

Squib

When Sound Change Obscures Morphosyntax: Insights from Seediq

Victoria Chen

VICTORIA UNIVERSITY OF WELLINGTON

In memory of Dakis Pawan

The Formosan language Seediq displays an understudied case of morphological opacity, where a single phonological innovation has resulted in the syncretism of five Proto-Austronesian functional affixes in affirmative declaratives. How and why these affixes remain functionally intact in modern Seediq has important implications for understanding the tension and interplay between semantic transparency and morphological opacity. In this squib, I demonstrate that the marginal overlap of these affixes' lexical subcategorization may have reduced obstacles to learnability and processing, enabling them to remain functionally distinct despite the absence of morphological distinctions. The case of Seediq therefore highlights the often-neglected fact that sound change-induced morphological opacity may obscure but not necessarily obliterate syntax.

Keywords: Seediq; Morpho-Phonological Interface; Proto-Austronesian; Morphological Change; Atayalic

1. INTRODUCTION.¹ In principle, language, as a tool of communication, is most effective when meaning is conveyed through one-to-one correspondence with form (i.e., morpheme). However, it has been well documented that phonological innovations may obscure morphosyntax, giving rise to ambiguity or discontinuity in morphology (see, e.g., [Baerman, Brown, and Corbett 2005](#); [Baerman 2007](#); [Tammimga 2013](#); [Bybee 2017](#); [Mckinnon et al. 2018](#)). Whether or not the need to maintain grammatical intelligibility blocks regular sound change has remained an ongoing question over the past few decades ([Melchert 1975](#); [Blevins 2004](#); [Hill 2014](#); a.o.). This squib investigates two questions related to this theme:

1. This article is dedicated to the memory of Dakis Pawan, a highly respected language consultant who passed away in November 2021. Dakis made an incredible contribution to the documentation of the Seediq language, and his expertise and passion for language were unparalleled. As his student and someone who had the privilege of working with him, I am grateful for the time he spent with me, sharing his knowledge and insights. Dakis was a true inspiration, and his legacy will continue to live on through the countless individuals whose lives he touched. May his memory be a blessing and an inspiration to all those who knew him.

- (1) a. To what extent can phonological innovation obscure morphosyntax?
 b. Where sound changes are completely regular and immune to grammatical factors, does the elimination of morphological distinction entail the loss of grammatical function?

I explore these questions through the lens of Seediq (ISO 639-3 *trv*), a first-order language of the Atayalic primary branch of the Austronesian family. Synchronically, Seediq possesses an affix *m-* that serves a wide range of functions. On some occasions, it combines with dynamic verbs and functions like a run-of-the-mill Actor Voice allomorph, as in (2a). It may also combine with stative stems and functions as a stative affix, as in (2b). It may also mark nominal stems and serve as a possessive-like marker denoting the meaning of ‘to wear’, ‘to carry’, or ‘to have’, as in (2c).²

(2) SEEDIQ³

- a. Gaga **[m]**-ege Ø pila Ø dakis ka robo.
 PROG AV-give ACC money ACC Dakis PIVOT Robo
 ‘Robo is handing money to Dakis now.’ (*m-* as an Actor Voice affix)
- b. Nii ku **[m]**-ure riyung.
 PROG 1SG.PIVOT STAT-hungry very
 ‘I am very hungry.’ (ODFL) (*m-* as a stative affix)
- c. Dangi na swai mu ga **[m]**-lukus qalux ga.
 friend POSS younger.brother 1SG.POSS that POSS-clothes black that
 ‘The person in black is my younger brother’s friend.’
 (*m-* as a possessive affix)

This affix may also function like an opaque detransitivizer. When attached to certain two-place verbs, it obligatorily introduces a one-place agentless construction (3a). This contrasts with this affix’s apparent allomorphic infix *<m>*, which, when attached to the same stem, introduces a two-place construction, (3b).

(3) SEEDIQ

- a. **[M]**-takur ka Robo di.
 M-trip PIVOT Robo PRE.PART
 ‘Robo tripped.’ (*m-* as an opaque detransitivizer)
- b. T**[<m>**]akur Ø Walis *(ka Temi).
 <AV>trip CM₁ Walis *(PIVOT Temi)
 ‘Temi tripped Walis.’ (<*m*>-marked two-place construction)

Besides these four uses, *m-* serves at least two other functions in Seediq, to be discussed later in this paper. A better understanding of how its multifunctionality arises and whether that has impacted the learnability of these affixes

2. List of abbreviations: AV: Actor Voice; CM: case marker; COMT: comitative; DEF: definite; DETR: detransitivizer; INDF: indefinite; LK: linker; LOC: locative; OBL: oblique; PROG: progressive; POSS: possessive; PART: particle; PRF: perfective; PROJ: projective; REC: reciprocal; RED: reduplication; STAT: stative.

3. The verb *mure* was originally written as *muure* in ODFL. Following an anonymous reviewer’s suggestion, I have pre sented it as *mure* (with *ure* as the stem, as reported exactly in the dictionary). The first *u* is likely to be an epenthetic vowel with no phonemic value that was accidentally transcribed.

would enable a better understanding of the questions in (1). In this squib, I argue that this morphological opacity is “affordable,” as the affixes impacted by the merger have little overlap in their lexical subcategorization. Their syncretism has therefore not triggered major issues in learnability and processing. The case of Seediq thus highlights the often-neglected fact that sound change-induced morphological opacity may obscure but not necessarily obliterate syntax.

This squib is organized as follows: Section 2 introduces the basic function and etymology of the five affixes, as well as evidence that they have remained functionally distinct in modern Seediq. Section 3 discusses the distribution of these affixes and their implications for the questions in (1). The Seediq data presented in this paper comes from the following sources: the Austronesian Comparative Dictionary (Blust and Trussel, *ongoing*), the Online Dictionary of Formosan Languages (ODFL), and Holmer (1996), as well as the author’s primary fieldwork on Tgdaya Seediq.

2. FIVE FUNCTIONS OF *m-* IN SEEDIQ AND THEIR DIACHRONIC SOURCES.

Seediq is a severely endangered language spoken in central Taiwan with less than 10,000 speakers (Sung 2018). It displays a vowel neutralization process that affects all prestress (pre-penultimate) syllables (Yang 1976; Li 1991; Holmer 1996; Tsukida 2009; Ochiai 2016; Sung 2018). In Tgdaya Seediq, pre-penultimate vowels are consistently realized (pronounced) as the high back vowel [u], except when separated from the stressed vowel by an /h/ or a glottal stop. Such vowels are omitted in Seediq orthography as their phonetic value is robust and predictable. This is illustrated in (4).

	<i>Phonetic Form</i>	<i>Seediq Orthography</i>	<i>Gloss</i>
a.	[runge <i>di</i>]	rngedi	‘fly’
b.	[hun <i>gedan</i>]	hngedan	‘to cook’
c.	[dun <i>gese</i>]	dn’gese	‘to wrap (in weaving)’
d.	[tu <i>makur</i>]	tmakur	‘to slip’
e.	[su <i>numalu</i>]	snmalu	‘to have made’

This phenomenon is well documented in the literature. Some researchers describe it as vowel neutralization (Yang 1976; Holmer 1996), and others analyze it as vowel deletion followed by the insertion of an epenthetic vowel (Li 1977, 1991). I adopt the former approach and schematize this change as vowel neutralization (5).

(5) $V \rightarrow u_CV$ (Yang 1976:661)

This rule consistently applies to Proto-Austronesian (PAN) reflexes in Seediq. As a result, various PAN prefixes in the form of CV- surface as C- in Seediq. As stems in Seediq are usually disyllabic or trisyllabic, prefixes are usually in pre-penultimate positions, subject to the rule in (5). Consider the examples in (6).

	<i>PAN Form</i>	<i>Reflex in Seediq</i>	<i>Function</i>
(6) a.	*Si-/Sa-	s-	circumstantial voice prefix
b.	*pa-	p-	causative prefix
c.	*pa-ka-	p-k-	causative of abilitative
d.	*ma-ka-	m-k-	abilitative
e.	*Ca-reduplication	C-reduplication	reduplication for plurality

This rule predicts that PAN prefixes that contrast only in vocalic value would be morphologically undistinguished in Seediq. One class of such examples are five affixes that share the same onset *m*: PAN **m*- (Actor Voice), **ma*- (AV stative), **mi*- (AV possessive), **mu*- (AV motion or detransitive), and **maCa*- (AV reciprocal). Below, I scrutinize the synchronic use of each affix in Seediq.

2.1. REFLEX OF PAN ACTOR VOICE AFFIX *<um> IN SEEDIQ. The PAN Actor Voice (AV) affix *<um> is a verbal affix attached to dynamic verbs. It displays an allomorphy rule commonly found in higher-order Austronesian languages: when attached to vowel-initial stems, it surfaces as a prefix *m*-; when attached to consonant-initial stems, it surfaces as an infix <*m*>. This rule is schematized in (7) and illustrated with the Seediq examples in (8).

(7) PAN Actor Voice infix *<um> → {*m*- on V-initial stems}

(8) SEEDIQ

	<i>Allomorph</i>	<i>Environment</i>	<i>Example</i>	<i>Gloss</i>
a.	< <i>m</i> >	/ _ C	c < <i>m</i> >eka	‘to hack-AV’
			g < <i>m</i> >eeguy	‘to steal-AV’
			k < <i>m</i> >etuy	‘to pick-AV’
			q < <i>m</i> >teqic	‘to bite, to nibble-AV’
			t < <i>m</i> >akur	‘to slip-AV’
b.	<i>m</i> -	/ _ V	m -aduk	‘to let leave-AV’
			m -eyah	‘to come-AV’
			m -urux	‘to do alone-AV’
			m -ilo	‘to jump-AV’

Given this rule, *m*- is conventionally described as an AV allomorph in the Seediq literature (Holmer 1996; Chang 2000; Sung 2018). The examples in (9) illustrate this use. In all three sentences, *m*- appears on a verbal stem that has a vocalic onset.

(9) SEEDIQ

- a. Wada **m**-uyas Ø saph pyasan laqi-mu.
 PRF AV-sing LOC school child-1SG.POSS
 ‘My child went to school to sing.’ (ODFL) (Actor Voice *m*-)
- b. **m**<*n*>alix Ø parih nii ka Pawan.
 AV<PRF>abandon ACC hoe this PIVOT Pawan
 ‘Pawan abandoned this hoe.’ (Actor Voice *m*-)

- c. \boxed{M} -eyah m-imah \emptyset qsiya \emptyset ruru ka boyak.
 AV-come AV-drink ACC water LOC stream PIVOT boar
 ‘Boars often come to the river to drink water.’ (ODFL) (Actor Voice *m*-)

However, a closer look at the distribution of *m*- reveals that the PAN affix **<um>* is not the only diachronic source of this affix. Below I discuss four other uses of this morpheme and their etymology.

2.2. REFLEX OF PAN STATIVE PREFIX *ma- IN SEEDIQ. PAN possessed a distinct affix **ma-* that marks stative verbs (Blust 2013 [2009]; Ross 2015). Consider the examples in (10).

- (10) PAN stative prefix **ma-*

<i>PAN Form</i>	<i>Examples</i>	<i>Gloss</i>
* <i>ma-</i> ‘stative prefix’	* <i>ma-demdem</i>	‘dark’
	* <i>ma-dalis</i>	‘smooth, slippery’
	* <i>ma-keseR</i>	‘strong’
	* <i>ma-Niwang</i>	‘thin’
	* <i>ma-quSaw</i>	‘thirsty’
	* <i>ma-qataq</i>	‘raw, unripe’

Given the vowel neutralization rule in (5), this affix is expected to surface only as a bilabial prefix *m-* in Seediq. This prediction is borne out by a series of *m*-marked stative/adjectival verbs in the language, as in (11).

- (11) Reflex of PAN **ma-* in Seediq

	<i>Reflex in Seediq</i>	<i>Gloss</i>
<i>m-</i> ‘stative’ < PAN * <i>ma-</i>	m-basi	‘sour’
	m-damac	‘excited’
	m-gigi	‘short’
	m-budu	‘blind’
	m-sekuy	‘cold’
	m-tbeno	‘strong’

Truncation/neutralization of the vowel /a/ has resulted in the absence of a formal distinction between the stative affix and the AV prefix *m-*. As seen below, stative constructions (12a) in Seediq are morphologically undistinguished from sentences formed with AV-marked dynamic verbs (12b), where the verb has a vocalic onset.

- (12) SEEDIQ

- a. Nii ku \boxed{m} -ure riyung.
 PROG 1SG.PIVOT STAT-hungry very
 ‘I am very hungry.’ (ODFL) (stative construction)

- b. Wada **m**-uyas Ø sapah pyasan laqi-mu.
 PRF AV-sing LOC school child-1SG.POSS
 ‘My child sang at school.’ (ODFL)

(AV construction with a vocalic onset)

This ambiguity does not exist in the majority of higher-order Austronesian languages, where these two constructions are marked by two distinct markers, as seen in (13) and (14).

(13) PUYUMA

- a. **Ma**-ebut na lrawlaw.
 STAT-extinguished DEF.PIVOT light
 ‘The light is extinguished.’ (ODFL) (stative construction)

- b. **M**-alrup dra marenem i inbu.
 AV-hunt INDEF.AC sambar SG.PIVOT Inbu
 ‘Inbu hunted sambar.’ (ODFL) (AV clause with a vocalic onset)

(14) PAIWAN

- a. **Ma**-cula=sun manu **ma**-zeLi=sun
 STAT-hungry=2SG.PIVOT or STAT-weary=2SG.PIVOT
 ‘Are you hungry or are you tired?’ (Chang 2006:273)
 (stative construction)

- b. **M**-aLap timadju ta za taki nimadju ka=ta za
 AV-take 3SG.PIVOT OBL.CM that aboriginal.knife 3SG.GEN COMT=OBL.CM that
 turivecan.
 tool
 ‘He took his aboriginal knife together with those tools.’
 (Chang 2006:180) (AV construction with a vocalic onset)

2.3. REFLEX OF PAN POSSESSIVE PREFIX *mi- IN SEEDIQ. Yet another PAN affix with the form of *mV-* is the possessive affix *mi-. This affix combines with a nominal stem and denotes the meaning of ‘to have X’ (or ‘to wear X’, depending on the context), as seen in the data below from Puyuma and Yami/Tao (Rau and Dong 2006; Teng 2008, 2014; ODFL).

		<i>Stem</i>	<i>Gloss</i>	<i>mi-form</i>	<i>Gloss</i>
(15)	a. Puyuma	kiping	‘clothes’	mi -kiping	‘to wear clothes’
		paysu	‘money’	mi -paysu	‘to have money’
		ruma’	‘house’	mi -ruma’	‘to have a house’
		walak	‘child’	mi -walak	‘to have children’
		kaput	‘peer’	mi -kaput	‘to have peers’
	b. Yami	ayob	‘clothes’	mi -ayob	‘to wear clothes’
		sakop	‘hat’	mi -sakop	‘to wear a hat’
		ipos	‘tail’	mi -ipos	‘to have a tail’
		panid	‘wing’	mi -pani-panid	‘to have wings’
		isis	‘scale’	mi -isis	‘to have fish scales’

As expected from the vowel truncation rule (5), the possessive affix *mi- is reflected as *m-* in Seediq. This is illustrated in (16) with three cognate constructions from Seediq, Puyuma, and Yami (Tao).

- (16) a. SEEDIQ
 Dangi na swai mu ga **mi**-lukus qalux ga
 friend POSS younger.brother 1SG.POSS that M-clothes black that
 ‘The person in black is my younger brother’s friend.’ (ODFL)
 (*m*-N as ‘to have N’)
- b. PUYUMA
 Adri=ku sagar **mi**-kiping dra meneene.
 NEG=1SG.PIVOT like.AV POSS-clothes INDF.ACC Tight
 ‘I don’t like to wear tight clothes.’ (ODFL) (*mi*-N as ‘to have N’)
- c. Yami (Tao)
mi-ayob namen so cininon.
 POSS-clothes 1PL.PIVOT ACC traditional.costume
 ‘We put on traditional costumes (to celebrate a new-born).’ (ODFL)
 (*mi*-N as ‘to have N’)

2.4. REFLEX OF PAN DETRANSITIVE PREFIX *mu- IN SEEDIQ

Given the rule in (5), the PAN detransitivizer *mu- also surfaces as *m-* in Seediq. This morpheme functions as a valency-decreasing affix that denotes a one-place construction with demotion of the agent/initiator (Chen 2020). Its use is illustrated below with data from Thao (17) and Saaroa (18).⁴

- (17) THAO
 a. Yaku t<**m**>uqris takic.
 1SG.(PIVOT) <AV>catch.with.a.nose.trap barking.deer.ACC
 ‘I caught a barking deer with a snare trap.’
 b. **Mu**-tuqris iza na lhizashan.
 AV.DETR-catch.with.a.snare.trap this LK pheasant.(PIVOT)
 ‘The pheasant is caught with a snare trap.’ (Blust 2003b:1020)
 (*mu*- as a detransitiviser)
- (18) SAAROA
 a. C<**um**>acuhlu a tamalungaluna hliasasapa.
 burn-<AV> PIVOT uncle field
 ‘Uncle used fire to burn the field.’
 b. **Mu**-cacuhlu-a kiu’u naka manganicu.
 AV.DETR-burn-PROJ wood AUX be.dry
 ‘Dry wood is easy to be burned.’ (ODFL) (*mu*- as a detransitiviser)

Recent work has shown that the sequence *mu-* is bimorphemic, consisting of an AV allomorph *m-* and the detransitivizer *u-* (Chen 2020). Accordingly, the valency-decreasing function of this sequence is denoted by the vocalic affix *u-*. This affix is, however, morphophonologically opaque in modern

4. The affix *mu-* has yet another function reconstructable to PAN: as a motion prefix attached to nominal stems (Blust 2003a; Adelaar 2014). As this function is not reflected in Tgdaya Seediq, I do not discuss it in this paper.

Seediq due to the vowel neutralization rule noted in (5). Consequently, the prefix *m-* functions as an opaque detransitivizer in Seediq. As seen in (19) and (20), the alternation between the *m-* vs. <*m*>- marked constructions is analogous to that observed in Thao and Saaroa—in the apparently unmarked detransitive (19b) and (20b), the sentence is obligatorily one-place with the theme in subject-marking, parallel to the *mu*-marked detransitized clause in (17b) and (18b). This suggests that the diachronic source of *m-* in this specific environment should be the detransitivizing sequence **m-u-*.⁵

(19) SEEDIQ

- a. T<**m**>ggequq Ø huling nii *(ka Watan).
 <AV>drown CM₁ dog this *(PIVOT Watan)
 ‘Watan drowned the dog.’
- b. **M**-tggequq ka huling nii di.
 M-drown PIVOT dog this PRF.PART
 ‘This dog drowned.’
- c. ***M**-tggequq Ø huling nii ka Watan.
 M-drown CM₁ dog this PIVOT Watan
 (intended: ‘Watan drowned the dog.’)

(20) SEEDIQ

- a. **M**-takur ka Robo di.
 M-trip PIVOT Robo PRF.PART
 ‘Robo tripped.’
- b. ***M**-takur Ø Walis ka Temi.
 M-trip CM₁ Walis PIVOT Temi
 (intended: ‘Temi tripped Walis.’)
- c. T<**m**>akur Ø Walis *(ka Temi).
 <AV>trip CM₁ Walis *(PIVOT Temi)
 ‘Temi tripped Walis.’

2.5. REFLEX OF PAN RECIPROCAL AFFIX **maCa-* IN SEEDIQ.

Yet another diachronic source of *m-* is the PAN reciprocal affix **maCa-* (see, e.g., Zeitoun 2002; Sung and Shen 2006; Blust 2013 [2009]; Ross 2015). In languages with a reflex of this affix, reciprocal verbs are affixed by the sequence *ma-* followed by a syllable whose onset copies that of the

5. One anonymous reviewer asked whether PAN **m-* and **mu-* have only be merged orthographically and not phonemically in Seediq. Accordingly to primary fieldwork, these two morphemes are phonologically and phonetically undistinguished in Tgdaya Seediq, and speakers do not considered the two to be functionally different sound sequences. Linguists would know that these two are phonemically distinguished given their functional and etymological differences, but that phonemic distinction is synchronically not realized phonetically—which is the main topic of this squib. This observation follows from the consensus among previous works that pre-penultimate vowels are consistently realised as [u] following vowel neutralization (see, e.g., Yang 1976; Li 1991; Holmer 1996; Tsukida 2009; Ochiai 2016; Sung 2018). This predicts that all *m*-initial prefixes would be phonologically undistinguished in Seediq—which is exactly what is observed in primary fieldwork.

stem. This is exemplified with the data below from Paiwan and Puyuma (21a,b).

- (21) a. PAIWAN
Mala-laing tiamadu.
 REC-chase 3PL/PIVOT
 ‘They chased each other.’ (Sung and Shen 2006:51)
- b. PUYUMA
Mara-raip mi me-rebay dra radis.
 REC-exchange.labour 1PL.PIVOT AV-weed INDF.ACC peanut.plant
 ‘We took turns to pull out peanut plants for each other.’ (ODFL)

Under the vowel truncation rule in (5), this affix surfaces in the form of *m-C-* in Seediq, as in (22)–(23).

- (22) SEEDIQ
Mg-gaalu kanna seediq m-eniq alang cbeyo.
 REC-love all people AV-live village Past
 ‘In the past, everyone living in the village loved each other.’

(23)

		<i>Example</i>			
		Stem	<i>Gloss</i>	Reciprocal form	<i>Gloss</i>
<i>m-C-</i> ‘reciprocal’	cebu	‘to shoot’	m-c-cebu	‘to shoot each other’	
	gatak	‘to peck’	m-g-gatak	‘to peck each other’	
< PAN *maCa-	hetun	‘to obstruct’	m-h-hetun	‘to obstruct each other’	
	bege	‘to give’	m-b-bege	‘to give each other’	
	sneeru	‘to mimic’	m-s-sneeru	‘to mimic each other’	

Although the reciprocal affix above (*m-C-*) does not fully merge with those surfacing as *m-*, multiple instances of C omission are attested in Seediq’s reciprocal constructions, adding to the morphological opacity of *m-*. Consider the examples below, excerpted from the Seediq Dictionary (ODFL) (24a–c).

- (24) SEEDIQ
- a. **M**-traqil hari mtsuwai tnsapah gaga.
 REC-dislike more brothers.and.sisters family.members that
 ‘The children in that family dislike each other.’ (ODFL)
- b. M-suupu=naq theyaq de **m**-stuku=naq hari duri.
 AV.get.together=self visit CONJ REC-collaborate.for.work=self more as.well
 ‘People who have frequent contact work for each other as well.’
- c. **M**-skuxun weewa daha ka riso ga.
 REC-like young.lady two PIVOT young.man that
 ‘The young girl and the young man like each other.’ (ODFL)

The PAN reciprocal affix *maCa- thus constitutes a fifth possible function of *m-* in Seediq.

Finally, it is noteworthy that *m-* has yet another possible function in Seediq. It appears as a prefix for counting 6–10, and this *m-* might be some kind of prefix since it generally precedes a consonant (i.e., *mataru* (for six), *mpitu* (for seven), *maspat* (for eight), *mengari* (for nine), and *maxal* (for ten)). However, other counting numbers do not contain this morpheme. Consider the example below from ODFL.

- (25) **M**-pitu dapa ga meniꞑ Sipo.
 M-seven cattle PROG stay the.opposite.bank
 ‘There are seven cattle on the opposite bank.’ (ODFL)

Although its etymology remains unclear, this affix is clearly distinct from the AV allomorph *m-*, as it is not subject to the phonotactic constraint in (7) and consistently surfaces as a prefix *m-* (and never an infix $\langle m \rangle$) even when attached to C-initial stems. Synchronically, however, this (unproductive) use does constitute the sixth function that is morphologically realized as *m-* in Seediq.

3. WHY CAN SEEDIQ AFFORD THIS MERGER? To summarize, pre-stress vowel neutralization has resulted in the syncretism of five grammatical affixes—PAN **m-* (Actor Voice allomorph), **mu-* (detransitivizer), **maCa-* (reciprocal affix), **ma-* (stative affix), and **mi-* (possessive affix), as well as the affix *m-* used in counting—in Seediq. Our perception of morphology involves generalizations of the mapping between sound segments and functions. When a sound segment representing a certain function is lost, can the function of the morpheme survive in speakers’ minds and be comprehended appropriately?

A key debate related to this question is whether functional morphemes would resist regular sound changes in order to maintain their function (Kiparsky 1982; Hill 2014; Bybee 2017). An implicit assumption behind this inquiry is that the absence of morphological distinction entails the loss of meaning/function. However, as seen above in section 2, the morphological opacity of these five affixes (as well as the affix *m-* used for counting 6–10) has not triggered the loss of these functions. Instead, *m-* serves as a multifunctional affix and is cognate with five different affixes in other Austronesian languages. Why can Seediq afford this degree of morphological opacity? This question is particularly relevant for this specific case, as all five morphemes that have undergone syncretism share a general Actor Voice component *m* (see Ross 2015, for a further discussion), with the very vowel that carries the functional distinction obscured by the vowel neutralization process.

Scrutinizing the lexical subcategorization of these affixes offers a logical account of why the merger is “affordable.” Consider the table in (25), which outlines the mapping between each affix and the type of stem it selects.

(26) Lexical subcategorization of the six affixes

PAN	<i>Reflex</i>	<i>Function</i>	<i>Phonological Condition</i>	<i>Lexical Subcategorization</i>	
a. *m-	<i>m-</i>	AV allomorph	vowel-initial stems	verbal stems	dynamic (1-place, 2-place, 3-place)
b. *mu-	<i>m-</i>	Detransitivizing affix	–	verbal stems	dynamic (2-place)
c. *maCa-	<i>m-</i>	Reciprocal affix	–	verbal stems	dynamic (2-place, 3-place)
d. *ma-	<i>m-</i>	Stative affix	–	verbal stems	stative
e. *mi-	<i>m-</i>	Possessive affix	–	nominal stems	non-numerals
f. –	<i>m-</i>	Affix for counting 6–10	–	nominal stems	numerals (see section 2.5)

Consider, firstly, the two affixes that select nominal stems, *mi- (possessive affix) and *m-* (affix for counting; see footnote 6). Their lexical subcategorizations are mutually exclusive. The former selects stems denoting physical entities (26e), whereas the latter combines only with the numerals 6–10 (26f). The use of the same morpheme for these two functions can therefore be viewed as economic, as the actual function can be inferred by context and by which stem is attached. Furthermore, as these two functions apply only to nominal stems, they would not conflict with the other four functions of *m-*, which apply only to verbal stems.

The homophony of the stative affix (26d) with the detransitivizer (26b) and the reciprocal affix (26c) is also non-problematic, as stative stems constitute a distinct class and are compatible neither with detransitivization nor with reciprocal expressions. Similarly, its lack of distinction from the AV prefix *m-* would not trigger issues in processing or learnability, as dynamic stems and stative stems are semantically distinguished and need not rely on morphological marking to form a distinction.

The key question thus boils down to the homophony of the AV prefix (26a), the detransitivizer (26b), and the reciprocal affix (26c). Does their merger trigger issues with comprehension or learnability? A closer look at their use suggests the answer is negative. Reciprocal constructions are characterized by a plural, animate agent, and the absence of the direct object (unless in constructions formed with a ditransitive verb). This contrasts with detransitivized constructions, whose sole argument is an undergoer, with the agent obligatorily absent. The homophony of these two functions in morphological marking therefore creates little problem in comprehension.

The only environment that triggers the loss of grammatical distinction concerns the detransitive of dynamic verbs with a vocalic onset. Recall that in Seediq, the alternation between <*m*> and *m*- (reflex of PAN **mu*-) indexes the argument structure alternation between a detransitive construction and its two-place counterpart (section 2.4). This is because the prefix *m*- is derived from the detransitivizing sequence *m-u*-. Given the rule in (7), the vocalic detransitizer *u*- altered AV morphology to a prefix prior to its loss due to vowel neutralization. Consequently, this argument structure alternation is indexed by the infixal vs. prefixal distinction of the affix *m*. This alternation is illustrated in (26).

- (27) a. Wada s<m>etuq Ø negul nii ka Watan.
 PRF <AV>break CM₁ string this PIVOT Watan
 ‘Watan broke this string.’ infix (<*m*> as an AV allomorph)
- b. Wada m{tu}-setuq ka hako=ta.
 PRF AV-~~DETR~~-break PIVOT bridge=1PL.POSS
 ‘Our bridge broke.’ (prefix *m*- as a reflex of PAN **m-u*-)

However, for two-place verbs with a vocalic onset, their AV-marked construction would originally be reflected as *m*- due to the allomorphic rule in (7), as in (27a). The detransitized form of such verbs would therefore be predicted to be morphologically undistinguished from the *m*-marked two-place counterpart, as illustrated in (27b).

- (28) a. Wada m-aduk Ø rodux-mu di ka dakis.
 PRF AV-drive.out ACC chicken-1SG.POSS PART PIVOT Dakis
 ‘Dakis drove out my chicken.’ (ODFL) (Actor Voice, *m*-)
- b. Wada m{tu}-aduk ka rodux-mu.
 PRF AV{-~~DETR~~}goad PIVOT chicken-1SG.POSS
 (Intended: ‘My chicken was goaded.’)
 (Hypothetical detransitive counterpart, *m*-)

According to primary fieldwork, for ambiguous cases such as (27), the intended detransitive construction is consistently unacceptable, as speakers would interpret the *m*-initial verb form (e.g., *maduk*) as its default (two-place) form. In addition, Seediq does possess an alternative structure available for denoting similar information structure. Like many other Formosan languages, Seediq has a Philippine-type voice system with a specific voice type for highlighting the undergoer of an event. This construction is known as the Patient Voice (PV). Similar to the detransitive construction (28b), the undergoer in a PV-marked sentence carries PIVOT-marking, and is prominent in discourse. Consider below the alternation between the AV, detransitive, and PV-marked construction of the same stem *dengu* ‘roast’.

- (29) a. Wada d<m>engu Ø qhuni ka Dakis.
 PRF <AV>roast CM₁ wood PIVOT Dakis
 ‘Dakis roasted/dried the wood.’ (Actor Voice)
- b. Wada m-dengu ka qhuni.
 PRF M-roast PIVOT wood
 ‘The wood has been roasted/dried.’ (*m*-marked detransitive clause)

- c. Psa-un daha gigan **ka macu** han ma, dngey-un daha
 place-PV 3PL.GEN basket **PIVOT millet** first CONJ roast-**PV** 3PL.GEN
 bobo bagah puniq (pro).
 then charcoal fire (pro.PIVOT)
 ‘They first put the millet in a basket, and then they dry (it) with fire.’
 (ODFL) (Patient Voice)

Speakers report that the detransitive construction (28b) and the PV construction (28c) both place an emphasis on the undergoer and could be used as an approximate alternative structure. This suggests that a detransitive construction (e.g., the unacceptable example (27b)) could be replaced with a PV clause where it is morphologically undistinguished from its two-place counterpart. Vowel deletions triggered by the process in (5) therefore do not trigger major issues in eliminating grammatical distinctions, as the morphological opacity can be easily circumvented by the use of a different structure.

Before concluding, a theoretical question deserves a note. So far, we have seen that the functions carried by the five morphemes have survived as they apply to recognizably different contexts. While their etymology can be made transparent through comparative data from sister languages, how do we rule out an alternative analysis that these morphemes have merged not only in the surface realization but also in syntax, with *m-* as a multifunctional affix whose function is conditioned by context? In other words, is there any independent way of testing whether the morphemes have not merged entirely?

It turns out there is. In Atayalic languages, the basic negation triggers obligatory connegative prefixal alternation on the following verb (see [Holmer 1996](#), [Ochiai 2016](#), and [Tsukida 2009](#) for details). One way of testing if the five morphemes remain distinct is, therefore, to elicit the connegative form of each type and see if they vary systematically. As pointed out by an anonymous reviewer, for *mu-* marked detransitive verbs, the connegative is predicted to be the stem without *m-* (as it is for verbs with the *-m-* infix), whereas for the majority of statives, the connegative is formed by replacing *m-* with *k-*. For reciprocals, the connegative is usually formed by replacing *mC-* with *pC-*. Similar tests could also be made for the other functions. This would enable a testable prediction for the current analysis and clarify whether the various functions carried by *m-* have remained morphosyntactically distinct in modern Seediq. I leave this question for future investigation.⁶

4. CONCLUSION. The Formosan language Seediq displays an instructive case of morphological opacity, where a single phonological constraint has obscured—but not obliterated—five grammatical affixes. The prefix *m-* in the language is known for denoting a large number of functions (and sometimes does not seem to have a clear function at all). We have seen that the morpheme is diachronically linked to at least five different morphemes, which have become morphologically undistinguished due to a phonotactic constraint of

6. My thanks to an anonymous reviewer for bringing up this testable hypothesis.

the language, which regularly elides pretonic vowels and replaces them with an epenthetic [u]. However, despite their morphophonological opacity, the marginal overlap between these affixes' lexical subcategorization may have reduced obstacles to learnability and processing, enabling them to remain functionally distinct despite the absence of morphological distinctions. The case of Seediq thus highlights the often-neglected fact that morphological opacity does not necessarily obliterate grammatical distinctions.

REFERENCES

- Adelaar, Alexander. 2014. The morphological expression of location and motion in Siraya. *NUSA: Linguistic Studies of Languages in and around Indonesia* 56:101–18.
- Baerman, Matthew. 2007. Syncretism. *Language and Linguistics Compass* 1(5): 539–51.
- Baerman, Matthew, Dunstan Brown, and Greville Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.
- Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Blust, Robert. 2003a. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42(2): 438–78.
- . 2003b. *Thao dictionary*. Taipei: Academia Sinica.
- . 2013 [2009]. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert, and Stephen Trussel. Ongoing. The Austronesian Comparative Dictionary (ACD). <http://www.trussel2.com/ACD/>.
- Bybee, Joan. 2017. Grammatical and lexical factors in sound change: A usage-based approach. *Language Variation and Change* 29(3): 273–300.
- Chang, Anna Hsiu-chuan. 2006. A reference grammar of Paiwan. PhD diss., The Australian National University.
- Chang, Henry Yung-li. 2000. *A reference grammar of Seediq*. Taipei: Yuanliu.
- Chen, Victoria. 2020. The derived intransitive in Formosan and its implication for the nature of Proto-Austronesian Actor Voice. *Oceanic Linguistics* 59(1/2): 59–90.
- Hill, Nathan. 2014. Grammatically conditioned sound change. *Language and Linguistics Compass* 8(6): 211–29.
- Holmer, Arthur. 1996. *A parametric grammar of Seediq*. Lund: Lund University Press.
- Kiparsky, Paul. 1982. *Explanation in phonology*. Dordrecht: Foris.
- Li, Paul Jen-kuai. 1977. Morphophonemic alternations in Formosan languages. *Bulletin of the Institute of History and Philology, Academia Sinica* 48(3): 375–413.
- . 1991. Vowel deletion and vowel assimilation in Seediq. In *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, ed. by R. Blust, 163–69.
- Mckinnon, Timothy, Gabriella Hermon, Yanti, and Peter Cole. 2018. From phonology to syntax: Insights from Jangkat Malay. In *Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics*, ed. by H. Bartos et al., 349–71. Cham: Springer.
- Melchert, Craig. 1975. 'Exceptions' to exceptionless sound laws. *Lingua* 35(2): 135–53.
- Ochiai, Izumi. 2016. Seediq reciprocals and their history. *Kyoto University Linguistic Research* 35:161–93.
- Online Dictionary of Formosan languages (ODFL). <http://e-dictionary.apc.gov.tw>.
- Rau, Victoria and Maa-neu Dong. 2006. *Yami texts with reference grammar and dictionary*. Taipei: Academia Sinica.
- Ross, Malcolm. 2015. Reconstructing proto Austronesian verb classes. *Language and Linguistics* 16(3): 279–345.

- Sung, Li-May. 2018. *A grammar of Seediq*. Council of Indigenous Peoples.
- Sung, Li-May, and Chia-chi Shen. 2006. Reciprocals in Kavalan and a typological comparison. In *Streams Converging into an Ocean: Festschrift in Honor of Professor Paul Jen-kuei Li on His 70th Birthday*, ed. by H. Chang et al., 239–77. Taipei: Academia Sinica.
- Tamma, Meredith. 2013. Phonology and morphology in Dutch indefinite determiner syncretism: Spatial and quantitative perspectives. *Journal of Linguistic Geography* 1(2): 115–24.
- Teng, Stacy Fang-ching. 2008. *A grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: The Australian National University.
- . 2014. Grammaticalization of predicative possession in Nanwang Puyuma and as a basis for reconstruction in PAN. *Oceanic Linguistics* 53(1): 136–54.
- Tsukida, Naomi. 2009. A grammar of the Seediq language (in Japanese). PhD diss., The University of Tokyo.
- Yang, Hsiu-fang. 1976. The phonological structure of the Paran Dialect of Seediq. *Bulletin of the Institute of History and Philology, Academia Sinica* 47:611–706.
- Zeitoun, Elizabeth. 2002. Reciprocals in the Formosan languages: A preliminary study. Paper presented at the ICAL 9, Canberra, January 8–11.