Roadblocks in the Grammaticalization Highway: When Phonology Gets in the Way

BRADLEY MCDONNELL

University of California, Santa Barbara

1 Introduction

In his famous characterization of syntactic change, Langacker likens languages to “gigantic expression compacting machines” (1977:106). He writes,

The machine does whatever it can to wear down the expressions fed into it. … It attacks expressions of all kinds by phonetic erosion. It bleaches lexical items of their semantic content and forces them into service as grammatical markers. It chips away at the boundaries between elements and crushes them together into smaller units. The machine has a voracious appetite. (Langacker 1977:106)

When a lexical item is fed into the machine, the resulting formative is only a fraction of the original in both form (by phonetic erosion) and meaning (by semantic bleaching). In a well-known case, the Old English noun *litc meaning ‘body’ was fed into the machine and, over centuries of use, the machine bleached its semantic content, wore down its phonetic form, and compacted it with the adjective that preceded it. These processes worked together to create a vaguely reminiscent adverbial suffix -ly in present day English (Lehmann 1995). This coupling of semantic, morphosyntactic, and phonetic processes in a single

---

1 Thanks to Marianne Mithun and the audience at the 39th Annual Berkeley Linguistics Society Meeting for helpful feedback on this paper. All mistakes are my own.
Bradley McDonnell

*machine* has been a foundational principle in studies of grammaticalization and has been rigorously demonstrated in studies like Heine and Reh (1984) and Bybee, Perkins, and Pagliuca (1994). Furthermore, observed trends in language after language have given strong support to a related hypothesis of unidirectionality, which states that the concomitant changes subsumed under grammaticalization move in one direction. In particular, the unidirectionality hypothesis has been robustly exemplified on the formal dimension of grammaticalization with the cline: lexical item > clitic > affix (Hauspelmuth 2004).

However, a number of studies have pointed out examples of changes in the opposite direction, where grammatical elements have become more autonomous (i.e., infinitival markers in English (Fischer 2000)), gained semantic content (i.e., modal to full verb in Pennsylvania German (Burridge 1998)), or developed from suffixes to clitics (i.e., English *s*-genitive clitic from genitive case suffix (Norde 2009). Although examples of degrammaticalization have not derailed notions of unidirectionality in grammaticalization, they have revealed that changes in one area of a linguistic system can have an unpredictable effect on the grammaticalization process in another. For example, it was the loss of case in Old English that led to the reanalysis of *-s* as a clitic.

What is more interesting is the case presented by Bisang (2004), who shows that grammaticalization in the geographic regions of East and Southeast Asia need not include these concomitant processes of semantic generalization and phonetic reduction. Bisang proposes, ‘[a]lthough the coevolution of form and meaning in terms of Bybee (1985) and C. Lehmann (1995) works for a large number of languages, it does not seem to be universal as we can see from the perspective of East and mainland Southeast Asian languages…’ (2004: 109). In a similar vein, Schiering proposes that phonetic erosion ‘is not universally associated with grammaticalization’ and that cross-linguistic evidence suggests that stress-based as opposed to syllable-based languages show higher degrees of phonetic erosion (2010:74-75). Schiering demonstrates that a stress-based language (German) undergoes more phonetic reduction diachronically than a syllable-based language (Turkish).

In light of these more recent proposals that question the unidirectionality of grammaticalization and the necessity of phonetic erosion (or more broadly autonomy), the third person pronominal suffix *-nye* in Besemah, a little-known Malay language of southwest Sumatra, presents an interesting puzzle. In the majority of constructions *-nye* attaches directly to the root and has undergone phonetic erosion (i.e., *-nye* > *-e*): *anak-nye* > *anak-e* ‘their child’. In deverbal noun constructions suffixed with *-an*, or circumfixed with *peN- -an*, however, the *-nye* is separated from the root and surfaces as a separate phonological word (i.e., *-nye* > *anye*): *langkah-an-nye* > *langkah-an anye* ‘his stride’ (from the root *langkah* ‘to step’). One might expect that the most phonetically robust and prosodically independent form (*anye*) is the oldest, but there is strong evidence going back to Proto-Austronesian that it is not. Therefore, what do we make of a case where a suffix by all other accounts is grammaticalizing, but in some
Roadblocks in the Grammaticalization Highway

constructions is phonetically strengthened and phonologically less bound (i.e., suffix > word)? Arguably, such a case should be added to the growing list of examples of degrammaticalization because -nye is apparently climbing up the cline. However, what is more interesting than this taxonomy is determining the conditions, pressures, and/or motivations by which this change emerged. It turns out that this unexpected alternation between -e, -nye, and anye can be attributed to a morphophonological (or more specifically morpho-prosodic) preference in Besemah for no more than one suffix per word. This preference alongside restrictions on phonological words and routinization has resulted in a number of unusual patterns, most importantly the alternations resulting from the suffix -nye attaching to other suffixes.

The discussion here proceeds as follows. Section 2 presents an overview of relevant aspects of the phonological and morphological systems of Besemah. Section 3 presents the synchronic description and diachronic development of -nye > anye, including the morphosyntactic distribution, semantics, morphophonemics, and etymology of -nye. Section 4 then provides the diachronic explanation for -nye > anye, drawing on evidence from a corpus of naturalistic speech. Section 5 describes other cases in the language where previously dependent forms become autonomous, and section 6 presents a possible grammaticalization pathway for these changes. Section 7 concludes the study.

2 Overview of Besemah

Besemah is a Malay language spoken in the remote highlands of southwest Sumatra by approximately 400,000 people. Aside from work in the early 20th century, including a dictionary, short text collection, and basic grammar sketch (Helfrich 1904), little recent linguistic work has been done on the language. Besemah is similar to well known varieties of Malay, namely Standard Malay-Indonesian, but it differs in interesting ways. For example, Besemah has roughly the same consonant inventory with the addition of voiced velar fricative /ɣ/, but it has a much more conservative vowel inventory (i.e., /a, i, u, ə/) because the language never developed mid vowels /e, o/ (McDonnell 2008).

Morphologically, Besemah has a slightly reduced inventory of grammatical morphemes. Prefixes include the active voice N-, passive voice di-, non-volitional te-, middle voice be-, and agent nominalizer peN-. The suffixes include the nominalizer -an, causative/applicative -kah, and the locative applicative -i. The circumfixes include the locative nominalizer peN- -an. Other morphemes only occur as fossilized forms in a small set of words (i.e., nominalizing circumfixes ke- -an and per- -an). The majority of these morphemes do not co-occur on the same word, with the exception of the voice prefixes and applicative suffixes.

Besemah has a slightly reduced set of pronominal affixes with only two forms: the first person =ku and third person -nye. These morphemes have

2 The N- represents an underspecified nasal that is homoroganic to the first consonant of the root.
traditionally been referred to as enclitics, but to my knowledge this position has never been actively defended. Although there is some evidence that the first person =ku is an enclitic, there is little evidence that -nye in Besemah should be characterized as an enclitic (see section 3 below). What is important to note about these pronominal formatives is that both forms occur freely after nominal roots, but after verbs only the third person -nye occurs with several syntactic restrictions. However, the following sections will demonstrate that the pronominal formatives behave erratically when attached to morphologically complex forms.

Lastly, it is useful in Besemah to distinguish between phonological and grammatical words (cf. Dixon and Aikhenvald 2003); only the former is discussed here. Lexical items in Besemah are largely disyllabic. There are a small number of lexical items that are monosyllabic, but these are typically borrowed words and consist of closed (CVC) syllables. Besemah words that are larger than two syllables have a number of restrictions. By and large any prepenultimate syllable is reduced to schwa, a historical process that occurred in all of the Malay languages of Sumatra (Adelaar 1992). In Besemah, this process did not apply to vowel sequences, so there are high vowels /i, u/ in the prepenultimate syllable in words like siamang [sijamaŋ] ‘gibbon.’ There are some marginal examples of lexical items that are longer including mataghi [mataɣi] ‘sun.’ However, this case is quite clearly a fused compound, mate [mata] ‘eye’ and aghi [aɣi] ‘day’. As in many Malay languages, word-level stress in Besemah is not entirely clear, but it appears that stress falls on the final syllable of the word (McDonnell 2013). If a root is suffixed, stress shifts to that suffix (i.e., the final syllable of the word).

3 The Form, Function, and Origin of -nye

The -nye suffix in Besemah is cognate with the well-documented enclitic/suffix -nya in Standard Malay-Indonesian (Macdonald 1976: 71-74, 126-127, Cumming 1991: 25-28, Ewing 1995: 237-238). As with its cognates in other Malay varieties, the meaning/function of the Besemah -nye is difficult to pin down. Nevertheless, it occurs primarily as the patient in an active transitive construction (1), the agent in a passive construction (2), and the possessor of an NP (3).

(1) Aku gale jiku (m)-masuk-i-nye,
    ISG all say.ISG AV-enter-APPL-3
    ‘I alone, I said, entered them into the group,’

(2) Dide di-ruru-i-nye agi li Tabran.
    NEG UV-take.care-APPL-3 again by T.
    ‘(It) is not taken care anymore by him, by Tabran’

(3) Sape dame-nye?
    who name-3
    ‘What was his name?’
However, it is important to note that the distribution of cognate forms of -nya in Malay languages is not uniform. For example, Ewing (1995: 237) reports that -nya in Colloquial Indonesian marks definiteness, (associative) possession, and gerunds. Similarly, Engelbreton (2003) reports that -nya acts as a possessive marker, identifiability marker, nominalizer, pronominal marker, and adverbial marker (153-171). For definiteness, -nya marks NPs that are identifiable in the discourse. For possession, -nya acts a ligature between the possessed and possessor (i.e., POSSESSED-nya POSSESSOR). For nominalization, -nya forms gerunds from verbal elements. In Besemah, however, -nye does not mark this type of (associative) possession nor does it form gerunds; it does appear to mark definiteness, which is demonstrated in (4).

(4) Puntung dindak mbagal **sumpit**.
firewood not.want av-hit picher.bug
‘The firewood didn’t want to hit a **pincher bug**.’

... 
Njadi puntung tadi,
so firewood just.now
‘So, the firewood just now,’
.. ude kate-nye mane sumpit-(ny)e?
whatever say-3 where picher.bug-3
‘.. that’s it, he said, where’s the **pincher bug**?’

This example from a narrative shows that **sumpit** ‘pincher bug’, suffixed with -nye (in the third line) has already been introduced earlier in the narrative (in the first line) and is thus clearly identifiable.

From the examples above, it is fairly evident that this Malay suffix has undergone semantic generalization. To my knowledge, however, no one has proposed a specific grammaticalization pathway for the Malay -nya, and to do so here would be beyond the scope of this study. What is clear is that the Malay -nya began as a third person genitive pronoun (see below) and underwent a number of semantic and morphosyntactic changes. The remainder of the study will focus on the morphological and phonological structure of -nye from synchronic and diachronic perspectives. The semantics and pragmatics of -nye are not affected by the alternation of -nye and anye; rather the alternation results primarily from morphophonemic and prosodic factors. It is important to note, however, that -nye has undergone considerable grammaticalization with its multitude of meanings, pragmatic functions, and syntactic distributions.

Unlike its cognates in other well-known Malay languages, the Besemah suffix -nye [ŋa] is phonologically conditioned; if the root to which the -nye is attached

---

3 In other varieties of Malay-Indonesian (including Besemah), nominal possession is simply marked by the juxtaposition of the possessed NP before the possessor NP.
ends in a consonant, the initial palatal nasal [ɲ] is lost as in (5a), but if the root ends in a vowel, the initial full form is realized as in (5b).

(5) a. batak [bataʔ] + -nə [+nə] → batake [bataʔnə] ‘brought by them’

Aside from the obvious reduction, when the root ends in a consonant as in example (5a), the final consonant undergoes gemination, a process that occurs with other vowel initial suffixes, such as the locative applicative -i and the nominalizer -an. Gemination most likely occurs across syllable boundaries to preserve the syllable structure of the root (i.e., the final closed syllable). In the examples written orthographically below, gemination is omitted.

The reduction of -nə > -e described thus far is not too surprising for two reasons. The first reason is that similar processes occur in other Malay varieties of Sumatra: Southern Barisan Malay (personal knowledge), Tanjung Raden Malay and Mudung Darat Malay in Jambi in central eastern Sumatra (Yanti 2010: 507-510), and Minangkabau in western Sumatra (Moussay 1981). However, because the Malayic languages of western Indonesia have not received much description, little is known about the extent and nature of these processes. Second, from a grammaticalization perspective, these data are not surprising because the loss of the initial segment of the suffix allows the root and suffix to form a tighter bond. This becomes even clearer in the next subsection once the etymology of -nə in Proto-Malayic and Proto-Austronesian is outlined, which demonstrates that this morphophonemic process is in fact phonetic erosion in progress.

When -nə occurs after the suffix -an or circumfix peN- -an, a third form surfaces. Consider the example in (6) below.

(6) a. langkah-an [laŋkahhan] ‘stride’
   langkah anye [(laŋkahkan)ə (aŋə)ə] ‘her stride’
   b. pe-langkah-an [pəlaŋkahhan] ‘threshold’
   pe-langkah-an anye [(pəlaŋkahkan)ə (aŋə)ə] ‘their bathing place’

In (6a), the nominalizer -an is suffixed to the root langkah ‘to step’, which results in the noun langkah-an ‘stride’. From the example above one might expect that -e will surface, resulting in the form langkah-an-e [laŋkahhamə] ‘her stride’, but the form anye surfaces instead, exemplified in the second line of (6a). What might be even more unexpected is the fact that anye appears to be a phonological word in its own right, denoted by the (...)ə, surrounding each phonological word. The same pattern occurs with the locative nominalizing circumfix peN-an, which means something like ‘the place of [verb]’. In (6b), the same root combines with the circumfix, resulting in pe-langkah-an ‘threshold.’ Subsequently, when the third person formative attaches to pe-langkah-an, the resulting form of -nə is anye.
From the data presented thus far, one might postulate that the internally reconstructed form of the third person suffix is \(*anye\). The synchronic description of -\(\text{n}\)ye, then, would simply state that there is a suffix-word alternation in the third person suffix in Besemah. This process, of course, is common (i.e., English \(\text{will}/\text{ll}, \text{him}/-(\text{e})\text{m}, \text{them}/-(\text{e})\text{m}\)). However, reconstructions of this form in Proto-Malayic (PM), Proto-Malayo-Polynesian (PMP), and Proto-Austronesian (PAn) show that this is not the case. It is clear from reconstructions by Adelaar (1992) for PM, Blust (1977a) and Ross (2006) for PMP and PAn that \(\text{anye}\) is not the older form. The etymology of the Besemah -\(\text{n}\)ye is presented in (7) below.

(7) Besemah -\(\text{n}\)ye < PM *\(-\text{n}\)a < PMP *\(\text{n}\)i-a (*\(\text{n}\)iya) < PAn *\(\text{n}\)(i)-\(\text{i}\)a or *\(\text{n}\)-\(\text{i}\)-a

Working from the Besemah reflex -\(\text{n}\)ye, the PM form *\(-\text{n}\)a was reconstructed straightforwardly by Adelaar (1992: 125) on the basis of six Malayic isolects from Java, Sumatra, and Borneo. Adelaar proposes that *\(-\text{n}\)a results from the merger of *\(\text{n}\) and *\(\text{i}\) in the PMP form *\(\text{n}\)iya, which was reconstructed by Blust (1977: 10-11) and Ross (2006: 530-531, 536-537). This etymology is controversial only at the level of PAn, where Blust (1977) reconstructed the genitive phrase marker *\(\text{n}\)i- and the third person pronoun *\(-\text{a}\), and Ross (2006) reconstructed the genitive phrase marker *\(\text{n}\)(i)- with the third person pronoun *\(-\text{i}\)a. However, this does not affect the analysis of -\(\text{n}\)ye here. What is most important is the fact that there is no possible etymology for *\(\text{anye}\). How then can the change from -\(\text{n}\)ye to \(\text{anye}\) be explained? The unexpected answer is that Besemah has a one-suffix-per-word preference, which is motivation for -\(\text{n}\)ye to make the leap from a dependent formative to an independent word. The next section further explains this claim, providing evidence for the unlikely change -\(\text{n}\)ye > \(\text{anye}\).

### 4 A Diachronic Explanation

A survey of probable explanations of the unexpected shift from -\(\text{n}\)ye > \(\text{anye}\) leads nowhere. For example, one expected source would be the fusion of -\(\text{n}\)ye with another meaningful formative \(-\text{a}\) (or something even more phonetically robust). Although it is difficult to argue for the absence of a morpheme, such a formative does not appear to exist in the history of Malay(ic) languages. Furthermore, it is also unlikely that any element would intervene between the suffix -\(\text{an}\) and the third person suffix, which does not appear to occur in any Malay(ic) language that I am aware of. Another explanation may be a synchronic phonological account. That is, the morphophonemics of -\(\text{n}\)ye above show that there is a dispreference for consonant-nasal sequences across morpheme boundaries. In the case of (7) above, the \(-\text{a}\) would serve as an epenthetic vowel to break up a consonant-nasal sequence. Again this is implausible for two reasons. First, in Besemah, it is schwa and not /\(\text{a}\)/ that is used to break up clusters of consonants. Second, this explanation would not explain why \(\text{anye}\) appears to occur as a separate phonological word (see discussion below). With no straightforward explanation
Bradley McDonnell

for the -nye/anye alternation, there is a rather unexpected explanation: Besemah prefers one suffix per word. The remainder of this section defends this admittedly unusual assertion drawing on data from a corpus of spontaneous speech.

4.1 The Prosodic Status of anye

Thus far, anye has been assumed to be an independent phonological word. However, in order to provide clear evidence in support of the one-suffix-per-word preference, it is necessary to show that anye is a prosodic word in its own right. From the outset, anye is suspicious as a suffix for a number of reasons. First, anye is disyllabic, and no other bound affix/clitic in Besemah is disyllabic. In fact, anye is homophonous with a word that is clearly phonologically independent, namely the adversative anye ‘but’ (cf. Standard Indonesian hanya ‘only’). Second, as was mentioned briefly in section 3, vowel-initial suffixes trigger gemination of the consonants preceding them, but in this construction the final /n/ in the suffix -an does not geminate before anye. The best evidence for the prosodic status of anye should come from stress patterns in Besemah. However, stress in Besemah is not straightforward. Preliminary evidence suggests that stress falls on the final syllable and is cued by intensity and less reliably pitch (McDonnell 2013). Impressionistic judgments support the recognition of anye as a separate phonological word based on the stress patterns demonstrated in (8).

(8) a. tulis [tu’IIs] ‘write’
   b. tulisan [tulis’san] ‘writing’
   c. tulisan anye [tulis’san a’ɲə] ‘her writing’

As noted in section 2, stress falls on the final syllable of words whether they contain a suffix or not. So, in (8a) stress appears to fall on the final syllable of the monomorphemic word, and in (8b) on the final syllable of the suffixed root. However, when anye follows the affixed root in (8c), both the final syllable of anye and tulisan are stressed. It is important to note, however, that anye is still a grammatical formative that is syntactically bound to the deverbal noun that precedes it. As such, it is hard to imagine cases where it would receive additional prosodic prominence, such as contrastive focus. This also means that anye does not occur in a separate intonational unit from the noun before it.

4.2 The -nye/anye Alternation in Discourse

From the morphophonemic, morphosyntactic, and prosodic properties of anye described thus far, it may or may not be surprising that in spontaneous speech speakers do not produce anye one hundred percent of the time. Instead speakers occasionally produce the -e form after the suffix -an. In a 60,000-word corpus of conversations and narratives in Besemah, there were a total of 51 occurrences of -nye after roots with the suffix -an or the circumfix peN- -an. In only eight
cases did the form -e surface. In the remaining 43 cases, anye surfaced. Of the 15 speakers who used a form of -nye after -an or peN- -an, only six speakers used -e; the other nine speakers exclusively used anye. Interestingly, five of these six speakers who used -e were younger than 30 years old, which suggests that anye may be disappearing with the younger generation. These data show that, although both -e and anye occur after an affixed word, the best predictor of variation is age and not some other morphosyntactic, pragmatic, or phonological phenomenon.

4.3 Diachronic motivations for -nye > anye

What is particularly strange about the change -nye > anye is the addition of phonetic material (a-), which has no clear origin in other formatives (as was shown above) and no clear pragmatic or syntactic motivation from the corpus. Without any other motivating factors, can the one-suffix-per-word preference in Besemah adequately account for the change -nye > anye?

If such a preference exists in Besemah, one might expect a number of possible consequences. One might expect speakers to use the full third person pronoun die instead of the suffix -nye. In the same corpus of spontaneous speech, only one possible example of this was found, shown here in (9).

(9) Anu die tu nak n-damping-i peng-gawih-an die, 
    uhm 3 that FUT AV-close-APPL NMLZR-work-NMLZR 3
    ‘Uhm they wanted to be close to (their) work,’

In Besemah, as in many Malay languages, possession can be marked either by bound enclitics/suffixes (=ku, -nye) or by free pronouns (aku, die) immediately following the possessed NP (e.g., NP=ku or NP aku). It is possible then that the die at the end of the example in (9) is the possessor. However, this example is in fact ambiguous; die could also be considered an echoed subject of the active voice clause, a fairly common phenomenon in spontaneous speech in Besemah. One reason that speakers may not choose this option is because the free pronoun die is only available for possession and not definiteness.

While the third person does not rely on this strategy, there is evidence from the first person enclitic-word alternation =ku/aku, which does not suffer from the same semantic incompatibility that die and anye do. The evidence suggests that speakers do in fact use the strategy discussed above by employing the full pronoun aku after an already suffixed word. This is a useful diagnostic for the one-suffix-per-word preference in Besemah, which would predict that =ku does not occur after an already suffixed word. In the same corpus, the first person possessive occurred after a root suffixed with -an or circumfixed with peN- -an a total of 27 times, while it occurred after unsuffixed words 96 times.
The table in (10) presents the number of occurrences of the free form *aku* and the bound form =*ku*. It is readily apparent that speakers prefer =*ku* after unsuffixed forms and *aku* after roots that are already suffixed or circumfixed.

<table>
<thead>
<tr>
<th></th>
<th><em>aku</em></th>
<th>=<em>ku</em></th>
<th></th>
<th><em>aku</em></th>
<th>=<em>ku</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsuffixed</td>
<td>8</td>
<td>88</td>
<td>Suffixed</td>
<td>23</td>
<td>4</td>
</tr>
</tbody>
</table>

From this pattern, it appears that Besemah speakers avoid the clitic pronoun when the lexical item is already suffixed. However, the third person free form *die* is not semantically similar enough to -*nye* to be an appropriate substitute. Therefore, it is probable that the one-suffix-per-word preference is met because *anye* is a separate prosodic word, just as in the case of *aku*. This provides good evidence for the one-suffix-per-word preference and, even further, it provides motivation for the change of -*nye* from a bound suffix to free form. Now that there is a motivating force behind the development of a bound form (-*nye*) to an autonomous form (*anye*), the question becomes: what motivates the epenthesis of *a*- in *anye*?

Under pressure from the one-suffix-per-word preference, it is likely that the motivation to append *a*- to -*nye* was to meet a bimoraic minimal word requirement (see section 2). Cross-linguistically, this is a quite common synchronic process. Applebaum and Gordon (2010) lay out a typology of languages that use processes such as consonant epenthesis (i.e., Cupeño), vowel lengthening (i.e., Northern Sámi), and most importantly here, vowel epenthesis (i.e., Minto) to meet minimal word requirements. It is then conceivable that historically, -*nye* did not satisfy the minimal word requirement, which was then resolved by the epenthetic vowel *a*-.

Where, then, does the *a*- come from? One might posit that this vowel epenthesizes on analogy from the first person singular *aku*. The similar phonological shape of *anye* and *aku*, and the fact that they are pronouns with a reduced and full form provide motivation for this analysis.

Without historical records, it is impossible to provide undeniable evidence of this diachronic process. However, there is clear evidence from (1) the distribution of *anye* and *aku* after suffixed words in the corpus and (2) the prosodic requirements of phonological words to motivate the change -*nye* > *anye*. Therefore, the driving force behind this unusual change is the preference in Besemah for one suffix per word. As it turns out, the one-suffix-per-word preference that motivates this unusual change actually surfaces in other constructions in the language.

5 The Causative Suffix -*kah* and the One-Suffix-Per-Word Preference

Aside from the examples where the Besemah pronominal formatives attach to deverbal nouns, -*nye* attaches to verbs as either agents in the passive voice (1) or patients in the active transitive voice (2). In these two constructions, -*nye* can occur after either the causative/applicative suffix -*kah* or the locative-applicative
Roadblocks in the Grammaticalization Highway

suffix -i. When -nye occurs after the locative-applicative suffix -i, it forms a single phonological word with the root, suffix -i, and -nye shown in (11).

(11) a. di-batak-i [(dibataʔi)ᵣ] ‘(it) was brought (somewhere)’
    b. di-batak-i-nye [(dibataʔiŋa)ᵣ] ‘(it) was brought (somewhere) by her’

In (11a), the root batak ‘bring’ attaches to the passive prefix di- and the applicative suffix -i. This is uncontroversially a phonological word in Besemah. This is also true of the example in (11b), where -nye is suffixed after -i. Consequently, the prosodic status of dibatakanye in (11b) challenges the one-suffix-per-word preference presented thus far. However, there may be good reason that speakers allow (11b) to be a phonological word. First, -i is phonologically ‘light’ as a high vowel with no onset or coda. Secondly, -i is historically older than suffixes like -kah and is therefore more closely bound to the root. This is precisely why the one-suffix-per-word preference is best considered a preference rather than a constraint or a rule.

What is more interesting is the causative/applicative construction. When -nye follows the causative/applicative suffix -kah in (12), it patterns quite differently.

(12) a. di-batak-kah [(dibataʔkah)ᵣ] ‘(it) was brought (for someone)’
    b. di-batak-kah-nye [(dibataʔ)ᵣ (kaŋa)ᵣ] ‘(it) was brought (for someone) by her’

In (12a), batak ‘bring’ combines with the passive prefix di- and the suffix -kah, resulting in dibatakakah ‘bring for’, which again is uncontroversially a phonological word. When the agentive -nye is suffixed in (12b), one might expect from the morphophonemic patterns in section 3, either -e after the word-final consonant /h/ or anye after the suffix -kah, resulting in batakahē [(bataʔkahē)ᵣ] or batakah anye [(bataʔkah)ᵣ (aŋa)ᵣ]. Instead, the suffixes -kah and -nye fuse together to form an autonomous phonological word, kanye. This shift from two bound suffixes to a single autonomous prosodic unit is quite unexpected, but as in the case of anye can be motivated by the one-suffix-per-word preference in Besemah. That is, section 4 showed that -nye cannot be suffixed to an already suffixed root. This, in turn, forced -nye to occur on its own, resulting in the autonomous form anye. In the case of kanye, instead of the suffix -nye being forced to occur on its own, the one-suffix-per-word preference and the minimal word constraint are satisfied by the combination of the two suffixes. In both cases the one-suffix-per-word preference is resolved. Even though the same motivating force is at work, the development of kanye is actually quite different from that of anye. Most importantly, kanye developed from the fusion of two suffixes, while anye developed from the epenthesis of a- to fulfill prosodic word requirements.

The history of -kah is more controversial than that of -nye presented in section 3, with two potential sources. According to Adelaar (1992: 149), the two probable
sources for -kah in Besemah are either the preposition *ka ‘to’ or the preposition *akan ‘about, concerning’. There are a number of reasons for adopting either etymology, and neither etymology has more explanatory power than the other for kanye. Let us first consider *akan. While the change *akan > -kan is found in many Malay varieties, including Standard Indonesian, the change from -kan > -kah is motivated by a number other instances of -n > -h. This change occurred in Besemah, but not Standard Indonesian (e.g., *dengan ‘companion’ in Old Malay > dengah 2SG in Besemah, dengan ‘with’ in Standard Indonesian). Alternatively, it is possible that the suffix -kah in Besemah is derived from the PM allative preposition *ka. The only issue that needs to be motivated is h-epenthesis to -kah in Besemah. Incidentally, this type of process occurs with modern borrowings from other varieties of Malay-Indonesian, where -h is appended to a word that ends in -a. For example, the Indonesian word desa is appended with a word-final -h to preserve the final -a. The -a is preserved because all other cases of word-final -a went to -ə in Besemah (i.e., *-a > -ə, Tadmor 2002). The choice of either of these etymologies has an interesting impact on the analysis of kanye, which is discussed in the next section.

6 Possible Pathways of Change

Thus far, there has been good evidence to support the notion that the one-suffix-per-word preference in Besemah led to two unexpected developments in Besemah, anye < -nye and kanye < *akan + -nye or < *ka + -nye. Both cases involve what might be deemed degrammaticalization, where autonomous forms appear to have developed from previously bound forms. Even though it is not possible to trace these changes through history, it is worth discussing probable and improbable grammaticalization pathways, focusing on areas where the one-suffix-per-word preference blocked the grammaticalization processes. It is likely that anye developed directly from -nye (or PM *-nya) and not from -e. As PM *-nya became more and more tightly bound to the root in Besemah, it is likely that the one-suffix-per-word preference disallowed the -nye from attaching. The minimal word requirement was then resolved by epenthesis of a-. This means that while -nye was in the process of becoming a bound suffix, anye most likely developed from -nye when it was still an enclitic. It is improbable that anye developed from the tightly bound form -e. From this diachronic perspective, the leap from bound to autonomous form was not so big (i.e., enclitic > word). One probable explanation for this change is that it resulted from two competing pressures: (1) the grammaticalization pressure to compact enclitics into suffixes and (2) the language specific pressure to prefer only one suffix per word.

In the case of the fused causative/applicative -kah and third person -nye suffixes (i.e., kanye), the grammaticalization pathway in many ways depends on the reconstruction of -kah. If *akan is the reconstructed form, the expected cline is *akan > *-kan > -kah. If this were the case, the fusion of -nye to -kah most
likely came after -kah had already become a causative suffix. The grammaticalization pathway of kanye is shown in (13) below.

(13) ROOT-ka{h,n}-nye > ROOT-kanye (or ROOT-ka{h,n})-nye > ROOT kanye

However, if the source of -kah is the allative pronoun *ka, it is possible that -nye fused with *ka prior to -h epenthesis (i.e., ROOT ka-nye > ROOT kanye) and prior to it attaching to the root (i.e., ROOT ka > ROOT-kah). Thus, it is likely that -nye first attached to the autonomous preposition *ka. Subsequently, *ka became dependent on the root and appended -h. If this is indeed the case, the one-suffix-per-word preference in Besemah prevented kanye from attaching to the root. It0 may also be the case that the one-suffix-per-word preference did not interfere here, but kanye was already a frozen form, when -kah developed into a suffix.

This discussion demonstrates a number of important points concerning these apparent cases of degrammaticalization in Besemah. First, language specific constraints or preferences can impact the gradual processes of grammaticalization. For example, the development of anye resulted from the one-suffix-per-word preference as -nye was steadily moving down the grammaticalization cline. Second, what are on the surface seemingly drastic cases of degrammaticalization may emerge from ‘roadblocks’ in the grammaticalization cline that result from more minor changes.

7 Conclusion

This study of the suffix -nye in Besemah has focused on explaining rather than categorizing the unusual development of phonologically autonomous forms from previously bound forms. Although many of the debates over examples of degrammaticalization have focused on challenging the unidirectionality hypothesis, the present study demonstrates that these cases of apparent degrammaticalization can be accounted for by combining language specific characteristics and the principles of grammaticalization. As Heine (2003) notes ‘most of the counterexamples [to unidirectionality] … can be described as being “idiosyncratic” in the sense that they do not allow for cross-linguistic generalizations on the directionality in the rise and development of grammatical categories’ (582). He goes on to say that cases of degrammaticalization are also idiosyncratic language internally in the sense that ‘it involves isolated instances within a given language’ (582). This does not mean, however, that cases of degrammaticalization are uninteresting. On the contrary, the unusual changes in Besemah point to an interesting characteristic of the language, a preference that allows only one suffix per word. It also points to an interesting interaction between grammaticalization and other language internal processes.
References


Bradley McDonnell
University of California, Santa Barbara
Department of Linguistics
South Hall 3432
Santa Barbara, CA 93106-3100

bradley.mcdonnell@gmail.com