consonant clusters. There is a significant correlation between both productivity scores. (b-d) Effect of voicing difference, manner-of-articulation difference and place-of-articulation difference on the relationship between productivity in Dutch and Afrikaans, respectively.

Figure 3c shows $AfrProd$ as a two-dimensional function of $DutProd$ and ΔMoA , illustrated by a curved surface. If ΔMoA is held constant at a certain value, say $\Delta MoA = 0.2$, one can inspect the relationship between $DutProd$ (horizontal) and $AfrProd$ (vertical), just as in Figure 3a, represented by one of the solid black lines in the grid superimposed on the curved surface. For low ΔMoA scores, this functional relationship is decreasing. Clusters which do not or only slightly differ with respect to manner of articulation are less productive in Afrikaans than expected based on historical Dutch, i.e. under the null-hypothesis that there was no change between historical Dutch and contemporary Afrikaans. In contrast, for high ΔMoA scores, the relationship is increasing. Clusters exhibiting large manner-of-articulation differences are more productive in Afrikaans than expected.

The second GAM, which analyses place of articulation, also shows a significantly non-zero intercept at 13.49 ($SE = 2.49$; $t = 5.42$; $p < 0.001$), a significant tensor-product term of the interaction of $DutProd$ and ΔPoA ($edf = 6.68$; $F = 10.11$; $p < 0.001$) and slightly reduced explanatory power ($R^2 = 0.68$; $N = 33$).

In contrast to the first GAM, however, the relationship between $DutProd$ (horizontal in Figure 3d) and $AfrProd$ (vertical) is roughly increasing for all ΔPoA scores (with a tiny dip in the lower-mid-frequency range). Additionally, the strength of this relationship seems to be strongest (i.e. showing the steepest slope) for small differences in place of articulation and

comparably weak if consonants differ in their place of articulation. Thus, increasing Δ_{POA} weakens the relationship between both productivity scores. The larger the difference in place of articulation, the smaller the productivity of Afrikaans clusters than what would be expected by the Dutch data. It can be concluded that, overall, Δ_{POA} exerts a diachronically destabilising influence on word-final clusters.⁹

Finally, the third GAM yields two separate one-dimensional curves (smooth terms) for the relationship between *DutProd* (horizontal in Figure 3b) and *AfrProd* (vertical), one for *same* voicing and one for *different* voicing (upper and lower graph in Figure 3b, respectively). Overall, the model shows an intercept of 19.30 ($SE = 2.49$; $t = 5.42$; $p < 0.001$) and, again, good fitting properties ($R^2 = 0.89$; $N = 33$). Both smooth terms show significantly non-trivial behaviour (*same*: $edf = 5.18$; $F = 4.78$; $p = 0.002$; *different*: $edf = 3.17$; $F = 63.36$; $p < 0.001$), the latter staying more constant than the former, as can be seen from the two graphs in Figure 3b. In the low- and mid-frequency range, voicing does not seem to differentially determine cluster productivity. However, in the high frequency range, clusters that have no voicing difference are more productive than their differently voiced counterparts. This shows that at least highly frequent clusters do not benefit from voicing contrasts, diachronically speaking; the opposite seems to be the case. This goes in line with various voicing-assimilation processes discovered in historical language research (e.g. Horobin and Smith 2002, Colantoni and Steele 2007). Infrequent clusters, however, do not seem to show this differential behaviour.

We thus conclude that, if diachronic stability of word-final clusters is assessed by comparing productivity in historical Dutch and contemporary Afrikaans, increasing manner-of-articulation differences between the building blocks of a cluster seem to have a promoting (or stabilising) effect, while increasing place-of-articulation differences and voicing differences exhibit demoting (or destabilising) effects on word-final consonant clusters. This effect is enhanced the more productive and, consequently, the more frequent a cluster is. Thus, it is clusters like /rt/, /lt/ and /mp/ which are diachronically most stable, and clusters such as /pt/, /kt/ or /lm/ which are expected to undergo diachronic deletion processes. In summary, this is in line with research hypothesis (2) in 1.4.

4. Discussion and conclusion

Based on considerations about the differential relevance of manner of articulation, place of articulation and voicing in phonotactic production and perception, we addressed the question of whether articulatory difference between the segments of a consonant sequence exerts a promoting effect in phonotactic acquisition and change. More specifically, we hypothesised difference in manner of articulation to have a stronger promoting effect than difference in

⁹ A clarifying note on the interpretation of the plots is in order. Although the relationship between productivity in Dutch and Afrikaans is (approximately) positive for all Δ_{POA} values, the overall effect of Δ_{POA} is negative because the relationship between Dutch and Afrikaans productivity gets weaker. Likewise, the effect of Δ_{MOA} in Figure 3c is positive even though the relationship between Dutch and Afrikaans productivity is negative for low Δ_{MOA} because the relationship gets stronger as Δ_{MOA} increases. What the negative slope for Δ_{MOA} shows, nevertheless, is that low differences in manner of articulation are substantially bad for the survival of clusters. In contrast, the positive slopes in Figure 3d indicate that the effect of Δ_{POA} is not substantial (because for all Δ_{POA} values, the relationship between Dutch and Afrikaans productivity does not differ much from the general positive relationship between Dutch and Afrikaans productivity shown in Figure 3a).

voicing and in place of articulation, as suggested by research on the neuro-cognitive organisation of phonemes (Mesgarani et al. 2014). We tested this set of hypotheses against Dutch acquisition data (word-final consonant clusters) and by means of a (pseudo-historical) comparative study of Dutch and Afrikaans. That is, we sought to identify the articulatory factors in the acquisition of Dutch coda phonotactics and the diachronically-stabilising determinants in the history of the Afrikaans phonotactic system.

Indeed, it was found that manner-of-articulation difference incurs a stronger promoting effect on phonotactic acquisition and diachronic stability than place-of-articulation difference does. That is, it is the set of clusters in the lower-right corner in Figure 4 below (dashed box) which is most successful in acquisition and change. Consequently, developments such as t-deletion in *specht* (/xt/ > /x/) or schwa-epenthesis in *worm* (/rm/ > /rəm/), as shown in Figure 1, follow a systematic and articulatorily as well as perceptually motivated trend rather than occurring just by coincidence.

The behaviour of voicing difference is slightly more complicated. While voicing difference has turned out to yield the strongest enhancing effects in the acquisition of Dutch, it shows demoting effects on the diachronic time scale in the history of Afrikaans. Interestingly, voicing difference seems to suffer from utterance frequency in the sense that clusters in the upper part of the frequency spectrum face the strongest demoting effects – both in acquisition and in diachrony – if their constituents differ in voicing. With respect to manner and place, the effect of frequency goes in precisely the opposite direction. This is surprising, given that sonority (which is linked to manner of articulation) – by definition, at least – slightly correlates with voicing.

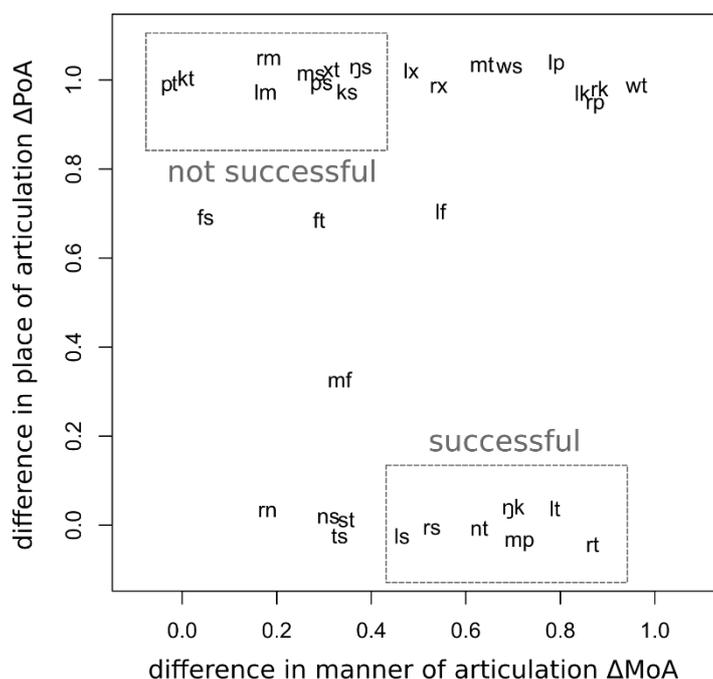


Figure 4. Dutch/Afrikaans word-final cluster types in the space defined by ΔMoA (horizontal axis) and ΔPoA (vertical axis). Data points are randomly jittered to make the labels readable. The dashed box in the lower right corner denotes clusters which are particularly successful in acquisition and change, i.e. clusters with segments showing similar place and different manner of articulation. Less successful types are located in the dashed box in the upper left corner.

Our results seem to converge with established phonotactic principles (cf. 1.2). The importance of manner-of-articulation difference in phonotactic acquisition and, more evidently, in phonotactic change is in line with the principle of sonority sequencing which favours decreasing sonority in coda position (Clements 1990). The diverging effect of voicing difference in acquisition and change, respectively, is admittedly more puzzling and probably hints at a conflict between voicing agreement and final devoicing in Afrikaans, especially in adult speech. Finally, the fact that large place-of-articulation differences are neither particularly beneficial to phonotactic acquisition nor to the diachronic stability of word-final clusters (in fact, the opposite seems to be the case) challenges net auditory distance (Dziubalska-Kołodziej 2014) as an overall measure of phonotactic well-formedness (because NAD puts equal weights on differences in MoA and PoA). Rather, a more differentiated approach is required to assess whether or not consonant clusters are preferred, if only in the case of Afrikaans and Dutch coda phonotactics.

It is important to remark that our results are restricted to word-final sequences of two consonants in the coda. Neither did we specifically address longer sequences of consonants (although it is reasonable to assume that similar restrictions with respect to manner, place and voicing hold in longer phonotactic items), nor did we account for dynamics in the onset position. Similarly, we did not address any additional factors that potentially determine phonotactic evolution in Afrikaans. Even if principles like sonority sequencing represent (statistical) universals, it cannot be ruled out that their effects are overshadowed by factors like morphology or language contact (and concomitant influx of phonotactic sequences). Both may apply to the history of Afrikaans: compared to Dutch, the language shows major morphological restructuring as well as contact with a number of Non-European languages. In fact, it is very unlikely that phonotactic evolution is not influenced by phonology-internal factors. Otherwise it would be difficult to explain why Afrikaans phonotactics changed considerably while Dutch phonotactics did less so.

Apart from the empirical results about the differential roles that manner, place and voicing play in the acquisition and change of Dutch and Afrikaans phonotactics, the present study more generally contributes to the discussion about the link between language acquisition and change (Bybee 2010, Diessel 2012, MacNeilage and Davis 2000, Monaghan 2014). Whether or not language change primarily happens during first-language acquisition cannot be clearly answered by our results. What can be confirmed is that in the domain of phonotactics, similar articulatory determinants influence acquisition and change.

Acknowledgements

We would like to thank Katarzyna Dziubalska-Kołodziej for many comments, particularly on NAD; Christina Prömer for proof reading; Freek Van de Velde for providing us with Dutch data; the participants of the first Afrikaans Grammar Workshop (Johannesburg, August 2016) for a fruitful discussion of this research project, as well as an anonymous reviewer for many valuable suggestions.

This article was made possible with the support from the South African Centre for Digital Language 379 Resources (SADiLaR). SADiLaR is a research infrastructure established by the Department of 380 Science and Technology of the South African government as part of the South African Research 381 Infrastructure Roadmap (SARIR). This research was in part financially supported by the National 382 Research Foundation (NRF) of South Africa.

Additionally, this research was supported by the Austrian Science Fund (FWF grant P27592-G18) and the Virtuele Instituut vir Afrikaans.

References

- Baayen, H. 2013. Multivariate Statistics. In R. Podesva and D. Sharma (eds.). *Research Methods in Linguistics*. Cambridge: Cambridge University Press. pp. 337–372.
- Baayen, H., R. Piepenbrock and L. Gulikers. 1995. CELEX2 LDC96L14. Available online: <https://catalog ldc.upenn.edu/LDC96L14> (Accessed 18 October 2017).
- Ballard, K.J., D. Robin, G. Woodworth and L. Zimba. 2001. Age-related changes in motor control during articulator visuomotor tracking. *Journal of Speech, Language, and Hearing Research* 44: 763–777. [https://doi.org/10.1044/1092-4388\(2001/060\)](https://doi.org/10.1044/1092-4388(2001/060))
- Berent, I., D. Steriade, T. Lennertz and V. Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104 (3): 591–630. <https://doi.org/10.1016/j.cognition.2006.05.015>
- Blevins, J. (ed.). 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486357>
- Box, G. and D. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society Series B* 26 (2): 211–252. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- Brysbaert, M., M. Stevens, S. de Deyne, W. Voorspoels and G. Storms. 2014. Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta psychologica* 150: 80–84. <https://doi.org/10.1016/j.actpsy.2014.04.010>
- Burnham, K.P. and D.R. Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer-Verlag.
- Burnham, K.P., D.R. Anderson and K. Huyvaert. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65: 23–35. <https://doi.org/10.1007/s00265-010-1029-6>
- Burquest, D.A. and D.L. Payne. 1993. *Phonological analysis: A functional approach*. Dallas, TX: SIL International.
- Bybee, J. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Bybee, J. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Bybee, J.L. 1994. A view of phonology from a cognitive and functional perspective. *Cognitive Linguistics* 5: 285–306. <https://doi.org/10.1515/cogl.1994.5.4.285>

Cho, Y.-m. 1991. A typology of voicing assimilation. In Halpern, A. (ed.). *Proceedings of the 9th West Coast Conference on Formal Linguistics*. Stanford: CSLI Publications. pp. 141–155.

Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, G.N. 1990. The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (eds.). *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press. pp. 282–333. <https://doi.org/10.1017/CBO9780511627736.017>

Coetzee, A. 1969. *Dictionary: Afrikaans–English, English–Afrikaans. Pronouncing dictionary containing more than 39,000 entries*. Johannesburg: William Collins Sons & Company.

Coetzee, A. 2014. Grammatical change through lexical accumulation: voicing cooccurrence restrictions in Afrikaans. *Language* 90 (3): 693–721. <https://doi.org/10.1353/lan.2014.0066>

Colantoni, L. and J. Steele. 2007. Voicing-dependent cluster simplification asymmetries in Spanish and French. In P. Prieto, J. Mascaró, and M.-J. Solé (eds.). *Segmental and prosodic issues in Romance phonology*. Amsterdam: John Benjamins. pp.109–130. <https://doi.org/10.1075/cilt.282.09col>

Conradie, C. 2017. Taalverandering in Afrikaans. In W. Carstens and N. Bosman (eds.). *Kontemporêre Afrikaanse taalkunde*. Pretoria: Van Schaik.

Croft, W. 2000. *Explaining language change: An evolutionary approach*. London: Longman.

Daland, R. and J. Pierrehumbert. 2011. Learning diphone-based segmentation. *Cognitive Science* 35: 119–155. <https://doi.org/10.1111/j.1551-6709.2010.01160.x>

Diessel, H. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New ideas in psychology* 25: 108–127. <https://doi.org/10.1016/j.newideapsych.2007.02.002>

Diessel, H. 2012. Diachronic change and language acquisition. In A. Bergs, and L.J. Brinton (eds.). *English Historical linguistics*. Berlin: Mouton de Gruyter. pp. 1599–1612.

Donaldson, B.C. 1993. *A Grammar of Afrikaans*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110863154>

Dressler, W.U. 1989. Markedness and naturalness in phonology: The case of natural phonology. In O.M. Tomić (ed.). *Markedness in synchrony and diachrony*. Berlin: Mouton de Gruyter. pp. 111–120. <https://doi.org/10.1515/9783110863154>

Dziubalska-Kołączyk, K. 2014. Explaining phonotactics using NAD. *Language Sciences* 46: 6–17. <https://doi.org/10.1016/j.langsci.2014.06.003>

Dziubalska-Kołączyk, K., D. Pietrala and G. Aperliński. 2014. The NAD Phonotactic Calculator – an online tool to calculate cluster preference in English, Polish and other

languages. Available online: http://wa.amu.edu.pl/wa/files/The_NAD_Phonotactic_Calculator.pdf (Accessed 18 October 2017)

Eiselen, E.R. and M.J. Puttkammer. 2014. Developing text resources for ten South African languages. In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk and S. Piperidis (eds.). *Proc. of the 9th International Conference on Language Resources and Evaluation*. Paris: European Languages Resources Association. pp. 3698–3703.

Grijzenhout, J. and M. Krämer. 2000. Final devoicing and voicing assimilation in Dutch. In B. Stiebels and D. Wunderlich (eds.). *Lexicon in Focus*. Berlin: Mouton de Gruyter. pp. 55–83.

Hansson, G.O. 2001. The phonologization of production constraints: Evidence from consonant harmony. *Chicago Linguistic Society* 37 (1): 187–200.

Hansson, G.O. 2004. Long-distance voicing agreement: An evolutionary perspective. *Berkeley Linguistics Society* 30: 130–141. <https://doi.org/10.3765/bls.v30i1.946>

Heine, B. and T. Kuteva. 2007. *The Genesis of Grammar*. Oxford: Oxford University Press.

Hogg, R.M. and C.B. McCully. 1987. *Metrical Phonology: A Course Book*. Cambridge: Cambridge University Press.

Horobin, S. and J.J. Smith. 2002. *An Introduction to Middle English*. Oxford: Oxford University Press.

Johnson, J. and K. Omland. 2004. Model selection in ecology and evolution. *Trends in Ecology and Evolution* 19 (2): 101–108. <https://doi.org/10.1016/j.tree.2003.10.013>

Kotzé, E.F. 1984. Afrikaans in die Maleierbuurt: 'n Diachroniese perspektief. *Tydskrif vir Geesteswetenskappe* 24 (1): 41–73.

Kuijpers, C., W. van Donselaar and A. Cutler. 1996. Phonological variation: epenthesis and deletion of schwa in Dutch. *ICSLP 96: Proceedings, Fourth International Conference on Spoken Language Processing*: 149–152. <https://doi.org/10.1109/ICSLP.1996.607060>

Kuperman, V., H. Stadthagen-Gonzalez and M. Brysbaert. 2012. Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods* 44: 978–990. <https://doi.org/10.3758/s13428-012-0210-4>

Langacker, R.W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195331967.001.0001>

Linke, K. 2017. Coda. *Taalportal*. Available online: <http://www.taalportaal.org/taalportaal/topic/pid/topic-14008431983824103> (Accessed 4 October 2017)

MacNeilage, P.F. and B.L. Davis. 2000. On the origin of internal structure of word forms. *Science* 288: 527–531. <https://doi.org/10.1126/science.288.5465.527>

- Marecka, M. and K. Dziubalska-Kołaczyk. 2014. Evaluating models of phonotactic constraints on the basis of sC cluster acquisition data. *Language Sciences* 46: 37–47. <https://doi.org/10.1016/j.langsci.2014.06.002>
- Mesgarani, N., C. Cheung, K. Johnson and E.F. Chang. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343: 1006–1010. <https://doi.org/10.1126/science.1245994>
- Monaghan, P. 2014. Age of acquisition predicts rate of lexical evolution. *Cognition* 133: 530–534. <https://doi.org/10.1016/j.cognition.2014.08.007>
- Nakawaga, S. and I. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews* 82 (4): 591–605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x>
- Pagel, M., Q.D. Atkinson and A. Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449 (7163): 717–720. <https://doi.org/10.1038/nature06176>
- Ponelis, F.A. 1989. Ontwikkeling van klusters op sluitklanke in Afrikaans. *South African Journal of Linguistics* 7 (1): 1–5. <https://doi.org/10.1080/10118063.1989.9723783>
- Prince, A. and P. Smolensky. 2002. Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers Optimality Archive. Available online: <http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF> (Accessed 12 June 2013.)
- R Core Team. 2018. R: A Language and Environment for Statistical Computing, Vienna, Austria. Available online <https://www.R-project.org/> (Accessed 18 October 2017)
- Ritt, N. 2004. *Selfish sounds and linguistic evolution: A Darwinian approach to language change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486449>
- Roberge, P. 1993. The formation of Afrikaans. *Stellenbosch Papers in Linguistics* 27: 1–111.
- Roberge, P. 2002. Afrikaans: considering origins. In R. Mesthrie (ed.). *Language in South Africa*. Cambridge: Cambridge University Press. pp. 79–103. <https://doi.org/10.1017/CBO9780511486692.005>
- Roberts, I. 2007. *Diachronic syntax*. Oxford: Oxford University Press.
- Rosenbach, A. 2008. Language change as cultural evolution: Evolutionary approaches to language change. In R. Eckardt, G. Jäger and T. Veenestra (eds.). *Variation, Selection, Development: Probing the Evolutionary Model of Language Change*. Berlin: Mouton de Gruyter. pp. 23–74.
- Schouten, M. 1982. T-deletie in de stad Utrecht: schoolkinderen en grootouders. *Forum der Lettern* 23: 282–291.

- Seliger, H. and R. Vago (eds.). 1991. *First Language Attrition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620720>
- Szemerényi, O. 1996. *Introduction to Indo-European linguistics*. Oxford: Oxford University Press.
- Välismaa-Blum, R. 2005. *Cognitive phonology in construction grammar: Analytic tools for students of English*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110920598>
- Van de Velde, F. 2009. *De nominale constituent: structuur en geschiedenis*. Leuven: Universitaire Pers Leuven.
- Van Veen, T. 1964. *Utrecht tussen oost en west: Studies over het dialect van de provincie Utrecht*. Assen: Van Gorcum.
- Watermeyer, S. 1996. Afrikaans English. In V.A. de Klerk (ed.). *Focus on South Africa* Amsterdam: John Benjamins. pp. 99–148. <https://doi.org/10.1075/veaw.g15.08wat>
- Wedel, A. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23 (3): 247–274. <https://doi.org/10.1515/TLR.2006.010>
- West, B., K. Welch and A. Gałecki. 2015. *Linear mixed models*. Boca Raton, Florida: CRC Press.
- Wikimedia Commons. 2005. File:Pseudoceros dimidiatus.jpg. Author: Richard Ling. Available online: https://commons.wikimedia.org/wiki/File:Pseudoceros_dimidiatus.jpg (Accessed 10 March 2014)
- Wikimedia Commons. 2012. File:Rufous Piculet (male) - Sasia abnormis.jpg. Author: Michael Gillam. Available online: [https://commons.wikimedia.org/wiki/File:Rufous_Piculet_\(male\)_-_Sasia_abnormis.jpg](https://commons.wikimedia.org/wiki/File:Rufous_Piculet_(male)_-_Sasia_abnormis.jpg) (Accessed 10 March 2017)
- Wood, S. 2006a. *Generalized Additive Models: an introduction with R*. Boca Raton, Florida: CRC Press. <https://doi.org/10.1201/9781420010404>
- Wood, S. 2006b. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. *Biometrics* 62: 1025–1036. <https://doi.org/10.1111/j.1541-0420.2006.00574.x>
- Yang, C. 2000. Internal and external forces in language change. *Language Variation and Change* 12: 231–250. <https://doi.org/10.1017/S0954394500123014>

APPENDIX

Table A1. Ordinal scores for articulatory features of first and second consonant in the cluster (analogously as in Dziubalska-Kołodziejczyk 2014; Dziubalska-Kołodziejczyk, Pietrala and Aperliński 2014).

cluster	MoA1	MoA2	PoA1	PoA2	Voice1	Voice2
fs	3	3	2	4	0	0
ft	3	1	2	4	0	0
ks	1	3	7	4	0	0
kt	1	1	7	4	0	0
lf	6	3	4	2	1	0
lk	6	1	4	7	1	0
lm	6	5	4	1	1	1
lp	6	1	4	1	1	0
ls	6	3	4	4	1	0
lt	6	1	4	4	1	0
lx	6	3	4	7	1	0
mf	5	3	1	2	1	0
mp	5	1	1	1	1	0
ms	5	3	1	4	1	0
mt	5	1	1	4	1	0
ŋk	5	1	7	7	1	0
ns	5	3	4	4	1	0
ŋs	5	3	7	4	1	0
nt	5	1	4	4	1	0
ps	1	3	1	4	0	0
pt	1	1	1	4	0	0
rk	6	1	4	7	1	0
rm	6	5	4	1	1	1
rn	6	5	4	4	1	1
rp	6	1	4	1	1	0
rs	6	3	4	4	1	0
rt	6	1	4	4	1	0
rx	6	3	4	7	1	0
st	3	1	4	4	0	0
ts	1	3	4	4	0	0
ws	7	3	1	4	1	0
wt	7	1	1	4	1	0
xt	3	1	7	4	0	0

Table A2. Derived scores for the variables introduced in 2.2 (Study 1). Note that all numbers are rounded for two digits and represent raw (i.e. not Box-Cox transformed) data.

cluster	ΔMoA	ΔPoA	ΔVoice	frequency	length	AoA
fs	0	0.67	0	15464	6	8.02
ft	0.33	0.67	0	401	7	6.03
ks	0.33	1	0	2098	7	5.91
kt	0	1	0	588	7	7.51
lf	0.5	0.67	1	743.5	3.5	5.29
lk	0.83	1	1	1383	4	5.55
lm	0.17	1	0	551.5	4	6.01
lp	0.83	1	1	2490	4	4.87
ls	0.5	0	1	1253	7	7.13
lt	0.83	0	1	991	7	6.28
lx	0.5	1	1	5192	6.5	8.18
mf	0.33	0.33	1	512	7	9.78
mp	0.67	0	1	1178	4	5.49
ms	0.33	1	1	19260	4	6.78
mt	0.67	1	1	284	6.5	6.9
nk	0.67	0	1	618.5	5	5.72
ns	0.33	0	1	1175	7	6.61
ns	0.33	1	1	269	5	7.06
nt	0.67	0	1	664.5	7	5.69
ps	0.33	1	0	159	5	10.56
pt	0	1	0	215.5	5	10.65
rk	0.83	1	1	593	4	5.58
rm	0.17	1	0	3545.5	4	4.3
rn	0.17	0	0	1393	4	8.36
rp	0.83	1	1	2360	7	6.41
rs	0.5	0	1	208.5	5.5	6.91
rt	0.83	0	1	626	5	5.49
rx	0.5	1	1	380	4	5.76
st	0.33	0	0	774	5	5.94
ts	0.33	0	0	651.5	5	6.19
ws	0.67	1	1	1237	4	6.59
wt	1	1	1	946	7	7.9
xt	0.33	1	0	646.5	6	5.67

Table A3. Derived scores for the variables introduced in 3.2 (Study 2). Note that all numbers are rounded for two digits and represent raw (i.e. not Box-Cox transformed) data.

cluster	ΔMoA	ΔPoA	ΔVoice	DutProd	AfrProd
fs	0	0.67	0	2	1
ft	0.33	0.67	0	26	0
ks	0.33	1	0	13	18
kt	0	1	0	34	0
lf	0.5	0.67	1	4	12
lk	0.83	1	1	11	5
lm	0.17	1	0	4	0
lp	0.83	1	1	2	7
ls	0.5	0	1	9	6
lt	0.83	0	1	38	35
lx	0.5	1	1	2	5
mf	0.33	0.33	1	1	0
mp	0.67	0	1	4	3
ms	0.33	1	1	1	5
mt	0.67	1	1	8	0
ŋk	0.67	0	1	16	16
ns	0.33	0	1	38	49
ŋs	0.33	1	1	7	0
nt	0.67	0	1	182	131
ps	0.33	1	0	3	1
pt	0	1	0	2	0
rk	0.83	1	1	13	46
rm	0.17	1	0	8	0
rn	0.17	0	0	4	2
rp	0.83	1	1	7	8
rs	0.5	0	1	18	16
rt	0.83	0	1	73	39
rx	0.5	1	1	9	9
st	0.33	0	0	110	2
ts	0.33	0	0	46	28
ws	0.67	1	1	1	0
wt	1	1	1	3	0
xt	0.33	1	0	114	1