GRAMMATICAL CHANGE THROUGH LEXICAL ACCUMULATION: VOICING COOCCURRENCE RESTRICTIONS IN AFRIKAANS

Andries W. Coetzee

University of Michigan and *North-West University (South Africa)*

Two explanations are offered in the literature for the origin of lexical patterns of consonantal voicing cooccurrence: (i) speaker-oriented: a cooccurrence pattern may result from voicing assimilation under ease-of-articulation pressures, and (ii) listener-oriented: a cooccurrence pattern may result from systematic misperception by listeners. This article argues for a third possible origin of such patterns: (iii) lexical accumulation: a series of unrelated sound changes may conspire to create a lexical pattern of voicing cooccurrence. Once introduced into the lexicon of some language through any of these three routes, speakers can elevate such a pattern to a grammatical principle. A new voicing cooccurrence pattern in Afrikaans is presented as an example of a pattern that arose via this third route of lexical accumulation. Evidence is also presented that this pattern is being learned as a grammatical constraint by Afrikaans speakers.*

Keywords: Afrikaans, grammatical change, obstruent voicing, cooccurrence restrictions, lexicon

1. Introduction. The lexicons of many languages have restrictions on the cooccurrence of (possibly nonadjacent) consonants in some morphological or phonological domain. These patterns pertain to a variety of consonantal features, ranging from place of articulation to nasality, laterality, stridency, and various laryngeal features. For typological reviews, see Hansson 2001b and 2010, MacEachern 1999, Rose 2011, and Rose & Walker 2004. In addition to applying to a variety of features, these patterns are also observed in languages from many different language families, and with a wide geographical distribution.

Given the ubiquity of these patterns, it seems reasonable to assume that their existence is not accidental, but is based on general cognitive or physiological properties of language users. Ohala (1993:238), for instance, claims that sound patterns that are 'attested independently in substantially the same form in many unrelated languages … are most likely to arise from language universal factors, i.e. physiological and psychological factors common to all human speakers'. Proposals for such general cognitive and physiological origins for these cooccurrence patterns have indeed been offered in the literature, and include both speaker-oriented (Hansson 2007:82–83) and listener-oriented (Ohala 1993:251–57) accounts. The evidence in support of these explanations is sufficient that it can be accepted with reasonable certainty that most of these cooccurrence patterns are the result of one of these two mechanisms.

This does not mean, however, that all such patterns must have arisen through one of these two pathways. Blevins (2004:48) cautions that 'similar sound patterns may have arisen in different ways from different kinds of sound change'. It is hence possible that the lexicon of some language may have a consonantal cooccurrence pattern that did not enter the lexicon as a consequence of the general physiological and cognitive biases that are responsible for the majority of such patterns crosslinguistically.

* I want to express my appreciation to the many people with whom I have had the opportunity to discuss the ideas that went into this article. This list includes Pam Beddor, Ted Stern, Marc van Oostendorp, René Kager, Jaye Padgett, Junko Itô, and Armin Mester, among others. I also received valuable feedback from audiences at OCP Istanbul, the University of Michigan, the University of California, Santa Cruz, and the Meertens Institute. My experiences with the referees and the editors of *Language* have been some of the most positive that I have had in my career. They have truly contributed to making this a better article.

In this article, I present an example of a language, Afrikaans, in which a voicing cooccurrence pattern is robustly present in the lexicon, but for which there is clear evidence that this pattern did not enter the lexicon through the cognitive or physiological biases responsible for similar patterns in other languages. There was never a process in the history of Afrikaans that specifically eliminated words that would counter the cooccurrence pattern. What happened is rather that a set of independent sound changes, motivated by other general phonetic principles, applied. As a side effect of these changes having cooccurred in a single language, the lexicon has been rid of words that countered the cooccurrence pattern. The pattern has therefore been introduced accidentally into the Afrikaans lexicon, rather than as a direct result of the general cognitive or physiological forces usually responsible for such cooccurrence patterns.

A second independent, but related, question to be addressed about consonantal cooccurrence patterns in the lexicon of a language is whether a specific pattern will be learnable by language users as a grammatical principle. Not all patterns in the lexicon of some language necessarily become grammaticalized and influence language users' linguistic performance. In fact, it is typically only so-called 'natural' patterns, consistent with general cognitive and physiological biases, that are learnable and hence liable to grammaticalization (Becker et al. 2011, Becker et al. 2012, Hayes & White 2013, Peperkamp et al. 2006). Whether some consonantal cooccurrence pattern has entered the lexicon of a language through general cognitive or physiological pathways (the usual route) or through accidental lexical accumulation (as happened exceptionally in Afrikaans), once present in the lexicon of the language, the pattern is likely to be noticed by language users and to influence their linguistic behavior if it is consistent with the cognitive and physiological biases that underlie language. I also present evidence that, though the consonant cooccurrence pattern in Afrikaans did not enter the lexicon of this language via one of the usual routes, now that it is present in the lexicon, Afrikaans speakers have knowledge of the pattern, and it does influence their linguistic performance.

This introductory section first gives the relevant facts about Afrikaans, then reviews different explanations for the origin of consonant cooccurrence patterns, and finally reviews evidence with regard to the learnability of lexical patterns. The rest of the article consists of a detailed discussion of the Afrikaans cooccurrence pattern (§2), a discussion of the diachronic path that resulted in the Afrikaans pattern (§3), and evidence that the pattern has been elevated to a grammatical principle (§4). A final discussion concludes (§5).

1.1. AFRIKAANS, IN SHORT. Afrikaans, like its close ancestor Dutch, is a so-called 'final devoicing' language, meaning that voiced obstruents are not allowed in word-final position. This results in paradigms such as those in 1.[1] The nouns in 1a end underlyingly in voiceless obstruents and do not alternate between their unaffixed singulars and their affixed plurals. The nouns in 1b, however, end underlyingly in voiced obstruents, which surface intact when the noun roots are followed by a vowel-initial suffix in the plural.

(1) a. NONALTERNATING

SINGULAR		PLURAL		
kant	[kɑnt]	kante	[ˈkɑntə]	'side'
voet	[fut]	voete	[ˈfutə]	'foot'
kop	[kɔp]	koppe	[ˈkɔpə]	'head'
wip	[vəp]	wippe	[ˈvəpə]	'trap'

[1] All Afrikaans examples in this article are provided by the author, a native speaker of Afrikaans.

b. ALTERNATING	hand	[ɦɑnt]	hande	[ˈɦɑndə]	‘hand’
	hoed	[ɦut]	hoede	[ˈɦudə]	‘hat’
	rob	[rɔp]	robbe	[ˈrɔbə]	‘seal (animal)’
	web	[vɛp]	webbe	[ˈvɛbə]	‘web’

In this article, it is shown that the distribution of voiced obstruents in Afrikaans is, in fact, more restricted. In addition to being banned word-finally on the surface, voiced obstruents are also subject to restrictions on where they are allowed in a morpheme, that is, on the underlying level. Specifically, Afrikaans has a restriction that is very similar to that observed in Ngizim (Hansson 2001b:151–54, 317–34, Schuh 1978, 1997). In Ngizim morphemes, voiced obstruents cannot be preceded by voiceless obstruents. In morphemes with two consonants, the configurations in 2a are attested, but not those in 2b (with the exception of recent loans). Examples are from Schuh 1978 and 1997.[2]

(2) a. ATTESTED IN NGIZIM

No voiced obstruents (TT)	/pátà/	‘the bush’	(1978:256)
Voiced before voiceless (DT)	/bàkú/	‘roast’	(1997:3)
Voiced before voiced (DD)	/gàdú/	‘break (rope)’	(1997:4)
Sonorant before voiced (LD)	/mâbú/	‘calabash’	(1978:256)

b. UNATTESTED IN NGIZIM

Voiceless before voiced (TD)	*/kàdú/

A nearly identical pattern holds of Afrikaans morphemes that end in C_1VC_2, with one additional complication: the sequence /voiceless obstruent … voiced obstruent/ is allowed, as long as the voiceless obstruent in C_1 position does not have a voiced counterpart in Afrikaans—that is, SVD is allowed (see n. 2). That SVD sequences are allowed is not unexpected. As noted by Hansson (2008:125), this type of featural cooccurrence pattern ‘is typically also limited to segments which are contrastively specified for the feature in question’ (see also Mackenzie 2009). The Afrikaans pattern can be described in terms of where voiced obstruents appear in a morpheme, as in 3. Examples of the observed and nonobserved root types are given in 4. Although all of the examples here are of monosyllabic C_1VC_2 morphemes, this restriction also holds of longer morphemes that end in C_1VC_2 (see §2.4 for more on longer morphemes). The made-up examples in 4b show the kinds of morphemes that do not exist in Afrikaans. In these forms, C_1 is either /p/ or /t/, voiceless obstruents with voiced counterparts (Afrikaans has /b/ and /d/). Each of these examples has a morpheme-final voiced obstruent (C_2) that is preceded by a voiceless obstruent (C_1) that is contrastively specified for the feature [voice]. These examples crucially differ from the actual Afrikaans root /sɑ:d/ in 4a, where there is also a voiced obstruent in C_2 position, preceded by a voiceless obstruent in C_1. Since, however, Afrikaans lacks a voiced counterpart for /s/, C_1 in /sɑ:d/ is a voiceless obstruent that is NOT contrastively specified for [voice].

(3) OBSTRUENT VOICING GENERALIZATION: In an Afrikaans morpheme that ends in the sequence C_1VC_2, C_2 can be a voiced obstruent iff C_1 is voiced already (a voiced obstruent or sonorant) or cannot be specified as [+voice] (a voiceless obstruent that lacks a voiced counterpart).

[2] Throughout this article, the following notation is used: T: voiceless obstruent that has a voiced counterpart, S: voiceless obstruent that lacks a voiced counterpart, D: voiced obstruent, L: sonorant, V: vowel.

(4) a. Possible (actual) voiced obstruent-final morphemes (UF: underlying form)

	UF	SINGULAR		PLURAL		
Sonorant before voiced (LD)	/lɔb/	lob	[lɔp]	lobbe	[ˈlɔbə]	‘segment’
Voiced before voiced (DD)	/dɑːd/	daad	[dɑːt]	dade	[ˈdɑːdə]	‘deed’
Noncontrastive voiceless before voiced (SD)	/sɑːd/	saad	[sɑːt]	sade	[ˈsɑːdə]	‘seed’

b. Impossible voiced obstruent-final morphemes

Contrastive voiceless before voiced (TD)	/tɔb/	[tɔp]	[ˈtɔbə]
	/pɔd/	[pɔt]	[ˈpɔdə]

This article establishes the reality of the generalization in 3, and argues that this pattern is the result of lexical accumulation through a series of unrelated sound changes, rather than the direct result of perceptual or production-oriented forces.

1.2. Three routes to voicing cooccurrence patterns in the lexicon. Both a speaker-oriented articulatory and a listener-oriented perceptual route for the origin of consonant cooccurrence patterns have been proposed, and there is strong evidence that most examples of such patterns did come about through one of these two routes. In this section, I briefly review the existing proposals (see Hansson 2001b, 2004, 2007, 2008, 2010 for more details) and then propose a third, alternative route.

Speaker-oriented explanations. There are two ways in which the mechanisms of speech production can introduce consonant cooccurrence patterns into the lexicon: one relies on the local mechanics of speech production, and the other on global mechanisms of speech planning. In the articulatory phonology literature, consonant harmony is viewed as local spreading (Gafos 1998, 1999, 2002, Ní Chiosáin & Padgett 2001). Under this view, the consonantal features that participate in harmony do not skip over intervening segments but spread through them. Consonant harmony is then just a kind of assimilation. A related explanation could be offered from the perspective of Lindblom’s theory that speakers prefer to produce speech in a way that minimizes articulatory effort (Lindblom 1990). Producing two consonants with the same articulatory settings arguably requires less effort than changing articulatory settings. If, under articulatory-ease pressures, speakers regularly assimilate word-internal consonants in terms of some feature, a lexicon can result in which all relevant consonants in a word agree in terms of that feature.

Another speaker-oriented explanation relies not on the mechanics of speech production, but on the mechanisms of speech planning (Hansson 2001a:190–92). In most speech production models, a segment is activated several moments before it is actually produced. This anticipatory activation can interfere with the production of the present segment, resulting in anticipatory speech errors. That the speech production model relies on anticipatory activation is confirmed by the fact that most speech errors are anticipatory, rather than perseveratory—that is, speech errors such as *the Romney pampaign* (for *the Romney campaign*) are more common than ones like the *Romney camcaign* (Dell et al. 1997, Schwartz et al. 1994). As Hansson (2001a,b) shows, all (non-morphologically sensitive) consonant harmony processes are also anticipatory. For example, in the Chumashan language Ineseño, sibilants in a word agree in terms of the feature [palatal] with the right-most sibilant, resulting in alternations such as those in 5.

(5) Anticipatory sibilant harmony in Ineseño (from Applegate 1972:118–19)

/**s**-api-**tʃʰ**o-u**s**/	[**s**api**tsʰ**olu**s**]	‘he has a stroke of good luck’
/**s**-api-**tʃʰ**o-u**s**-wa**ʃ**/	[**ʃ**api**tʃʰ**olu**ʃ**wa**ʃ**]	‘he had a stroke of good luck’

Additionally, many of these anticipatory consonant harmony systems operate over long distances (Rose & Walker 2004), making the possibility of a localized assimilation origin for such processes unlikely. Hansson (2001a,b) argues that, if the speakers of some language regularly produce this type of speech error, it can over time result in a lexicon in which all relevant consonants in a word agree in terms of some feature.

LISTENER-ORIENTED EXPLANATIONS. Due to the dynamic nature of speech, a speech sound typically has an acoustic footprint that is wider than the discrete linear locus of the segment in the speech stream. For instance, voiced obstruents cause lower F0 and F1 in neighboring vowels compared to vowels adjacent to voiceless obstruents—see Bradshaw 1999:Ch. 6 for the F0 effect and Moreton 2008 for the F1 effect. When a listener is presented with a token consisting of the sequence [obstruent + vowel + obstruent], and with a depressed F0 and F1 in the vowel, the listener has to assign the F0/F1 effects to either the first or the second obstruent, or to both. On the one hand, if a speaker utters a token such as /bug/, with two phonologically voiced obstruents, but a listener assigns the F0/F1 effects to only the first or second consonant, the listener may settle on a representation with only one voiced obstruent, /buk/ or /pug/. If listeners in some speech community often make this kind of perceptual error, it can result over time in a lexicon that lacks words with multiple voiced obstruents. Gallagher (2010) presents evidence for such a perceptual origin for the lack of Quechua words with multiple aspirated segments; see the examples in 6.

(6) Aspirated plosives in Quechua roots (examples from Parker 1997:2)

a. Observed	[**pʰ**atay]	‘to explode’	[**tʰ**anta]	‘old’
	[**kʰ**anka]	‘slimy’	[**tʃʰ**atʃu]	‘ragged’
b. Not observed	[**pʰ**a**tʰ**ay], [**tʰ**an**tʰ**a], [**kʰ**an**kʰ**a], [**tʃʰ**a**tʃʰ**u]			

On the other hand, if a speaker utters a form with only one voiced obstruent, /buk/, and a listener erroneously ascribes the depressed F0 and F1 to both obstruents, the listener may settle on /bug/, and over time this may result in a lexicon where all obstruents in a word agree in voicing. Hansson (2004:132) offers this explanation for the restriction in the Chadic language Kera that requires all obstruents in a word to have the same specification for voicing, and that causes voicing alternation in affixes, as shown in 7.[3] This route to the origin of consonant cooccurrence patterns is most commonly associated with Ohala’s so-called ‘innocent misperception’ theory of sound change (Ohala 1993); see also Blevins 2004, 2006 for more recent similar ideas.

(7) Voicing alternations in Kera affixes (examples from Pearce 2005:68)

	SINGULAR	PLURAL ([k-]/[g-] prefix)	
a. Root with [−voice] obs	[**k**úmná]	[**kɨk**ámn**ɨ**]	‘chief’
	[**t**āa**ɖ**á]	[**kɨt**āatáw]	‘big pot’
b. Root with [+voice] obs	[**bɨ**rwá]	[**gɨb**ɨrwáŋ]	‘white’
	[**d**àarà]	[**gə̀d**àarà]	‘friend’

[3] See Pearce 2009 for a different view. She shows that Kera listeners rely on F0 rather than VOT as the primary cue for the contrast between historically voiced and voiceless plosives. VOT is used only as a secondary cue, when the F0 cue is ambiguous between the two categories. She therefore concludes that Kera no longer contrasts its plosives in terms of voicing, but that the contrast has been transferred onto the F0 of the following vowel. The vowels in the plural prefixes in 7a and 7b would then differ from each other in F0, while the velar plosives in these prefixes would not differ appreciably in terms of VOT.

LEXICAL ACCUMULATION. The literature on consonantal cooccurrence contains many examples of languages where it can be shown through historical reconstruction that the cooccurrence patterns were introduced into these languages via the perceptual and/or production routes discussed above. However, Blevins cautions against overgeneralization—even though a frequent sound pattern can be shown to have the same origin in many languages, it does not follow that it must have the same origin in every language in which it is observed: 'similar sound patterns may have arisen in different ways from different kinds of sound change' (Blevins 2004:48; see also Blevins 2006:123).

Blevins (2004, 2006) discusses several examples of languages with virtually the same synchronic patterns, but where it can be shown that these patterns were introduced into the languages via different routes. In Madurese, for instance, the first of two consonants in a cluster often assimilated historically to the second consonant, giving rise to geminates in the modern language, and to correspondences like the following between Proto-Austronesian and Madurese: **bakbak* > *babbaʔ* 'tree bark' and **paqpaq* > *pappa* 'coconut frond' (Blust 1995:125). In Hausa, by contrast, gemination arose as a consequence of vowel syncope between identical consonants, giving correspondences such as the following between Proto-Hausa and Hausa: **zoobàbaa* > *zôbbaa* 'rings' and **saasàsaa* > *sâssaa* 'section' (Newman 1990:47).

In this article, I show that there are routes other than the usual perception/production ones through which consonant cooccurrence patterns can enter the lexicon of a language. Under the perception/production accounts, the assumption is that words that counter the cooccurrence pattern are inherently marked in some way (articulatorily or perceptually), and that there was a historical process that specifically targeted these words due to their markedness. The alternative proposed here is that a cooccurrence pattern can be introduced into the lexicon of a language without being specifically targeted by a single historical process. It can be introduced as an accidental side effect of a collection of independent sound changes that happen to cooccur in a single language. Each of these processes can affect a broader group of words, including some of the words that counter the cooccurrence pattern. These changes accumulate over time in the lexicon, and eventually result in a lexicon lacking any words that counter the cooccurrence pattern. These different routes are represented schematically in Figure 1. The diagram on the left has one process that targets all and only the words that counter the cooccurrence pattern. The diagram on the right has two processes. Each of these targets a group of words, some of which counter the cooccurrence pattern and some of which do not. Together, however, these two processes still eliminate from the lexicon all words that counter the cooccurrence pattern. The end result of these different routes is the same: a lexicon lacking words that counter the cooccurrence pattern.

The usual route: One targeted process The alternative: Multiple, unrelated processes

FIGURE 1. Two different routes to a lexicon with a consonantal cooccurrence pattern. w'_i represents a word that counters the cooccurrence pattern, and w_i a word that agrees with the pattern. c_i stands for a historical sound change.

I show below that the voicing cooccurrence pattern of Afrikaans, stated in 3 above, entered the Afrikaans lexicon via this third possible route, the gradual accumulation of the pattern in the lexicon through the cooccurrence of a set of unrelated sound changes. Three independent sound changes that affected voiced obstruents more generally happen to have cooccurred in the history of Afrikaans, and the cooccurrence of these three processes resulted in the elimination from the Afrikaans lexicon of words that counter the generalization in 3.

1.3. GRAMMATICALIZATION OF LEXICAL PATTERNS. The previous section focused on how consonantal cooccurrence patterns may come to be established in the lexicon of some language. However, not every pattern present in the lexicon of a language will necessarily be linguistically relevant. Once a lexical pattern has been established, the next question is whether speakers will treat the pattern as accidental, or whether they will elevate it to a grammatical principle.

Hayes and White (2013), for instance, show that there are robust phonotactic patterns in the English lexicon that nonetheless do not influence how English speakers perform linguistically. They show that English lacks words with high round vowels followed by [h] (no [uh] or [ʊh] sequences) and also words that start with stop-nasal clusters (no sequences like [#kn…]). The former pattern (absence of [uh] and [ʊh]) is not crosslinguistically common and is most likely accidental. The second pattern (no [#kn…]) is robustly attested crosslinguistically and has a clear phonetic basis. Hayes and White show that, though both of these patterns receive equal evidence from the English lexicon, only the second pattern influences English speakers' word-likeness ratings. English speakers prefer nonce words with FRICATIVE-NASAL initial clusters over those with STOP-NASAL clusters (they prefer [**sn**ɛk] over [**pn**ɛk]). But they do not prefer nonce words with NONROUND high vowels followed by [h] over nonce words with ROUND high vowels followed by [h] (they do not prefer [l**ɪh**æləm] over [l**uh**æləm]). See also Becker et al. 2012 for additional, similar evidence from English, Becker et al. 2011 for evidence from Turkish, and Peperkamp et al. 2006 for evidence from French. Hayes and White use this as evidence that not all lexically robust patterns are equally learnable by speakers. Patterns that are grammatically natural in some sense are more easily learnable than grammatically unnatural or accidental patterns.[4]

I leave aside here the question of what determines the naturalness of a pattern (whether grammatical naturalness is determined by innate principles of UNIVERSAL GRAMMAR or by phonetic naturalness). Whatever determines whether a pattern is learnable, the consonantal cooccurrence patterns reviewed in §1.2 clearly are learnable. These patterns are robustly attested crosslinguistically, and even motivate active phonological alternations in some languages (see the Kera voicing alternations in 7 above). In any language where such a cooccurrence pattern is present in the lexicon, irrespective of how the pattern came to be established in the lexicon, it can therefore be expected that the pattern will influence how speakers perform linguistically. In §4, I present evidence that Afrikaans speakers are influenced by the consonantal cooccurrence pattern in 3 when evaluating nonce words. This will serve as evidence that any lexical pattern that is consistent with general

[4] The claim is not that unnatural patterns are unlearnable, but rather that such patterns are less easily learnable and less stable (more prone to be lost over time). For evidence that unnatural patterns are learnable, see, for instance, Hayes 1999:269–70, Hyman 2001, and Ohala 2005:35. As an example of an unnatural pattern that is being learned as a grammatical rule, see Coetzee & Pretorius 2010 for evidence that Tswana speakers do learn the phonetically unnatural process of postnasal devoicing and productively apply it to new words (also Coetzee et al. 2007, Solé & Hyman 2010).

grammatical principles can be grammaticalized, even if the pattern entered the lexicon via the accidental lexical accumulation route described above.

2. The distribution of obstruent voicing in afrikaans. This section documents the Afrikaans cooccurrence pattern by investigating the Afrikaans lexicon in detail. The Afrikaans consonant inventory is given in Table 1. Sounds in parentheses either appear only in recent loans or have questionable phonemic status. The obstruents are enclosed in a box—in Afrikaans [ɦ] functions as a glide rather than a fricative (Wissing 1982: 45), and is therefore not included in the obstruent box.

p	b	t	d				k	(g)	(ʔ)
f	v	s	(z)	(ʃ)	(ʒ)		x		
		(ts)		(tʃ)					
	m		n					ŋ	
			l						
			r						
	w				j				ɦ

Table 1. Afrikaans consonant inventory.

The focus in this article is on the three pairs of obstruents that contrast in voicing, that is, /p, b/, /t, d/, and /f, v/. Specifically, the claim is that a morpheme that ends in C_1VC_2 can have one of /b, d, v/ in C_2 position only if C_1 is not a potential sponsor of a voicing feature. If C_1 is a sonorant or a voiced obstruent, it is already voiced and cannot sponsor an additional voicing feature. If C_1 is a voiceless obstruent that lacks a voiced counterpart (/ʔ, k, x, s/), it also cannot sponsor a voicing feature. Morphemes with /b, d, v/ in C_2 position are hence possible when C_1 is a voiced obstruent, a voiceless obstruent without a voiced counterpart, or a sonorant (examples were given in 4a), but not when C_1 is one of /p, t, f/, voiceless obstruents that have voiced counterparts in Afrikaans (see §2.2 for one exception). Examples of the nonobserved TVD morpheme types are given in 8.

(8) Nonobserved Afrikaans TVD morphemes

/pVb/	/tVb/	/fVb/
/pVd/	/tVd/	/fVd/
/pVv/	/tVv/	/fVv/

2.1. Lexical evidence about the distribution of voicing in C_1VC_2 morphemes. This section starts by documenting the distribution of consonantal voicing in monosyllabic C_1VC_2 morphemes of Afrikaans. Section §2.4 below considers longer morphemes that end in C_1VC_2. With one marginal exception (on which more below), morphemes like those in 8 are absent from the Afrikaans lexicon. To confirm this, a list was compiled of all C_1VC_2 monomorphemic words in Bosman et al. 1984, a widely used Afrikaans-English bilingual dictionary. Homophones were removed from the list, so that there was only one representative for each underlying C_1VC_2 sequence. Obvious loans were also removed. This resulted in a corpus of 888 C_1VC_2 words. Table 2 shows how these words are distributed with regard to the identity of C_1 and C_2. Each cell contains the observed number of words of each type (O), the number expected under a chance combination of the consonants in C_1 and C_2 positions (E), and the ratio between these two numbers (O/E) (see Frisch et al. 2004 for how expected frequencies are calculated, and see also Coetzee & Pater 2008). A cell with an O/E ratio above 1 indicates that the word type in that cell occurs more often than expected by chance, and an O/E ratio below 1 indicates a word type that appears less often than expected. A χ^2-test con-

firms that the actual distribution of the morphemes is different from what would be expected under chance ($\chi^2(9) = 3.32$, $p < 0.001$). Inspection of the O/E ratios shows that TVD morphemes—that is, morphemes that counter the generalization in 3—deviate from their expected distribution more than any other morpheme type by a large margin. The O/E ratio for these morphemes is 0.06, indicating severe underrepresentation. The next most underrepresented morpheme type is DVT, with an O/E ratio of 0.69. Examples of each of these kinds of morphemes are given in Appendix A.

C_1 \ C_2	L			S			T			D		
	O	E	**O/E**	O	E	**O/E**	O	E	**O/E**	O	E	**O/E**
L	109	126.72	**0.86**	110	105.17	**1.05**	68	67.26	**1.01**	43	30.84	**1.39**
S	73	70.66	**1.03**	54	58.64	**0.92**	42	37.50	**1.12**	15	17.20	**0.87**
T	79	67.20	**1.18**	52	55.77	**0.93**	43	35.67	**1.21**	1	16.36	**0.06**
D	80	76.42	**1.05**	67	63.42	**1.06**	28	40.56	**0.69**	24	18.60	**1.29**

TABLE 2. Distribution of C_1VC_2 monomorphemes in Afrikaans. (L: sonorant /m, n, ŋ, l, r, j, w, ɦ/; S: noncontrasting voiceless obstruent /s, x, k, ʔ/; T: contrasting voiceless obstruent /p, t, f/; D: contrasting voiced obstruent /b, d, v/.)

The distribution of the C_1VC_2 morphemes is represented graphically in a mosaic plot in Figure 2. This plot shows that TVD morphemes are severely underrepresented in the lexicon of Afrikaans—and as is shown in §2.2, the single observed TVD morpheme has a marginal status in the Afrikaans lexicon, and may in fact not be an active part of the mental lexicon of many Afrikaans speakers.

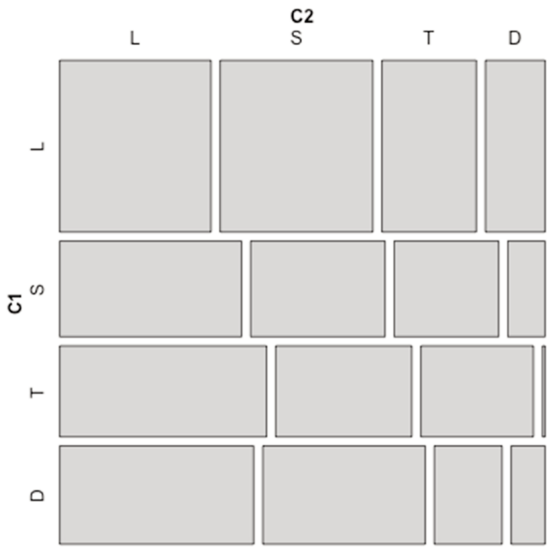

FIGURE 2. Mosaic plot of the distribution of C_1VC_2 morphemes in Afrikaans.

2.2. EXCEPTIONS. As just noted, there is one exception in Afrikaans to the generalization stated in 3. In addition to the one actual exception there is also one apparent exception. These two examples are given in 9, together with their morphologically related forms that seem to give evidence for the underlying voicing of their respective C_2s.

(9) Exceptions

BARE FORM			AFFIXED FORMS		
tyd	[teɪt]	‘time’	tydelik	[ˈteɪdələk]	‘temporary’
			tydig	[ˈteɪdəx]	‘timely’
voed	[fut]	‘feed (verb)’	voedende	[ˈfudəndə]	‘feeding (pres. participle)’
			voeding	[ˈfudəŋ]	‘feed (verbal noun)’

tyd: ONLY AN APPARENT EXCEPTION. Although the word *tyd* ‘time’ seems to be an exception to the voicing generalization, it is shown here that its underlying form is really /teɪt/, so that it is not a TVD morpheme. Afrikaans nouns that end in an alveolar plosive fall into one of two categories. Some preserve the alveolar plosive in the plural, while in others, addition of the plural suffix /-ə/ triggers two phonological alternations in the root: diphthongization of the root vowel, and replacement of the root-final alveolar stop with the glide [j]. Membership in these two classes cannot be predicted and has to be lexically specified. Several examples of each of these two classes of nouns are given in 10. The nouns in 10a retain the alveolar plosive in their plurals. For these nouns, the underlying voicing of the alveolar plosive can therefore be determined by consulting their plurals. In the nouns in 10b, however, the alveolar plosive is replaced by the glide [j] in the plural. The plurals of these nouns therefore do not give evidence for the underlying voicing of their final plosives.

(10) Alveolar plosive-final nouns

a. Alveolar plosive preserved in the plural

UF	SINGULAR		PLURAL		
/nuəd/	nood	[nuət]	node	[ˈnuədə]	‘need/emergency’
/ɦud/	hoed	[ɦut]	hoede	[ˈɦudə]	‘hat’
/nuət/	noot	[nuət]	note	[ˈnuətə]	‘note’
/fut/	voet	[fut]	voete	[ˈfutə]	‘foot’

b. Alveolar plosive replaced by [j] in the plural

UF	SINGULAR		PLURAL		
/pɑt/	pad	[pɑt]	paaie	[ˈpɑːɪjə]	‘road’
/skəlpɑt/	skilpad	[ˈskəlpɑt]	skilpaaie	[ˈskəlpɑːɪjə]	‘tortoise’
/krœɪt/	kruid	[krœɪt]	kruie	[ˈkrœɪjə]	‘herb’
/teɪt/	tyd	[teɪt]	tye	[ˈteɪjə]	‘time’

With the exception of the root for *tyd*, none of the roots belonging to the class represented in 10b appears in any form followed by a vowel-initial suffix other than the plural. For these roots, there is hence no evidence that the underlying plosive may be voiced. However, the situation for *tyd* is different. As shown in 9, what appears to be this root combines with two vowel-initial suffixes in addition to the plural /-ə/. With these other suffixes, the plosive is retained and surfaces as [d]. This appears to be evidence for the underlying voicedness of this plosive, and would imply that the root of this noun is really /teɪd/, and that it represents an exception to the generalization in 3.

It is not clear, however, that the words *tydig* and *tydelik* are actually derived from the root for the noun *tyd*. Of the two suffixes involved, /-ələk/ is very productive and combines with many different noun roots. In 11, examples are given of adjectives formed with this suffix. Addition of /-ələk/ changes the noun into an adjective with the meaning ‘like Noun’, as seen in all of these adjectives with the exception of *tydelik*. The meaning of *tydelik* is not merely ‘like time’ but ‘temporary’, and, in fact, it cannot be deduced from the combination of the meaning of the noun *tyd* and the suffix /-ələk/. Although *tydelik* did historically derive from *tyd*, it has gained partial independence from the root, and it may now be stored in fully derived form in the lexicon.

(11) /-ələk/ adjectives[5]

NOUN			ADJECTIVE		
vriend	[frint]	‘friend’	vriendelik	[ˈfrindələk]	‘friendly’
vader	[ˈfɑːdər]	‘father’	vaderlik	[ˈfɑːdərlək]	‘fatherly’
suster	[ˈsœstər]	‘sister’	susterlik	[ˈsœstərlək]	‘sisterly’
dier	[diːr]	‘animal’	dierlik	[ˈdiːrlək]	‘animal-like’
mens	[mɛns]	‘human’	menslik	[ˈmɛnslək]	‘humanly’

The other suffix with which what seems to be the noun *tyd* combines (/-əx/) is not productive, and its meaning is also less transparent. Examples of some of the adjectives formed with this suffix are given in 12. The meanings of the adjectives, though related to the nouns from which they are derived, are not merely ‘like Noun’. Given that /-əx/ is not productive and that the meanings of /-əx/ adjectives are not compositionally transparent, these adjectives are probably stored in fully derived form in the mental lexicon.

(12) /-əx/ adjectives

NOUN			ADJECTIVE		
skat	[skɑt]	‘treasure’	skattig	[ˈskɑtəx]	‘cute’
hand	[ɦɑnt]	‘hand’	handig	[ˈɦɑndəx]	‘useful/appropriate’
kop	[kɔp]	‘head’	koppig	[ˈkɔpəx]	‘stubborn’
haar	[ɦɑːr]	‘hair’	harig	[ˈɦɑːrəx]	‘full of hair/hairy’
vuur	[fyːr]	‘fire’	vurig	[ˈfyːrəx]	‘intense/serious’

The derivational relationship between *tyd* and *tydig*/*tydelik* is tenuous at best. If *tyd* is no longer derivationally related to *tydig* and *tydelik*, there is no evidence for the voicedness of the final obstruent in *tyd*, so *tyd* would end underlyingly in /t/, like the other noun roots in 10b. Although I therefore treat *tyd* as not representing an exception to the voicing cooccurrence restriction stated in 3, the main thrust of the argument made here does not depend on this assumption. Even if *tyd* were counted as an exception, there would still be only two such exceptions in the Afrikaans lexicon. TVD morphemes would still be highly underrepresented with an O/E ratio of 0.12 (rather than 0.06 when *tyd* is not counted as an exception). TVD roots would still be significantly more underrepresented than the next most underrepresented root type (DVT roots with an O/E ratio of 0.69).

voed: THE ONLY ACTUAL EXCEPTION. The only actual exception to the generalization on voicing cooccurrence in Afrikaans morphemes is the verb *voed* /fud/ → [fut] ‘feed’. This verb also combines with two vowel-initial suffixes, and in these suffixed forms the final obstruent of the root surfaces as voiced [d]. The first suffix, /-əŋ/, though not extremely productive, combines with several verbs and always creates a noun with the meaning ‘the act of performing Verb’. Some examples are given in 13a. The other suffix with which /fud/ combines is highly productive, and the meaning of the resulting word is always completely transparent: attaching /-əndə/ to a verb creates a present participle form. Some examples are given in 13b. There is no evidence that *voed* [fut] is not directly derivationally related to *voeding* [ˈfudəŋ] and *voedende* [ˈfudəndə], so this word is an exception to the generalization about the distribution of voiced obstruents in Afrikaans morphemes.

[5] With roots that do not end in /t/ and /d/, this suffix surfaces as /-ləк/.

(13) a. /-əŋ/ nouns

VERB			NOUN: 'act of performing Verb'	
buig	[bœɪx]	'bend'	buiging	[ˈbœɪxəŋ]
staak	[stɑːk]	'strike (labor)'	staking	[ˈstɑːkəŋ]
opwek	[ˈɔpvæk]	'wake someone'	opwekking	[ˈɔpvækəŋ]
styg	[steɪx]	'rise'	styging	[ˈsteɪxəŋ]
daal	[dɑːl]	'lower (intrans.)'	daling	[ˈdɑːləŋ]

b. Present participles formed with /-əndə/

VERB			PRESENT PARTICIPLE	
sit	[sət]	'sit'	sittende	[ˈsətəndə]
staan	[stɑːn]	'stand'	stanende	[ˈstɑːnəndə]
loop	[luəp]	'walk'	lopende	[ˈluəpəndə]
praat	[prɑːt]	'talk'	pratende	[ˈprɑːtəndə]
brei	[breɪ]	'knit'	breiende	[ˈbreɪjəndə]

Even though *voed* /fud/ is a counterexample to the generalization about the distribution of voiced obstruents in Afrikaans, it should be noted that this word is marginal in the Afrikaans lexicon. It is a very formal word that is used primarily in medical discourse. The more common way to express the meaning corresponding to the English verb 'feed' is with the phrase *gee kos* [xiə kɔs] 'give food'.

2.3. ON THE SPECIAL STATUS OF /f/ AND /v/. There are five C_1VC_2 morphemes in Afrikaans that start with a contrasting voiceless obstruent (a potential voicing sponsor) and end in a labiodental fricative; all five are given in 14. When these words appear without a suffix, the final fricative is realized as [f]. When followed by (some) vowel-initial suffixes, the fricative surfaces as [v]. These words therefore appear to be exceptions to the voicing cooccurrence generalization. However, this section shows that these roots do end underlyingly in /f/, and that this /f/ is realized as [v] before vowel-initial suffixes due to a process of intersonorant voicing.

(14) /f/-final nouns

UF	BARE FORM			BEFORE VOWEL-INITIAL SUFFIX		
/tiəf/	teef	[tiəf]	'bitch'	tewe	[ˈtiəvə]	'bitches'
/puf/	poef	[puf]	'beanbag chair'	poewe	[ˈpuvə]	'beanbag chairs'
/pɔf/	pof	[pɔf]	'puff (noun)'	powwe	[ˈpɔvə]	'puffs (noun)'
/feɪf/	vyf	[feɪf]	'five'	vywe	[ˈfeɪvə]	'fives'
/fœɪf/	fuif	[fœɪf]	'party (verb)'	fuiwery	[ˈfœɪvəreɪ]	'act of partying'

The distribution of [f] and [v] in Afrikaans interacts in complex ways with morphology. Both can appear word-initially and morpheme-internally, as shown in 15a and 15b. However, when followed by a vowel-initial suffix, the contrast between /f/ and /v/ is neutralized. Before the diminutive suffix /-i/, only [f] is observed (15c). Before all other vowel-initial suffixes, only [v] is observed. In 15d, for instance, the nouns from 15c are repeated, but this time in the plural. The fricative is realized as [f] in the diminutive and as [v] in the plural. The plural pattern is the more general pattern observed with other vowel-initial suffixes, as shown by the examples in 16.

(15) The distribution of [f] and [v]

a. Word-initial	[v]:	water	[ˈvɑːtər]	'water'
		woud	[vœʊt]	'forest'
	[f]:	vark	[fɑrk]	'pig'
		vuis	[fœɪs]	'fist'

b. Morpheme-internal	[v]:	oewer	[ˈuvər]	‘river bank’
		lower	[ˈluəvər]	‘greenery/green leaves’
	[f]:	offer	[ˈɔfər]	‘sacrifice’
		skoffel	[ˈskɔfəl]	‘hoe’
c. Before DIM /-i/	[f]:	grafie	[ˈxrɑːfi]	‘small spade’
		skroefie	[ˈskrufi]	‘small screw’
d. Before PL /-ə/	[v]:	grawe	[ˈxrɑːvə]	‘spades’
		skroewe	[ˈskruvə]	‘screws’

(16) More [v] before vowel-initial suffixes

a. /-ereɪ/: Changes verb to noun with meaning ‘act of performing Verb’

VERB			NOUN	
fuif	[fœɪf]	‘revel’	fuiwery	[ˈfœɪvəreɪ]
skryf	[skreɪf]	‘write’	skrywery	[ˈskreɪvəreɪ]
draf	[drɑf]	‘jog’	drawwery	[ˈdrɑvəreɪ]

b. /-ərəx/: Changes nouns or adjectives to adjectives meaning ‘a little like Noun’

NOUN/ADJECTIVE			ADJECTIVE	
teef	[tiəf]	‘bitch’	tewerig	[ˈtiəvərəx]
pof	[pɔf]	‘puff’	powwerig	[ˈpɔvərəx]
styf	[steɪf]	‘stiff’	stywerig	[ˈsteɪvərəx]

Given that, except before the diminutive suffix, only [v] appears before vowel-initial suffixes, I propose that there is a process of intersonorant voicing that applies to /f/.[6] The context in which the process applies is limited to morphologically derived environments—across morpheme boundaries—making this an example of a so-called ‘non-derived environment blocking’ effect (Kiparsky 1973, Łubowicz 2002, McCarthy 2003, etc.). What complicates the application of this process in Afrikaans is that it applies in all morphologically derived environments EXCEPT before the diminutive suffix /-i/.

The exceptional status of the diminutive with regard to this voicing rule indicates that diminutive formation is a late morphological process that applies only after the phonological rule of intersonorant /f/-voicing has already applied so that /-i/-suffixation cannot feed this rule. The late addition of the diminutive is confirmed by the interaction of this suffix with other phonological rules. For instance, Afrikaans does not allow tautosyllabic obstruent clusters. In nouns that end underlyingly in such clusters, the second member of the cluster is deleted in the singular. In the plural, however, the addition of the /-ə/ rescues the second member of the cluster. This gives rise to alternating singular ~ plural paradigms such as: *kas* /kɑst/ → [kɑs] vs. *kaste* /kɑst-ə/ → [ˈkɑs.tə] ‘cupboard’. The diminutive suffix /-i/, however, does not rescue the cluster. The diminutive form of *kas* is [ˈkɑ.si], not *[ˈkɑs.ti]. The diminutive suffix therefore attaches only after the intersonorant /f/-voicing and the cluster-simplification rules have applied.

Given the existence of intersonorant /f/-voicing in Afrikaans, there is no definitive evidence for the underlying voicing of the morpheme-final labiodental fricatives of the words in 14. Since the surface realization of the voicing of these fricatives is predictable, they could be either voiced or voiceless underlyingly. While acknowledging the indeterminacy of the available data, I follow other researchers (Grieshaber 1987, Lubbe & Zonneveld 1996) in assuming that these words end underlyingly in voiceless

[6] See den Besten 1999 for some speculation about the historical origin of this process in Afrikaans.

/f/, which surfaces as [v] in the plural as a result of intersonorant voicing. These words then do not represent exceptions to the voicing cooccurrence restriction stated in 3.

2.4. ON MORPHEMES OTHER THAN C_1VC_2. The previous sections discussed the distribution of obstruent voicing in simple C_1VC_2 morphemes. This section explores how obstruent voicing is distributed in roots with different structures. The first root type to be considered is roots of the form $C_1C_2VC_3$ and $C_1VC_2C_3$. A question to be answered is whether roots exist where C_3 is a voiced obstruent (/b, d, v/) while an earlier consonant (C_1 or C_2) is a potential sponsor for a voicing feature (/p, t, f/). Afrikaans does not have many roots of the form $C_1C_2VC_3$ or $C_1VC_2C_3$. Additionally, there are only a few clusters in Afrikaans in which a voiced obstruent is tolerated. In morpheme-initial position, the only such clusters are /br, bl, dr/. With regard to $C_1C_2VC_3$ roots, the relevant root types are hence /prVb, prVd, trVb, trVd, … / (roots that begin with a cluster that could sponsor a voicing feature and end in a voiced obstruent). To the best of my knowledge only one such root exists, given in 17a.

(17) Roots with onset or coda clusters

a. INITIAL CLUSTERS

tred	[trɛt]	‘step/tread’	trede	[ˈtriədə] ‘steps/treads’

b. FINAL CLUSTERS

pand	[pant]	‘surety’	pande	[ˈpandə] ‘sureties’
perd	[pært]	‘horse’	perde	[ˈpærdə] ‘horses’
pond	[pɔnt]	‘pound’	ponde	[ˈpɔndə] ‘pounds’
tand	[tant]	‘tooth’	tande	[ˈtandə] ‘teeth’
veld	[fælt]	‘field’	velde	[ˈfældə] ‘fields’
vind	[fənt]	‘find (verb)’	vindende	[ˈfəndəndə] ‘finding (pres. part.)’

In morpheme-final position, the only clusters with a voiced obstruent are /rd, nd, ld/. Here the relevant root types are hence /fVld, pVrd, tVnd, … / (roots that end in a voiced obstruent and have a consonant earlier in the root that could sponsor a voicing feature). An exhaustive list of such words from Bosman et al. 1984 is given in 17b. The words in 17 all have a morpheme-final voiced obstruent and a potential voicing sponsor earlier in the root. These roots seem to represent exceptions to the generalization on the distribution of voicing stated in 3. In these roots, however, the potential voicing sponsor is separated from the morpheme-final voiced obstruent by an intervening consonantal voicing feature (associated with the sonorants /n, l, r/). The generalization in 3 should therefore be stated more accurately as in 18. The adjacency requirement applies on the consonantal tier, ignoring intervening vowels, similar to the adjacency requirement that applies to consonantal OCP constraints (McCarthy 1986).

(18) UPDATED OBSTRUENT VOICING GENERALIZATION: An Afrikaans morpheme can end in a voiced obstruent iff any directly preceding segment on the consonantal tier is voiced already (a voiced obstruent or sonorant) or cannot be specified as [+voice] (a voiceless obstruent that lacks a voiced counterpart).

A second root type to consider is longer roots that end in the sequence C_1VC_2. The vast majority of Afrikaans morphemes are monosyllabic, so there are not many longer roots that end in C_1VC_2, and the longer roots that do exist are mostly loans. Some examples are given in 19. An exhaustive investigation of all such roots in Bosman et al. 1984 confirms that none of the final C_1VC_2 sequences in these words represent an exception to the generalization in 3. This generalization does not hold of just C_1VC_2 morphemes, but is true of Afrikaans morphemes in general.

(19) Longer Afrikaans roots ending in C_1VC_2

		SINGULAR	PLURAL	
kapok	/kɑpɔk/	[kɑˈpɔk]	no plural	‘light snow’
soldaat	/sɔldɑːt/	[sɔlˈdɑːt]	[sɔlˈdɑːtə]	‘soldier’
krawat	/krɑvɑt/	[krɑˈvɑt]	[krɑˈvɑtə]	‘cravat’
pandok	/pɑndɔk/	[pɑnˈdɔk]	[pɑnˈdɔkə]	‘hovel’
trompet	/trɔmpɛt/	[trɔmˈpɛt]	[trɔmˈpɛtə]	‘trumpet’

3. ON THE ORIGIN OF THE AFRIKAANS PATTERN. The previous section reviewed the evidence for the (near) absence of morphemes ending in TVD from the Afrikaans lexicon. As reviewed in §1.2, the literature acknowledges two routes through which such a pattern can be introduced into the lexicon of a language, an articulatory and a perceptual route. Both of these routes assume that there is something inherently marked (either perceptually or articulatorily) about morphemes that counter the generalization, leading to them being specifically targeted by a historical process. The usual way these types of patterns are introduced into a language’s lexicon is therefore through a sound change explicitly aimed at those morphemes that counter the generalization. This article proposes a third possibility, accidental lexical accumulation, and that this is the route through which the Afrikaans lexicon lost TVD morphemes. There was never a sound change in the history of Afrikaans that specifically targeted TVD morphemes for elimination. What happened instead is that a series of independent sound changes applied, each motivated for reasons unrelated to the markedness of TVD morphemes. Each of these sound changes targeted a larger group of morphemes that included some, but not all, of the TVD morphemes. Over time, the effects of these independent sound changes accumulated in the Afrikaans lexicon, resulting in the complete elimination of TVD morphemes (except for the one exceptional form *voed*; see §2.2). Unlike Afrikaans, its direct primary ancestor, Dutch, has multiple TVD morphemes. This section investigates the slow lexical accumulation that eliminated TVD morphemes from Afrikaans by tracing the development of Dutch TVD morphemes into Afrikaans. As in §2.1, focus is here on simple C_1VC_2 morphemes, although the reported patterns also hold for other roots ending in TVD (see §2.4).

A list was compiled of all Dutch C_1VC_2 morphemes in the CELEX lexical database (Baayen et al. 1995) that countered the generalization in 3—that is, where C_1 is a voiceless obstruent with a voiced counterpart in the Dutch consonantal inventory, and where C_2 is a voiced obstruent.[7] Given the special status of morpheme-final /f/ and /v/ in Afrikaans, Dutch morphemes that end in one of these sounds were excluded from the list. This left a list of fifteen such morphemes from CELEX. Although Dutch clearly has more such roots than Afrikaans (fifteen vs. one), even fifteen is a relatively low number, indicating that this type of root is underrepresented even in Dutch. The rest of this section discusses the different ways in which these morphemes survived (or not) into Afrikaans. A full list of the morphemes with explanatory notes on each is included in Appendix B.

3.1. DUTCH TVD MORPHEMES WITH NO COGNATES IN AFRIKAANS. Out of the fifteen Dutch TVD morphemes, three do not have Afrikaans cognates. These morphemes are listed in 20, together with the evidence that their final phonemes are underlyingly voiced. The Afrikaans word with the corresponding meaning is also given.

[7] Special thanks to René Kager for help in extracting this list of morphemes from CELEX, and for discussions about the distribution of obstruent voicing in Dutch.

(20) a. Nouns

DUTCH					AFRIKAANS			
ROOT	SG		PL		ROOT	SG		
/puz/	poes	[pus]	poezen	[ˈpuzən]	/kat/	kat	[kat]	‘cat’
/xøːz/	geus	[xøːs]	geuzen	[ˈxøːzən]	/fark/	vark	[fark]	‘hog/swine’

b. Verbs

DUTCH				AFRIKAANS			
ROOT	1SG.PRES		INFINITIVE	ROOT	PRESENT TENSE		
/terɣ/	tijg	[teɪx]	tijgen [ˈterɣən]	/aːnklaː/	aankla	[ˈaːnklaː]	‘accuse’

3.2. LOSS OF VOICED OBSTRUENTS FROM THE AFRIKAANS CONSONANT INVENTORY. As shown in the Afrikaans consonant inventory in Table 1, Afrikaans has only limited voicing contrasts among its obstruents. Three pairs of obstruents contrast for voicing (/p, b/ and /t, d/ in all positions, and /f, v/ only in some environments; see §2.3). The remaining three obstruents appear only as voiceless (/x, k, s/) in Afrikaans. Dutch has a richer obstruent inventory, containing voiced versions of /x/ and /s/ (i.e. /ɣ/ and /z/), though they are currently in the process of being lost.

Devoicing of /ɣ/ has been an ongoing change in Dutch, noted as early as 1928 (Zwaardemaker & Eijkman 1928:195) and seeming now to have all but reached completion. Based on recordings of radio news reports, Van de Velde and colleagues (1996:164) report that just over 35% of noninitial underlying /ɣ/s were realized with at least partial voicing in 1935. By 1993, however, this percentage had fallen to less than 9%. In a more recent study, van der Harst and colleagues (2007) found no evidence for a voiced realization of word-initial /ɣ/. Since word-initial /ɣ/ is generally more prone to devoicing than, for instance, word-medial intersonorant /ɣ/, this finding by van der Harst and colleagues does not necessarily mean that voiced velar fricatives are now completely absent from Dutch. Regardless, considering Van de Velde and colleagues’ findings about /ɣ/ in 1935, we can assume that the Dutch that was planted in South Africa in the mid-seventeenth century still had an active contrast between /x/ and /ɣ/.

The situation with /s/ and /z/ is somewhat similar. Van de Velde and colleagues (1996:163) report over 97% voiced realizations of noninitial /z/ in 1935. By 1993, this number had fallen to just below 76%. Although this indicates an ongoing change in Dutch that may eventually result in the loss of the contrast between /s/ and /z/, the contrast is still an active part of Modern Dutch. There is also evidence that, even in Dutch dialects where this voicing contrast is being lost, the contrast between these two segments is being retained in terms of length, with the originally voiceless /s/ realized as longer than the originally voiced /z/ (van Oostendorp 2003).

In Afrikaans, the loss of /ɣ/ and /z/ has been an across-the-board process. All occurrences of historical /ɣ/ and /z/ have been replaced by /x/ and /s/. These devoicing processes therefore did not specifically target TVD morphemes, and the elimination of some TVD morphemes by these devoicing processes is just a side effect of more general processes. In 21, for instance, examples are given of words in which /z/ has undergone devoicing in Afrikaans in non-TVD contexts.

(21)

DUTCH		AFRIKAANS		
zee	[zeː]	see	[siə]	‘sea’
ziel	[zil]	siel	[sil]	‘soul’
keuze	[ˈkøːzə]	keuse	[ˈkøəsə]	‘choice’
boezem	[ˈbuzəm]	boesem	[ˈbusəm]	‘bosom’

There are several Dutch TVD morphemes where the D in C_2 position is /ɣ/ or /z/, sounds that are no longer part of the Afrikaans consonant inventory. In the Afrikaans cognates of these roots, the final voiced obstruents have been replaced by their respective voiceless counterparts, so that these Dutch TVD roots are now TVS roots in Afrikaans. A few examples are given in 22. A full list of examples can be found in Appendix B.

(22)

	ROOT	SG ~ PL			
a.	/peːz/	pees ~ pezen	[peːs] ~ [ˈpeːzən]	‘sinew’	(Dutch)
	/piəs/	pees ~ pese	[piəs] ~ [ˈpiəsə]	‘sinew’	(Afrikaans)
b.	/tøɣ/	teug ~ teugen	[tøx] ~ [ˈtøɣən]	‘slug (of water)’	(Dutch)
	/tøəx/	teug ~ teue	[tøəx] ~ [ˈtøfiə][8]	‘slug (of water)’	(Afrikaans)

Similarly, in some of the Dutch TVD morphemes the T in C_1 position is /x/ or /s/, sounds that have voiced counterparts in Dutch but not in Afrikaans. In Afrikaans, these morphemes are therefore SVD rather than TVD morphemes, and hence not counterexamples to the generalization in 3. An example is given in 23.

(23)

ROOT	SG ~ PL			
/xɔd/	god ~ goden	[xɔt] ~ [ˈxoːdən]	‘god’	(Dutch)
/xɔd/	god ~ gode	[xɔt] ~ [ˈxuədə]	‘god’	(Afrikaans)

3.3. INTERVOCALIC GLIDING OF HISTORICAL /d/. As mentioned in §2.2, historical /d/ in some words undergoes gliding when in intervocalic position in Afrikaans. This process is lexically specific and observed in only a few Afrikaans words. It is most likely the remnant of a sound change that applied for a short period in the history of the language, but stopped short of spreading through the whole lexicon. Dutch has undergone a similar process (Zonneveld 1978:Ch. 2), but, curiously, the set of words to which the process applied is not identical in Afrikaans and Dutch. This process was most likely still active during the time that South Africa was settled by Dutch speakers during the seventeenth century. It applied both in Dutch as spoken in the Netherlands at that time, and in the early years of Dutch as spoken in South Africa. Since there was little to no linguistic contact between continental Dutch and the Dutch of the South African settlers, the process did not affect the same words in these two linguistic communities.

Some Dutch TVD morphemes with a final /d/ were subject to this change in Afrikaans. In these roots, the evidence for the underlying voicedness of the final obstruent has hence been lost in Afrikaans—there are no longer any surface forms of these morphemes in which the morpheme-final consonant is realized as [d]. When this consonant appears in word-final position (when the root morpheme is used without any suffix), it is realized as [t]. When followed by a vowel-initial suffix, it is realized as [j]. Relevant examples are given in 24, with additional examples in Appendix B.

(24)

	ROOT	SG ~ PL			
a.	/teɪd/	tijd ~ tijden	[teɪt] ~ [ˈteɪdən]	‘time’	(Dutch)
	/teɪt/	tyd ~ tye	[teɪt] ~ [ˈteɪjə]	‘time’	(Afrikaans)
b.	/pɑd/	pad ~ paden	[pɑt] ~ [ˈpɑːdən]	‘path’	(Dutch)
	/pɑt/	pad ~ paaie	[pɑt] ~ [ˈpɑːɪjə]	‘path’	(Afrikaans)

This intervocalic gliding of historical /d/ did not specifically target TVD roots. It was a more general process that applied to several /d/-final morphemes, irrespective of the form of the rest of the morpheme. It also applied to some morpheme-internal, intervo-

[8] Root-final /x/ in Afrikaans is always replaced by a glide after affixation of the plural /-ə/.

calic /d/s. Some examples of non-TVD roots that were subject to this change are given in 25. As with the loss of obstruent voicing, this is a process that did not apply in order to create a lexicon consistent with a generalization like that in 3. That morphemes that would have countered this generalization have been eliminated from the Afrikaans lexicon by this change is rather a mere side effect of a more general process.

(25)		ROOT	SG ~ PL			
		/blɑd/	blad ~ bladen	[blɑt] ~ [ˈblɑːdən]	‘sheet’	(Dutch)
		/blɑt/	blad ~ blaaie	[blɑt] ~ [ˈblɑːɪjə]	‘sheet’	(Afrikaans)
	b.	/xəbɔd/	gebod ~ geboden	[xəˈbɔt] ~ [xəˈboːdən]	‘command’	(Dutch)
		/xəbɔt/	gebod ~ gebooie	[xəˈbɔt] ~ [xəˈbuəjə]	‘command’	(Afrikaans)
	c.	/krœɪd/	kruid ~ kruiden	[krœɪt] ~ [ˈkrœɪdən]	‘herb’	(Dutch)
		/krœɪt/	kruid ~ kruie	[krœɪt] ~ [ˈkrœɪjə]	‘herb’	(Afrikaans)
	d.	/pudər/	poeder	[ˈpudər]	‘powder’	(Dutch)
		/pujər/	poeier	[ˈpujər]	‘powder’	(Afrikaans)

3.4. OTHER WORDS. There is also one example of a Dutch TVD morpheme that has an extra vowel in Afrikaans so that the Afrikaans morpheme is of the form TVDV and therefore not in violation of the restriction stated in 3. This example is given in 26.

(26)	ROOT	SG ~ PL			
	/pɑd/	pad ~ padden	[pɑt] ~ [ˈpɑdən]	‘toad’	(Dutch)
	/pɑdɑ/	padda ~ paddas	[ˈpɑdɑ] ~ [ˈpɑdɑs]	‘toad’	(Afrikaans)

Middle Dutch (*padde*, 1200–1500), Middle Low German (*padde*/*pedde*, 1100–1500), and Old English (*padde*, before 1100) all had forms of this noun with a final vowel (Philippa et al. 2007:487–88). Vowel-final forms are also attested in Modern Norwegian (*padde*) and Modern Swedish (*padda*). It is therefore possible that the Afrikaans vowel-final *padda* is the more conservative form, while the Modern Dutch *pad* resulted from loss of the final vowel. However, it is equally possible that Afrikaans *padda* developed from Dutch *pad* through addition of a final vowel. The existing evidence is not sufficient to decide between these two possibilities.

This section has shown that Dutch TVD roots either did not survive into Afrikaans, or are attested in forms that do not counter the generalization in 3. The changes to which they were subject were more general changes that applied throughout the lexicon and did not specifically target TVD morphemes. There was no specific process just for the elimination of these morphemes, and there is therefore no evidence that the generalization in 3 is the result of either the speaker- or listener-oriented forces discussed in §1.2. That the Afrikaans lexicon all but lacks TVD morphemes is rather epiphenomenal to a set of more general, independent changes that applied in the history of Afrikaans.

4. GRAMMATICALIZATION OF THE DISTRIBUTIONAL PATTERN. In §2, it was shown that the Afrikaans lexicon contains strong evidence for the generalization about the distribution of obstruent voicing stated in 3. Section 3 showed that this pattern was introduced into the lexicon through lexical accumulation, rather than through a specific targeted sound change. What is still unknown is whether this lexical pattern is treated by Afrikaans speakers as an accidental pattern (in accordance with how it was introduced into the lexicon), or whether they have elevated this pattern to a grammatical principle. As was reviewed in §1.3, not all lexical patterns are learned by speakers as grammatical principles. Typically, only patterns that are phonetically motivated in some sense are liable to grammaticalization. In fact, as shown in §1.2, phonetic naturalness is often the driving force that introduces patterns into the lexicon. A question that still needs to be

answered is whether there is a causal relationship between the introduction of a pattern into the lexicon and the grammaticalization of such patterns. Is any phonetically natural lexical pattern liable to be grammaticalized, irrespective of how it was introduced into the lexicon? Or are natural patterns that were introduced through accidental lexical accumulation treated just as accidental patterns?

Given that speakers typically do not have knowledge of the changes in the history of their language, it seems reasonable to assume that any phonetically natural pattern should be learnable as a grammatical principle. The expectation would therefore be that Afrikaans speakers should elevate the distributional pattern to a grammatical constraint, and that evidence for this should be observed in the linguistic performance of Afrikaans speakers. This section presents the results of an experiment in which Afrikaans speakers had to rate novel singular ~ plural pairs. If the speakers have internalized the generalization in 3 as a grammatical constraint, they should rate pairs like [tup] ~ [ˈtubə] as particularly bad. This singular ~ plural pair can only come from an underlying form /tub/, that is, a TVD morpheme that counters the generalization in 3.

4.1. DESIGN: TOKEN SELECTION. Forty nonce-word underlying forms with the structure /C_1VC_2/ were selected. The underlying forms were organized into ten quadruples, as shown in Table 3. In each quadruple: (i) the vowel remains constant; (ii) C_1 is a voiceless obstruent (/p, t, f/) in two members, and the corresponding voiced obstruent (/b, d, v/) in the other two members; (iii) each of the two C_1s is paired with both a voiced obstruent (/b/ or /d/) and the corresponding voiceless obstruent (/p/ or /t/) in C_2. The forms in the first column in Table 3 are all examples of the morpheme type that would counter the voicing cooccurrence pattern in Afrikaans. The nonce underlying forms were treated as noun roots, and a singular and plural form were created for each root. For those roots that end in voiceless obstruents, the singular was identical to the root, and the plural identical to the root plus the plural suffix /-ə/. For the root /pɑːp/, for instance, the singular was [pɑːp] and the plural [ˈpɑːpə]. For roots that end in voiced obstruents, the singular showed the effect of final devoicing, while the plural was identical to the root with the addition of /-ə/. For the root /pɑːb/, for instance, the singular was [pɑːp] (with final devoicing) and the plural [ˈpɑːbə].

C_1 = VOICELESS		C_1 = VOICED	
C_2 = VOICED	C_2 = VOICELESS	C_2 = VOICED	C_2 = VOICELESS
/pɑːb/	/pɑːp/	/bɑːb/	/bɑːp/
/pɛb/	/pɛp/	/bɛb/	/bɛp/
/tɑb/	/tɑp/	/dɑb/	/dɑp/
/tɛb/	/tɛp/	/dɛb/	/dɛp/
/tub/	/tup/	/dub/	/dup/
/fɑb/	/fɑp/	/vɑb/	/vɑp/
/fib/	/fip/	/vib/	/vip/
/fub/	/fup/	/vub/	/vup/
/fuəd/	/fuət/	/vuəd/	/vuət/
/fœɪd/	/fœɪt/	/vœɪd/	/vœɪt/

TABLE 3. Nonce-word underlying forms used in the experiment.

In addition to these forty token roots, ten fillers were also selected. The fillers all had as C_2 a consonant that does not contrast for voicing in Afrikaans—five ended in a nasal (/fɔm, fɛn, fun, pɔn, tɑːn/), and five in a velar plosive (/fik, piək, pæk, tiək, tuk/). Singular and plural pairs were also formed for each of the fillers with the addition of the plural suffix /-ə/. Since C_2 in the fillers is either voiced in all contexts (nasal) or voice-

less in all contexts (/k/), the singular was always equal to the root, and the plural to the root plus /-ə/.

In addition to the /-ə/-plural suffix, Afrikaans has another very productive plural suffix /-s/. Each of the nonce-word singulars, both fillers and tokens, was also paired with an /-s/-plural. Since adjacent obstruents in Afrikaans have to agree in voicing, an underlying form like /pɑːb-s/ will be realized as [pɑːps]. There will therefore be no surface difference in the singular ~ plural pairs of roots like /pɑːp/ and /pɑːb/ with an /-s/-plural suffix. Both roots will be realized as [pɑːp] ~ [pɑːps], so that there are only twenty distinct singular ~ plural pairs for the forty token roots with the /-s/-plural. All of the singular ~ plural pairs with /-s/-plurals also functioned as fillers in the experiment.

Stimulus creation. Each of the singular and plural forms was placed in a carrier phrase. Singulars were placed in the phrase *een X* [iən *X*] 'one *X*', and plurals in the phrase *tien X* [tin *X*] 'ten *X*'. These phrases were recorded by a male, adult native speaker of Afrikaans in a sound-attenuated booth in the Phonetics Laboratory at the University of Michigan. Recordings were made with Praat (Boersma & Weenink 2012) at a sampling rate of 44.1 kHz. Several recordings of each phrase were made. A recording that was judged to contain a clear production of the singular or plural was selected. The selected phrases were scaled to have an average intensity of 70 dB, using the Scale intensity… function in Praat. Each singular and plural carrier phrase pair was then spliced together, with the singular first, followed by 750 ms of silence, followed by the plural.

Procedure. Stimulus presentation and response collection were controlled with the web interface of the Qualtrics Labs, Inc., software (Qualtrics Labs 2012). Participants were instructed to do the experiment in a quiet room, using headphones. Each experiment trial consisted of the presentation of one of the singular ~ plural pair recordings described above. Participants were instructed to listen to the recording once and then to indicate how likely they would be to use the particular plural for the made-up word. They indicated their response on a seven-point scale where 1 corresponded to 'I will never use this plural', 4 to 'I may perhaps use this plural', and 7 to 'I am very likely to use this plural'. Each singular ~ plural pair was presented four times during the experiment. Stimuli were differently randomized for each participant. Given that the experiment contained both a /-ə/-plural and an /-s/-plural for each nonce root, participants' attention was focused on the choice between these two suffixes rather than on the voicing of the final root consonant.

Participants. Participants were recruited in two ways. First, some participants were recruited via Amazon's Mechanical Turk (Ipeirotis 2010). A link was provided in the Mechanical Turk interface to the Qualtrics online survey. In order to minimize the possibility that nonnative Afrikaans speakers might participate, all instructions were posted exclusively in Afrikaans. The Qualtrics survey also contained some biographical questions, including a question about the language background of the participants. Nine participants in total were recruited this way. An additional five participants were recruited from the faculty and students at the North-West University in South Africa. Participants received a small monetary reward for their participation. Out of the fourteen total participants, two were excluded for indicating in the biographical part of the Qualtrics survey that English, rather than Afrikaans, was their primary language.

4.2. Results and discussion. A linear mixed-effects model was fitted to the participants' plural ratings in R, using the lmer() function from the lme4 package (Bates et al. 2013). The fixed effects in the model were specified as C_1 voicing (voiced vs. voice-

less), C_2 voicing (voiced vs. voiceless), and their interaction.[9] The model also included a random intercept for each participant, as well as random slopes for participants by C_1 voicing, C_2 voicing, and their interaction. Initially, a random intercept for each nonce root was included in the model. This effect was estimated as having negligible variance, and was excluded from the final model.

The results of the model are reported in Table 4, with *t*-statistics > |2| assumed to correspond to significant effects (Baayen 2008:248). Both main effects and their interaction were found to contribute significantly to the participants' plural ratings. The average rating for each of the root types is represented in Figure 3. The reference value for both C_1 voicing and C_2 voicing was 'voiced', so the negative coefficient for C_1 voicing indicates that a voiceless consonant in C_1 position resulted in an overall lower rating of the plural. This effect was small, however, with a nonce root with a voiceless C_1 being rated on average only 0.31 lower on the seven-point scale. The positive coefficient for C_2 voicing indicates that a voiceless C_2 resulted in a significantly better plural rating. This effect was quite large, with a voiced C_2 resulting in a penalty of nearly 1.8 on the seven-point scale relative to a voiceless C_2. This penalty is expected, since roots with voiced obstruents in C_2 position alternate in their surface forms between the singular (with devoicing of C_2) and the plural (without devoicing). As has been shown by, among others, Coetzee (2009), Hayes (2004), McCarthy (1998), and Tessier (2007), there are learnability reasons why nonalternating morphemes are preferred over alternating morphemes in the absence of specific evidence for alternation. Of particular relevance here are the results of Kerkhoff (2007:198; see also Zamuner et al. 2012), who shows that both Dutch toddlers and adults prefer singular ~ plural nonce-word pairs that do not alternate in the voicing of the final root consonant. Given that all of the roots in this experiment were nonce words, the participants had no prior evidence that any of the roots are alternating morphemes. In agreement with learnability considerations and the results of Kerkhoff, it is to be expected that the participants will prefer nonalternating plurals.

SOURCE OF VARIANCE	COEFFICIENT	STD. ERROR	*t*-STATISTICS
(intercept)	4.32	0.37	11.56
C_1 voicing (voiceless)	−0.31	0.12	−2.54
C_2 voicing (voiceless)	1.79	0.42	4.25
Interaction	0.53	0.16	3.30

TABLE 4. Result of linear mixed model for the fixed effects.

What is of crucial relevance here is the significant interaction between C_1 voicing and C_2 voicing. As Fig. 3 shows, nonalternating pairs (with a voiceless obstruent in C_2 position) are penalized when they have a voiced obstruent in C_1 position—DVT roots are rated slightly worse on average than TVT roots. But in alternating pairs (with a voiced obstruent in C_2 position), it is pairs with a voiceless obstruent in C_1 position that are subject to an additional penalty—TVD roots are rated worse on average than DVD roots. Having a voiced obstruent in C_1 position is dispreferred when C_2 is voiceless, but is preferred when C_2 is voiced. Given that a voiceless C_1 has opposite effects when C_2 is voiced and voiceless, one would expect C_1 voicing not to have an overall significant

[9] Repetition (ranging from one to four) was originally also included. This factor, however, was not found to contribute significantly to the results. A comparison of the models with and without Repetition using the anova() function also confirmed that the full model with Repetition does not fit the data significantly better.

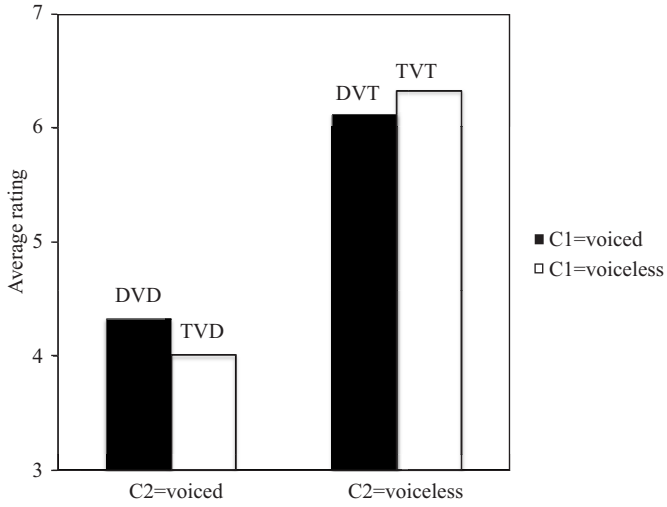

FIGURE 3. Average rating for each of the four singular ~ plural pair types.

effect. As reported above, however, C_1 voicing did have a significant main effect such that a voiceless C_1 on average resulted in a lower plural rating—that is, in agreement with the effect of a voiceless C_1 when C_2 is voiced, and counter to the effect of a voiceless C_1 when C_2 is voiceless. That a voiceless C_1 results in an overall penalty shows that the penalty for having a voiceless C_1 when C_2 is voiced outweighs the boost of having a voiceless C_1 when C_2 is voiceless. This result shows that the participants in the experiment relied on, among other things, a restriction such as that in 3.[10]

Although this experiment gives evidence that the participants have elevated the distributional pattern in 3 to a grammatical restriction that influences how they perform linguistically, Fig. 3 also shows that the effect of this restriction is small. The difference in average rating between TVD roots and DVD roots is approximately 0.25 on a seven-point scale. There are two possible explanations for the size of the effect. One possibility is that the effect is really very small, implying a weak, gradient effect. Another possibility is that there is significant variance between the participants, and that some participants show a stronger effect while others show a weaker or no effect. Both of these possibilities probably contribute to the size of the effect. The participants assigned TVD roots (i.e. roots that violate the constraint) an average acceptability rating of about 4 on the seven-

[10] A referee points out another possible interpretation of the results depicted in Fig. 3. First, it is clear that having a voiced C_2 is rather severely penalized, so that both forms with a voiceless C_2 are rated much better than both forms with a voiced C_2 (DVT, TVT > DVD, TVD). Once the effect of voiced vs. voiceless C_2 has been taken into account, then a smaller penalty against forms with obstruents that differ in voicing arises. Roots with two voiceless obstruents are rated better than roots with a voiced obstruent followed by a voiceless obstruent (TVT > DVT). Similarly, roots with two voiced obstruents are rated better than roots with a voiceless obstruent followed by a voiced obstruent (DVD > TVD). This would mean that Afrikaans is a language with voicing harmony, that is, with a preference for roots where all obstruents agree in voicing, similar to Kera discussed in §1.2. The results from this experiment are consistent with such an interpretation. Additionally, as the data in Table 2 show, roots with alternating voicing (TVD and DVT) are the most underrepresented of all C_1VC_2 root types in Afrikaans (with O/E ratios of 0.06 and 0.69, respectively). However, given that there are twenty-eight roots of the type DVT in the Afrikaans lexicon and only one marginal TVD root, the lexical evidence is still overwhelmingly in favor of a restriction specifically against TVD roots. Should Afrikaans lose more of its DVT roots, it may well develop into a language with voicing harmony—but it is not yet there.

point scale, indicating that they did not experience such roots as absolutely ill-formed (i.e. that the constraint is gradient rather than categorical). The results of the mixed model analysis also show that there was considerable between-participant variation—of the variance in the data not accounted for by the main effects and their interaction, roughly 67% is due to between-participant variance. In order to investigate this further, a difference score was calculated for each participant by subtracting the participant's rating for TVD roots from that for DVD roots. A positive difference score indicates a participant that relied on the restriction in 3—that is, that rated DVD roots better than TVD roots. These difference scores, represented in Figure 4, show that out of the twelve participants, half rated DVD roots better than TVD roots, while the other half showed little or no difference between these two root types.

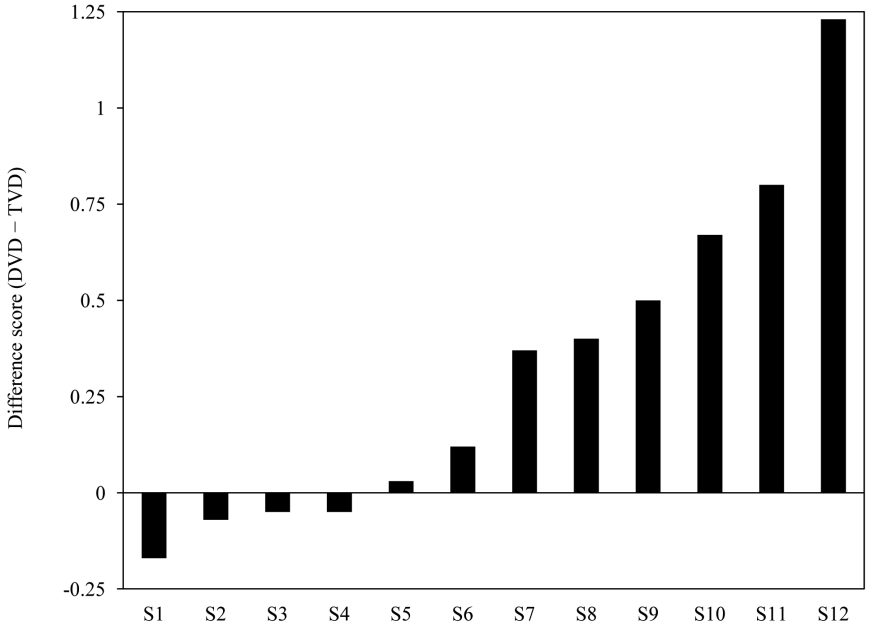

FIGURE 4. Difference scores for each participant, calculated by subtracting a participant's average rating for TVD roots from his/her average rating for DVD roots.

The results of this experiment are consistent with a scenario in which a restriction like that in 3 is psychologically real for at least some speakers of Afrikaans, so that this constraint influences how they rate the nonce-word singular ~ plural pairs in the experiment. The results must be interpreted with caution, however. First, as pointed out above, the average rating difference between TVD and DVD roots is small, and there are speakers that do not differentiate between these forms at all. This might be indicative of a newly arising constraint that is only now starting to affect how Afrikaans speakers evaluate TVD-type roots. Second, the current experiment does not directly address the question of the source of the rating differences for those speakers who do differentiate between TVD and DVD roots. Does this difference arise because, for such speakers, TVD roots violate a grammatical constraint? Or is the origin of the difference to be found in frequency effects—that is, do the DVD roots used in the experiment have higher lexical statistics (lexical neighborhood densities, biphone transitional probabilities, etc.) than the TVD roots? If this were the case, the higher ratings for the DVD roots

may just reflect their frequency advantage over the TVD roots. Such a frequency explanation for the effect is unlikely, however. As shown in Table 2, there are many more Afrikaans roots with a voiceless than a voiced C_1 (180 vs. 84). It is therefore more likely that TVD roots will have the upper hand over DVD roots in terms of lexical statistics such as biphone transitional probabilities and lexical neighborhood densities. Additionally, even if the advantage of DVD roots had such a frequency origin, it would be consistent with the main argument of this article, namely that a grammatical constraint active in some language can have its origin in (frequency) patterns in the lexicon of that language.

5. Concluding remarks. This article has documented a cooccurrence restriction on obstruent voicing in Afrikaans. In morphemes ending in the sequence C_1VC_2, C_2 can be a voiced obstruent (one of /b, d, v/) only if C_1 is not a viable voicing sponsor (i.e. when C_1 is already voiced or it is not voiceable). It was shown that the Afrikaans lexicon gives evidence for such a restriction, and that such a restriction influences how Afrikaans speakers rate nonce-word singular ~ plural pairs.

Afrikaans is not unique in having a restriction such as this. In fact, the Chadic language Ngizim has a restriction that is very similar to that in Afrikaans. However, Afrikaans does differ from most other languages with such cooccurrence patterns in terms of how the pattern arose. The typical route for the development of this type of pattern is either via a listener-oriented perceptual process (misperception) or via a speaker-oriented articulatory process (articulatory simplification or speech errors). It was shown here that the Afrikaans restriction arose via a different route, that of lexical accumulation. The effects of a series of unrelated sound changes that affected the Afrikaans lexicon across the board (i.e. that did not specifically target TVD morphemes) accumulated over time. As an accidental side effect of these changes, the Afrikaans lexicon came to lack TVD-final morphemes.

Over the past several decades, research in language processing has shown that language users are extremely skilled statistical learners. They are very adept at noticing statistical patterns in the lexicon, and such patterns can then be elevated to general principles or constraints, and eventually can be grammaticalized. Hayes and Wilson (2008) review the relevant literature, and also develop and test a computational implementation of a statistical learner that can do exactly this (see also Albright & Hayes 2001, Hayes et al. 2009, Hayes & White 2013, Peperkamp et al. 2006, etc.).

That the Afrikaans lexicon was shaped via lexical accumulation rather than the more direct speaker- or listener-oriented routes does not mean that the voicing cooccurrence pattern in the Afrikaans lexicon cannot be learned as a grammatical constraint. Since this pattern is consistent with phonetic naturalness, it should be just as easily learnable as similar patterns that were established directly through the listener- or speaker-oriented routes (Becker et al. 2011, Becker et al. 2012, Hayes & White 2013, etc.). This expectation was confirmed by the results of a nonce-word rating experiment, which showed that (at least some) Afrikaans speakers have elevated the distributional pattern to a grammatical principle that influences how they rate novel plurals.

Not all phonetically natural patterns in a language's lexicon were necessarily established in the lexicon as a consequence of their phonetic naturalness, even if this is the usual route through which such patterns come about. The usual route cannot be assumed correct in the absence of specific evidence for it. Such an approach is liable to result in incorrect reconstruction of the historical developments of a language's lexicon: 'similar sound patterns may have arisen in different ways from different kinds of sound change' (Blevins 2004:48).

APPENDIX A: EXAMPLES OF EACH OF THE MORPHEME KINDS REPRESENTED IN FIGURE 2

For the one example of the type TVD, *voed* /fud/ ‘feed’, see the discussion in §2.2.

TYPE	UNDERLYING	SINGULAR		PLURAL		
LVL	/mɑn/	man	[mɑn]	manne	[ˈmɑnə]	‘man’
	/rɑːm/	raam	[rɑːm]	rame	[ˈrɑːmə]	‘frame’
LVS	/reɪs/	reis	[reɪs]	reise	[ˈreɪsə]	‘journey’
	/leɪk/	lyk	[leɪk]	lyke	[ˈleɪkə]	‘corpse’
LVT	/rɑːt/	raat	[rɑːt]	rate	[ˈrɑːtə]	‘advice’
	/rup/	roep	[rup]	roepe	[ˈrupə]	‘call’
LVD	/rɑːd/	raad	[rɑːt]	rade	[ˈrɑːdə]	‘board’
	/lɔb/	lob	[lɔp]	lobbe	[ˈlɔbə]	‘cell’
SVL	/sɑːl/	saal	[sɑːl]	sale	[ˈsɑːlə]	‘hall/auditorium’
	/kɑm/	kam	[kɑm]	kamme	[ˈkɑmə]	‘comb’
SVS	/kaːs/	kaas	[kɑːs]	kase	[ˈkɑːsə]	‘cheese’
	/xæk/	gek	[xæk]	gekke	[ˈxækə]	‘fool’
SVT	/xɑt/	gat	[xɑt]	gate	[ˈxɑːtə]	‘hole’
	/sɔp/	sop	[sɔp]	soppe	[ˈsɔpə]	‘soup’
SVD	/sɑːd/	saad	[sɑːt]	sade	[ˈsɑːdə]	‘seed’
	/xɔd/	god	[xɔt]	gode	[ˈxuədə]	‘god’
DVL	/bɑl/	bal	[bɑl]	balle	[ˈbɑlə]	‘ball’
	/dul/	doel	[dul]	doele	[ˈdulə]	‘goal’
DVS	/bœs/	bus	[bœs]	busse	[ˈbœsə]	‘bus’
	/duk/	doek	[duk]	doeke	[ˈdukə]	‘diaper’
DVT	/dɔp/	dop	[dɔp]	doppe	[ˈdɔpə]	‘shell’
	/vəp/	wip	[vəp]	wippe	[ˈvəpə]	‘trap’
DVD	/dɑːd/	daad	[dɑːt]	dade	[ˈdɑːdə]	‘deed’
	/boʊd/	boud	[boʊt]	boude	[ˈboʊdə]	‘buttock’
TVL	/tɑːl/	taal	[tɑːl]	tale	[ˈtɑːlə]	‘language’
	/pɛn/	pen	[pɛn]	penne	[ˈpɛnə]	‘pen’
TVS	/fɑːs/	vaas	[fɑːs]	vase	[ˈfɑːsə]	‘vase’
	/tɑk/	tak	[tɑk]	takke	[ˈtɑkə]	‘branch’
TVT	/puət/	poot	[puət]	pote	[ˈpuətə]	‘paw’
	/feɪt/	feit	[feɪt]	feite	[ˈfeɪtə]	‘fact’

APPENDIX B: DUTCH TVD MORPHEMES IN AFRIKAANS (SEE §3)

There are no morphemes with a final /b/ in this list. Overall, /b/ is relatively rare in both Afrikaans and Dutch. Kager and Shatzman (2007; also Shatzman & Kager 2007) also show that labial segments prefer to occur as close as possible to the beginning of morphemes. These two facts together are the most likely reason for the absence of /b/-final morphemes from this list.

NOUNS

UF	SINGULAR		PLURAL			WHAT HAPPENED IN AFRIKAANS
/puz/	poes	[pus]	poezen	[ˈpuzən]	‘cat’	No Afrikaans cognate.
/teɪd/	tijd	[teɪt]	tijden	[ˈteɪdən]	‘time’	/teɪt/: Plural formed with glide [j], not [d]. (TVT in Afrikaans.)
/sɑʊz/	saus	[sɑʊs]	sauzen	[ˈsɑʊzən]	‘sauce’	/soʊs/: */z/ > /s/ in Afrikaans. (SVS in Afrikaans.)
/tøɣ/	teug	[tøx]	teugen	[ˈtøɣən]	‘slug (of beer)’	/tøəx/: */ɣ/ > /x/ in Afrikaans. (TVS in Afrikaans.)
/pɑd/	pad	[pɑt]	padden	[ˈpɑdən]	‘toad/frog’	/ˈpɑdɑ/: Afrikaans has additional vowel in singular. (TVDV in Afrikaans.)
/xɔd/	god	[xɔt]	goden	[ˈxoːdən]	‘god’	/xɔd/: */ɣ/ > /x/ in Afrikaans. (SVD in Afrikaans.)
/pɑd/	pad	[pɑt]	paden	[ˈpɑːdən]	‘path’	/pɑt/: Plural formed with [j], not [d]. (TVT in Afrikaans.)
/tœʏɣ/	tuig	[tœʏx]	tuigen	[ˈtœʏɣən]	‘harness’	/tœʏx/: */ɣ/ > /x/ in Afrikaans. (TVS in Afrikaans.)

(*Table continues*)

UF	SINGULAR		PLURAL			WHAT HAPPENED IN AFRIKAANS
/xøz/	geus	[xøs]	geuzen	[ˈxøːzən]	‘hog/swine’	No Afrikaans cognate.
/pez/	pees	[pes]	pezen	[ˈpezən]	‘sinew’	/piəs/: */z/ > /s/ in Afrikaans. (TVS in Afrikaans.)

ADJECTIVES

UF	PREDICATIVE		ATTRIBUTIVE			WHAT HAPPENED IN AFRIKAANS
/xɑʊd/	goud	[xɑʊt]	gouden	[ˈxɑʊdən]	‘gold’	/xoʊt/: Attributive formed with [j], not [d]. */ɣ/ > /x/ in Afrikaans. (SVT in Afrikaans.)
/xud/	goed	[xut]	goeden	[ˈxudən]	‘good’	/xut/: Attributive formed with [j], not [d]. */ɣ/ > /x/ in Afrikaans. (SVT in Afrikaans.)

VERBS

UF	ROOT/1.SG PRES		INFINITIVE			WHAT HAPPENED IN AFRIKAANS
/poːɣ/	poog	[poːx]	pogen	[ˈpoːɣən]	‘attempt/try’	/puəx/: */ɣ/ > /x/ in Afrikaans. (TVS in Afrikaans.)
/teɪɣ/	tijg	[teɪx]	tijgen	[ˈteɪɣən]	‘accuse’	No Afrikaans cognate.
/sœʏz/	suis	[sœʏs]	suizen	[ˈsœʏzən]	‘hiss’	/sœɪs/: */z/ > /s/ in Afrikaans. (SVS in Afrikaans.)

REFERENCES

ALBRIGHT, ADAM, and BRUCE HAYES. 2001. Learning and learnability in phonology. *The handbook of phonological theory*, 2nd edn., ed. by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu, 661–90. Cambridge: Blackwell.

APPLEGATE, RICHARD B. 1972. *Ineseño Chumash grammar*. Berkeley: University of California, Berkeley dissertation.

BAAYEN, R. HARALD. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

BAAYEN, R. HARALD; RICHARD PIEPENBROCK; and LEON GULIKERS. 1995. The CELEX lexical database [CD-ROM]. Philadelphia: Linguistic Data Consortium.

BATES, DOUGLAS; MARTIN MAECHLER; and BEN BOLKER. 2013. lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-2. Online: http://CRAN.R-project.org/package=lme4. Downloaded June 9, 2013.

BECKER, MICHAEL; NIHAN KETREZ; and ANDREW NEVINS. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 81.84–125.

BECKER, MICHAEL; ANDREW NEVINS; and JONATHAN LEVINE. 2012. Asymmetries in generalizing alternations to and from initial syllables. *Language* 88.231–68.

BLEVINS, JULIETTE. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

BLEVINS, JULIETTE. 2006. A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics* 32.117–65.

BLUST, ROBERT. 1995. Notes on Berawan consonant gemination. *Oceanic Linguistics* 34. 123–38.

BOERSMA, PAUL, and DAVID WEENINK. 2012. Praat: Doing phonetics by computer. Version 5.3.56. Online: http://www.praat.org.

BOSMAN, D. B.; I. W. VAN DER MERWE; and L. M. HIEMSTRA (eds.) 1984. *Tweetalige woordeboek/Bilingual dictionary*. 8th rev. and enlarged edn. Cape Town: Tafelberg.

BRADSHAW, MARY. 1999. *A crosslinguistic study of consonant-tone interaction*. Columbus: The Ohio State University dissertation.

COETZEE, ANDRIES W. 2009. Learning lexical indexation. *Phonology* 26.109–45.

COETZEE, ANDRIES W.; SUSAN LIN; and RIGARDT PRETORIUS. 2007. Post-nasal devoicing in Tswana. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, 861–64. Online: http://www.icphs2007.de/.

COETZEE, ANDRIES W., and JOE PATER. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26.289–337.

COETZEE, ANDRIES W., and RIGARDT PRETORIUS. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38.404–21.
DELL, GARY S.; LISA K. BURGER; and WILLIAM R. SVEC. 1997. Language production and serial order: A functional analysis and model. *Psychological Review* 104.123–47.
DEN BESTEN, HANS. 1999. Speculations on [χ]-elision and intersonorantic [ʊ] in Afrikaans. *Stellenbosch Papers in Linguistics (Language genesis)* 32.45–66.
FRISCH, STEFAN A.; JANET PIERREHUMBERT; and MICHAEL B. BROE. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22.179–228.
GAFOS, ADAMANTIOS. 1998. Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16.223–78.
GAFOS, ADAMANTIOS. 1999. *The articulatory basis of locality in phonology*. New York: Garland.
GAFOS, ADAMANTIOS. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20.269–337.
GALLAGHER, GILLIAN. 2010. Perceptual distinctness and long-distance laryngeal restrictions. *Phonology* 27.435–80.
GRIESHABER, N. J. 1987. Onstemming en /f/-stemgewing van nader beskou. *South African Journal of Linguistics* 5.41–67.
HANSSON, GUNNAR ÓLAFUR. 2001a. The phonologization of production constraints: Evidence from consonant harmony. *Chicago Linguistic Society* 37.1.187–200.
HANSSON, GUNNAR ÓLAFUR. 2001b. *Theoretical and typological issues in consonant harmony*. Berkeley: University of California, Berkeley dissertation.
HANSSON, GUNNAR ÓLAFUR. 2004. Long-distance voicing agreement: An evolutionary perspective. *Berkeley Linguistics Society* 30.130–41.
HANSSON, GUNNAR ÓLAFUR. 2007. On the evolution of consonant harmony: The case of secondary articulation agreement. *Phonology* 24.77–120.
HANSSON, GUNNAR ÓLAFUR. 2008. Effects of contrast recoverability on the typology of harmony systems. *Contrast in phonology: Theory, perception, acquisition*, ed. by Peter B. Avery, Elan Dresher, and Keren Rice, 115–41. Berlin: Mouton de Gruyter.
HANSSON, GUNNAR ÓLAFUR. 2010. *Consonant harmony: Long-distance interaction in phonology*. Berkeley: University of California Press.
HAYES, BRUCE. 1999. Phonetically-driven phonology: The role of optimality theory and inductive grounding. *Functionalism and formalism in linguistics, vol. 1: General papers*, ed. by Michael Darnell, Edith A. Moravscik, Michael Noonan, Frederick J. Newmeyer, and Kathleen Wheatly, 243–85. Amsterdam: John Benjamins.
HAYES, BRUCE. 2004. Phonological acquisition in optimality theory: The early stages. *Constraints in phonological acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 158–203. Cambridge: Cambridge University Press.
HAYES, BRUCE, and JAMES WHITE. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44.45–75.
HAYES, BRUCE, and COLIN WILSON. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.
HAYES, BRUCE; KIE ZURAW; PÉTER SIPTÁR; and ZSUZSA LONDE. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85.822–63.
HYMAN, LARRY. 2001. On the limits of phonetic determinism in phonology: *NC revisited. *The role of speech perception in phonology*, ed. by Elizabeth Hume and Keith Johnson, 141–85. New York: Academic Press.
IPEIROTIS, PANAGIOTIS G. 2010. Analyzing the Amazon Mechanical Turk marketplace. *XRDS* 17.16–21.
KAGER, RENÉ, and KEREN SHATZMAN. 2007. Phonological constraints in speech processing. *Linguistics in the Netherlands* 2007.99–111.
KERKHOFF, ANNEMARIE O. 2007. *Acquisition of morpho-phonology: The Dutch voicing alternation*. Utrecht: Utrecht University dissertation.
KIPARSKY, PAUL. 1973. Phonological representations. *Three dimensions of linguistic theory*, ed. by Osamu Fujimura, 1–136. Tokyo: TEC.
LINDBLOM, BJÖRN. 1990. Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modeling*, ed. by William J. Hardcastle and Alain Marchal, 403–39. Dordrecht: Kluwer.

LUBBE, JOHAN, and WIM ZONNEVELD. 1996. Ontstemming of stemgewing—weer eens oor die v/f-wisseling in Afrikaans. *South African Journal of Linguistics* 14.109–17.
ŁUBOWICZ, ANNA. 2002. Derived environment effects in optimality theory. *Lingua* 112. 243–80.
MACEACHERN, MARGARET R. 1999. *Laryngeal cooccurrence restrictions*. New York: Garland.
MACKENZIE, SARA. 2009. *Contrast and similarity in consonant harmony processes*. Toronto: University of Toronto dissertation.
MCCARTHY, JOHN J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17.207–63.
MCCARTHY, JOHN J. 1998. Morpheme structure constraints and paradigm occultation. *Chicago Linguistic Society* 32.2.123–50.
MCCARTHY, JOHN J. 2003. Comparative markedness. *Theoretical Linguistics* 29.1–51.
MORETON, ELLIOTT. 2008. Analytic bias and phonological typology. *Phonology* 25.83–127.
NEWMAN, PAUL. 1990. *Nominal and verbal plurality in Chadic*. Dordrecht: Foris.
NÍ CHIOSÁIN, MÁIRE, and JAYE PADGETT. 2001. Markedness, segment realization, and locality in spreading. *Segmental phonology in optimality theory: Constraints and representations*, ed. by Linda Lombardi, 118–56. Cambridge: Cambridge University Press.
OHALA, JOHN J. 1993. The phonetics of sound change. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 235–78. London: Longman.
OHALA, JOHN J. 2005. Phonetic explanations for sound patterns: Implications for grammars of competence. *A figure of speech: A festschrift for John Laver*, ed. by William J. Hardcastle and Janet Mackenzie Beck, 23–38. London: Lawrence Erlbaum.
PARKER, STEPHEN. 1997. An OT account of laryngealization in Cuzco Quechua. *Working Papers of the Summer Institute of Linguistics* 41.1–11.
PEARCE, MARY. 2005. Kera tone and voicing. *University College London Working Papers in Linguistics* 17.61–82.
PEARCE, MARY. 2009. Kera tone and voicing interaction. *Lingua* 119.846–64.
PEPERKAMP, SHARON; ROZENN LE CALVEZ; JEAN-PIERRE NADAL; and EMMANUEL DUPOUX. 2006. The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101.B31–B41.
PHILIPPA, MARLIES; FRANS DEBRABANDERE; AREND QUAK; TANNEKE SCHOONHEIM; and NICOLINE VAN DER SIJS. 2007. *Etymologisch Woordenboek van het Nederlands*. Amsterdam: Amsterdam University Press.
QUALTRICS LABS, INC. 2012. Qualtrics research suite. Provo, UT. Online: http://www.qualtrics.com.
ROSE, SHARON. 2011. Long-distance assimilation of consonants. *The Blackwell companion to phonology, vol. 3: Phonological processes*, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 1811–37. Oxford: Blackwell.
ROSE, SHARON, and RACHEL WALKER. 2004. A typology of consonant agreement as correspondence. *Language* 80.475–531.
SCHUH, RUSSELL G. 1978. Bade/Ngizim vowels and syllable structure. *Studies in African Linguistics* 9.247–83.
SCHUH, RUSSELL G. 1997. Changes in obstruent voicing in Bade/Ngizim. Los Angeles: University of California, Los Angeles, MS.
SCHWARTZ, MYRNA F.; ELEANOR M. SAFFRAN; DIANE E. BLOCH; and GARY S. DELL. 1994. Disordered speech production in aphasic and normal speakers. *Brain and Language* 47.52–88.
SHATZMAN, KEREN, and RENÉ KAGER. 2007. A role for phonotactic constraints in speech perception. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, 1409–12. Online: http://www.icphs2007.de/.
SOLÉ, MARIA-JOSEP, and LARRY HYMAN. 2010. More on post-nasal devoicing: The case of Sekgalagari. *Journal of Phonetics* 38.604–15.
TESSIER, ANNE-MICHELLE. 2007. *Biases and stages in phonological acquisition*. Amherst: University of Massachusetts dissertation.
VAN DE VELDE, HANS; MARINEL GERRITSEN; and ROELAND VAN HOUT. 1996. The devoicing of fricatives in Standard Dutch: A real-time study based on radio recordings. *Language Variation and Change* 8.149–75.

VAN DER HARST, SANDER; HANS VAN DE VELDE; and BERT SCHOUTEN. 2007. Acoustic characteristics of standard Dutch /ɣ/. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, 1469–72. Online: http://www.icphs2007.de/.

VAN OOSTENDORP, MARC. 2003. Ambisyllabicity and fricative voicing in West Germanic dialects. *The syllable in optimality theory*, ed. by Caroline Féry and Ruben van de Vijver, 304–37. Cambridge: Cambridge University Press.

WISSING, DAAN. 1982. *Algemene en Afrikaanse Generatiewe Fonologie*. Johannesburg: Macmillan.

ZAMUNER, TANIA S.; ANNEMARIE O. KERKHOFF; and PAULA FIKKERT. 2012. Phonotactics and morphophonology in early child language: Evidence from Dutch. *Applied Psycholinguistics* 33.481–99.

ZONNEVELD, WIM. 1978. *A formal theory of exceptions in generative phonology*. Lisse: Peter de Ridder.

ZWAARDEMAKER, HENDRIK, and LEONAARD P. H. EIJKMAN. 1928. *Leerboek der Phonetiek, Inzonderheid met Betrekking tot het Standaard-Nederlandsch*. Haarlem: Erven F. Bohn.

Department of Linguistics
440 Lorch Hall
University of Michigan
611 Tappan Street
Ann Arbor, MI 48109
[coetzee@umich.edu]

[Received 6 December 2012;
revision invited 15 July 2013;
revision received 20 August 2013;
accepted 12 January 2014]