

## Original Paper

# Lexical Incorporation of Loanwords in the Sengwato Dialect of Setswana

Mahmoud S. Al Mahmoud<sup>1\*</sup>

<sup>1</sup> Imam University, Riyadh, Saudi Arabia

\* Mahmoud S. Al Mahmoud, Imam University, Riyadh, Saudi Arabia

Received: July 6, 2020

Accepted: July 14, 2020

Online Published: July 23, 2020

doi:10.22158/sll.v4n3p100

URL: <http://dx.doi.org/10.22158/sll.v4n3p100>

### Abstract

*This paper reports on the incorporation of lexical items into one of the well-known dialects of Setswana, namely Sengwato. A native speaker of the Sengwato dialect provided much of the data for this study. The data show that borrowings from English and Afrikaans Dutch adhere to the phonological canons of Sengwato via the application of a number of sound change processes such as lenition, fortition, affrication, palatalization and obstruent devoicing to name a few. In addition, the data demonstrate how lexical borrowings undergo gender-based classification when incorporated into Sengwato and are categorized into fourteen gender classes identified from the Swadesh list provided by the informant. The paper discusses how the semantic content of loanwords plays a vital role in their incorporation into the Sengwato dialect.*

### Keywords

*Sengwato, loanword phonology, lexical incorporation*

### 1. Introduction

Much of the linguistic aspects of the Niger-Congo languages remain largely undocumented. The present research paper deals with a well-known dialect of the Setswana language, a Bantu branch belonging to the larger Niger-Congo family, namely the Sengwato dialect. A Sengwato native speaker from the town of Shoshong in central Botswana provided the data for this study. Even though Sengwato is one of the indigenous dialects spoken in Shoshong, the informant in this study reported exposure and knowledge of other dialects in the region such as Segwatketsi during her early years of schooling. However, the informant explained that all dialects of the Setswana language are mutually intelligible and that it was hard to find an individual speaking only one variety of the language. The Sengwato dialect presents a particularly intriguing case to investigate since it is heavily borrowed into

from English and Afrikaans Dutch besides the fact that much linguistic inquiry is lacking.

The goal of this short paper is to draw on linguistic data from the native speaker of Sengwato in an attempt to describe and generalize certain linguistic aspects relevant to the incorporation of lexical borrowings from English and Afrikaans Dutch into the Sengwato dialect. Such characteristics include, but are not limited to, the linguistic processes in the adaptation of phonological, syllable structure, and morphosyntactic loanwords. The analysis of the linguistic incorporation of loans into Sengwato is framed within the rule-based as well as the Optimality-theoretic approach.

## 2. Incorporation of Loanwords in Sengwato

### 2.1 Phonological Adaptation

In reviewing the English and Afrikaans Dutch loanword data provided by the Sengwato native speaker, a number of phonological processes affecting consonants can easily be identified; in what follows I highlight these processes and provide illustrative examples to manifest whatever phonological changes occurred due to frequent borrowing from English and Afrikaans Dutch.

The first process to note is that of *lenition*. Lenis is a sound pronounced with less intensity and more laxly than a plain sound (Spencer, 1996, p. 24). Lenition refers to the process by which a consonant is abated and weakened. One instantiation of lenition is spirantization or fricativization which is turning an obstruent sound into continuant, and it is very commonly found in loanwords. Trask (1996, p. 56), for example, explains how Italian made use of spirantization in borrowing the word *fava* from the Latin *faba* “bean”. In the limited Sengwato data provided by the informant, the following example of lenition, and more particularly spirantization, is observed:

- (1) /se.ge.re.se/ < Eng. /sɪ.gə.ɹet/ “cigarette”

In (1) the Sengwato word *segerese* is borrowed from the English word *cigarette*, with the [-continuant] voiceless alveolar obstruent /t/ changed into the closest [+continuant], namely the voiceless alveolar fricative /s/.

Further examination of the data revealed a number of cases of another widely attested process, *final obstruent devoicing*:

- (2) /se.to.fo/ < Eng. /stov/ “stove”  
 (3) /bo.ro.to/ < Eng. /bord/ “board”  
 (4) /bo.ro.t<sup>h</sup>o/ < Afrikaans Dutch /bro:d/ “bread”  
 (5) /pɪ.li.si/ < Eng. /pɪlz/ “pills”

It is important to note that due to certain syllable form restrictions, e.g., a clear preference for a consonant vowel sequence (CV), Sengwato, in its borrowing of foreign words, consistently copies the main word vowel for the third syllable it creates (section 2.2 for further discussion). The origin for the Sengwato words in (2-5) is monosyllabic words from English and Afrikaans. In their adaptation, not only did Sengwato devoice the last consonant in the lending language, /v/ to /f/ in (2), /d/ to /t/ in (3),

/d/ to /t<sup>h</sup>/ in (4), and /z/ to /s/ in (5), but also managed to turn monosyllabic words into trisyllabic ones via vowel epenthesis. It would appear, however, that final obstruent devoicing takes precedence over vowel epenthesis application in order to derive the examples in (2-5):

(6) Final obstruent devoicing and vowel epenthesis in the Sengwato loanwords /se.to.fo/ “stove” and /bo.ro.to/ “board”

<b>U(nderlying) R(epresentation)</b>	stov	bord
<b>Final obstruent devoicing</b>	stof	bort
<b>V-epenthesis</b>	setofo	boroto
<b>Resyllabification</b>	se.to.fo	bo.ro.to
<b>S(urface) R(epresentation)</b>	se.to.fo	bo.ro.to

If, on the other hand, vowel epenthesis were to apply before final obstruent devoicing, the outputs \*se.to.vo and \*bo.ro.do would be erroneously derived. It is worth noting that the pronunciation for the words “board” and “bread” in (3) and (4), distinguishable only by the aspiration of /t<sup>h</sup>/ in (4), demonstrates how Sengwato utilizes subtle allophonic variations contrastively to signal lexical distinctions.

In the native speaker’s speech, there was one instance of *affrication*, a process by which an affricate sound is formed:

(7) /me.tse/ < Eng. /mæθ/ “math”

In (7) the voiceless interdental fricative /θ/ was tuned into the voiceless dental affricate /ts/.

The next phonological process found in the data is *fortition*, whereby weak segments are turned into strong ones. Fortition or fortis is an opposing process to the effects of lenition and is by far less common than lenition in language (Trask, 1996). In the examples (8) and (9), the continuants /θ/ and /l/ undergo fortition:

(8) /ta.n.ki/ < Eng. /θænk/ “thank”

(9) /(le)po.di.si/ < Eng. /pəlis/ “police”

In (8), the voiceless interdental [+continuant] /θ/ is borrowed into the voiceless [-continuant] plosive /t/. In (9) the continuant lateral /l/ is turned into its allophonic non-continuant variant /d/. The /l/ to /d/ change appears to be precipitated by a Sengwato phonotactic restriction on the sequence \*li. This restriction, however, does not hold for other similar loanwords in the data which allow li sequences; why would the /l/ change to /d/ in /po.di.si/ < Eng. /pəlis/ “police” but not, for example, in /pɪ.li.si/ < Eng. /pɪlz/ “pills”?

The data suggest that the phonotactic restriction on the \*li sequence is enforced only in non-derived environments. In other words, li alternates with its allophonic variant di only if the vowel that triggers the delateralization is present in the source language and is not the outcome of epenthesis. This generalization can be formalized as follows and it shows how delateralization would apply prior to vowel epenthesis:

(10) l~d and vowel epenthesis in Sengwato loanwords /po.di.si/ “police” and /pɪ.li.si/ “pills”

<b>UR</b>	pə.lis	pɪlz
<b>Final obstruent devoicing</b>	N/A	pɪls
<b>Delateralization</b>	po.dis	N/A
<b>Vowel epenthesis</b>	po.disi	pɪlisi
<b>Resyllabification</b>	po.di.si	pɪ.li.si
<b>SR</b>	po.di.si	pɪ.li.si

It is important to maintain the ordering of the rules in (10). If the established ordering between delateralization and vowel epenthesis were to be reversed, the derivation would incorrectly yield \*pɪ.di.si.

The data also show one case of *voicing progressive assimilation* as is clear from the devoicing of /z/ in its assimilation to the preceding voiceless plosive /t/:

(11) /ka.tse/ < Germ. /ka.tze/ “cat”

*Palatalization* also is another process found in the speech of the native Sengwato informant:

(12) /en.je.la.ne/ < Eng. /ɪŋ.glənd/ “England”

In (12) the velar /g/ seems to undergo palatalization to /j/ in the environment of the front vowel /e/. Palatalization in the vicinity of front vowels is not uncommon among the world languages (e.g., see Rubach, 2002 on Russian palatalization, and Ito and Mester, 2003 on Japanese palatalization).

The data further reveal a number of other interesting phonological processes as a result of the Sengwato lexical incorporation of the following English and Afrikaans Dutch words:

(13) /u.ni.be.se.ti/ < Eng. /ju.ni.vɛr.si.ti/ “university”

(14) /ti.vi/ < Eng. /ti.vi/ “TV”

Sengwato lacks the labiodental continuant /v/ and is often compensated for by the bilabial non-continuant /b/ as shown in (13); this can be considered a clear case of fortition. In (14), however, /v/ retains its source language form in the incorporation of the word /ti.vi/ “TV”. The failure of fortition to apply in (14) may be an exception possibly due to potential confusability with other lexical words. In Sengwato, the word for the tuberculosis disease “TB” is incorporated as *ti.bi*. If fortition were to apply in the adaptation of “TV”, the Sengwato word would also be *ti.bi*. It is possible, therefore, that in order not to render the two words TB and TV confusable by maintaining lexical distinction between them in pronunciation, /v/ is adapted in Sengwato without a change for the word “TV”.

Further examination of the Sengwato data shows that whenever the rhotic /r/ occurs as the coda of the source language word syllable, it gets consistently deleted as in (15), or replaced by the lateral liquid /l/ as in (16):

(15) /mɪs.tə/ < Eng. /mɪs.tər/ “mister”

(16) /ha.mo.le/ < Afrikaans Dutch /hæ.mər/ “hammer”

The data also show that final clusters are reduced through consonant deletion:

(17) /po.so/ < Afrikaans Dutch. /post/ “post”

In (17) elision of the second element /t/ in the source language takes place to resolve the coda cluster.

## 2.2 Syllable Structure in Loanwords

The examples of *final consonant deletion* in Sengwato present a clear evidence for a preference of the unmarked open syllable CV. Preference for the syllable CV is argued to be universal as it is the most common and unmarked form of the syllable in the world’s languages (Cairns & Feinstein, 1982; Clements, 1990; Clements & Keyser, 1983; Greenberg, 1978). In addition, the CV form of the syllable seems to be perceptually more privileged than a CC sequence; a consonant in a prevocalic CV-position is argued to be auditorily more salient than a consonant in a preconsonantal CC-position (Wright, 2004).

One possibility is to treat cases of final consonant deletion discussed in section 2.1 as an overt effect of a restriction on the maximum number of syllables allowed in the borrowing process. In other words, if the source language word has two syllables, then the borrowing language will have two syllables as well. In (17), and given the preference for open syllables, retaining final /t/ would have resulted in a trisyllabic word, namely \*/po.so.to/. Thus, it would be more economical and faithful to the source language to keep the same syllable count of two. We can extrapolate that the same argument also applies to (15). It is not readily clear, however, why /r/, while deleted in (15), is retained in (18) where the condition on syllable max-count seems to be violated:

(18) /ti.tʃa.ra/ < Eng. /ti.tʃɜːr/ “teacher”

In general, there appears to be a tendency in Sengwato to delete word final /r/ whenever an [+ATR] vowel follows in the borrowing language.

Further examination of the data confirms that Sengwato has a well-attested preference for codaless syllables (CV); in fact, Sengwato appears to place a stipulation on the minimum number of moras each word must have. These two generalizations on Sengwato syllable structure can be summarized as follow:

- (19) a. Every word must have at least two moras  
 b. Closed syllables are not allowed; only CV or V syllable forms

The bi-moraic condition in (19a) accounts for CV monosyllabic words that undergo vowel lengthening as they are being incorporated into Sengwato. The generalization in (19b) specifies the permissible types of syllable structures in Sengwato. Preference for open CV syllables over closed VC ones is universally acknowledged. Not only is the sequence CV is the most common type of syllable structure in the world’s languages (Clements, 1990; Greenberg, 1978), it provides an aurally more privileged context for the perceptibility of adjacent consonants (Borden, Katherine, & Lawrence, 2003; Wright, 2004). The idea is that in CV sequences the consonant releases into the following vowel; this provides ample acoustic cues for the perceptibility of the consonant. On the other hand, in VC sequences the

consonant trails off the syllable with no possible audible release into a vowel; this impoverishes the consonant cues leading ultimately to misperception or imperceptibility altogether.

It may appear at first that the stipulation in (19a) is superfluous since most Sengwato words show a tendency towards open syllables. Thus, *vowel epenthesis* in words like /bu.ka/ “book” would be implemented to conform to the CV syllable condition in (19b):

(20) /bu.ka/ < Eng. /bʊk/ “book”

Note that in the incorporation of the word “book” in (20) the CV syllable condition can be satisfied alternatively through consonant deletion (i.e., \*/bu/ < Eng. /bʊk/ “book”). Yet in line with (19a), Sengwato adapts the word “book” by means of vowel epenthesis instead, forming a second syllable altogether and increasing the mora count from one to two. Further, the incorporation of other monosyllabic CV words by tensing and lengthening the vowel suggests that a condition on the mora count in Sengwato exists:

(21) /ti:/ < Eng. /ti/ “tea”

It is clear that the application of vowel epenthesis in (20), as well as in (22) below, works to avoid a violation of the mora minimality and the no word-final coda conditions stated in (19a, b), respectively:

(22) /i.n.ki/ < Eng. /ɪnk/ “ink”

In (22) a vowel is inserted to form a third CV syllable, namely *ki* (note that the nasal is syllabic here). This is also true for the adaptation of the word “table” in (23) where vowel epenthesis is utilized to create the simple form of the syllable structure CV (i.e., *le*):

(23) /ta.fo.le/ < Afrikaans Dutch /ta.fəl/ “table”

Vowel epenthesis can also be employed in Sengwato to resolve consonant clusters in loanwords; thus, another syllable well-formedness condition related to complex onsets and codas can be stated as follows:

(24) Complex onsets and codas are illicit in Sengwato, except for [strident+stop] onset clusters: {sp-, st-, sk-} for which cluster simplification is optional

The statement in (24) on syllable structure well-formedness is well-adhered to in the adaption of the words (25-30):

(25) /ka.ra.ta/ < Eng. /kɑrd/ “card”

(26) /ta.n.ka/ < Eng. /tænk/ “tank”

(27) /xa.la.si/ < Afrikaans Dutch. /xlas/ “glass”

(28) /se.ta.ra.ta/ or /ste.ra.ta/ < Afrikaans Dutch. /stra:t/ “street”

(29) /se.pu.ne/ or /spu.ne/ < Eng. /spun/ “spoon”

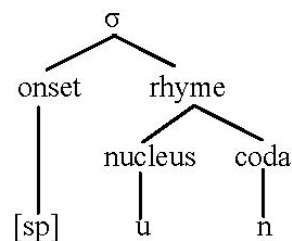
(30) /se.kwe.le/ or /skew.le/ < Eng. /skul/ “school”

In (25) the word for “card” is adapted as /ka.ra.ta/ with a vowel breaking up the final cluster (-rd); note that Sengwato also inserted another vowel after the final consonant /d/, which was devoiced to /t/, in line with the (19b) condition on syllable structure. In (26) the coda cluster in the source language

“tank” was avoided not via epenthesis but by means of resyllabification of the nasal. Once the nasal /n/ is treated as syllabic comprising a separate syllable by itself, it no longer became part of the coda cluster /-nk/. The vowel that is inserted after /-k/ is arguably then due to the condition in (19b). In (27) the complex onset /xl-/is obligatorily reduced through vowel epenthesis. The examples (28-30) show the optionality in the simplification of /st-, sp-, sk-/ clusters; Sengwato often alternates in the pronunciation of these words between epenthetic and non-epenthetic forms, as illustrated above.

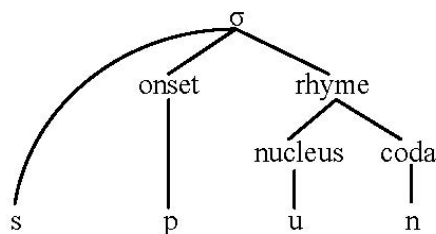
The fact that [strident+stop] clusters are treated differently from other onset clusters in Sengwato borrowings can be attributed to the status of these clusters in the world’s languages. Despite their prevalence universally, [strident+stop] clusters appear to violate the Sonority Sequencing Principle (SPP) which stipulates that in onsets sonority should rise toward the syllable peak or nucleolus (the vowel), and fall thereafter in the coda toward the end of the syllable. However, in onset [strident+stop] clusters the transition from the strident /s/ to the following stop /t/, /p/ or /k/ involves a sonority reversal since /s/ is more sonorous than the voiceless stops (Broselow and Finer, 1991, Clements, 1990, Selkirk, 1982).

It has also been argued typologically that in many languages [strident+stop] clusters resist vowel epenthesis more than other clusters. One explanation is that the onset sequence [strident+stop] is to be considered a one-unit cluster (Broselow, 1993; Kenstowicz, 1994) as outlined in Figure 1:



**Figure 1. Projection of /sp-/ Cluster as One Unit of the Onset**

Others like Giegerich (1992), and more recently Barlow (2001), treat the strident in the cluster as a leftmost adjunct adjoined to the syllable node as in Figure 2:



**Figure 2. Leftmost Adjunct Projection of the Strident /s/**

To sum up, [strident+stop] onsets' violation of the sonority principle, which regulates the sequencing of consonants and vowels in syllables, sets them apart from other clusters. Their treatment as one units in Figure 1, or as extrametrical with the strident appended prosodically to the initial syllable margin in Figure 2 offer possible insights as to why these types of onset clusters are less susceptible to vowel epenthesis, and may explain why Sengwato permits them in the adaption of loanwords.

Further examination of the Sengwato data reveals yet another syllable repair strategy, namely /r/-deletion, albeit less common than vowel epenthesis. Examples (31-33) demonstrate how Sengwato prefers to omit the postvocalic rhotic. In the words for "mister" and "professor" the word final /r/ is deleted; in (33) the /r/ is deleted after the vowel.

- (31) /mɪs.tə/ < Eng. /mɪs.tər/ "mister"  
 (32) /po.ro.fe.se/ < Eng. /prə.fɛ.sər/ "professor"  
 (33) /se.ke.te/ < Eng. /skɜrt/ "skirt"

Non-rhotic pronunciation of these words is quite characteristically common of the standard British Received Pronunciation (RP), amongst many other varieties of English (Roach, 2000). This pronunciation of English may have possibly affected how Sengwato adapts these words. Another possibility is that consonant deletion is triggered by syllable well-formedness conditions, namely Sengwato's dispreference for closed syllables stated in (19b). This would also explain why in Sengwato cases of consonant deletion are quite infrequent compared to vowel epenthesis, which is used for mora augmentation, cluster simplification, and avoidance of the coda condition as well. To conclude, not all loans in Sengwato undergo modification when adapted. Note how the adaptation for the English word "TV" in (34) is identical to the source language:

- (34) /ti.vi/ < Eng. /ti.vi/ "TV"

Similarly, in (35) and (36) the adaptation of the word "coffee" from Afrikaans Dutch, and the word "radio" from English appears unaltered:

- (35) /ko.fi/ < Afrikaans Dutch. /ko.fi/ "coffee"  
 (36) /re.dijo/ < Eng. /re.dijo/ "radio"

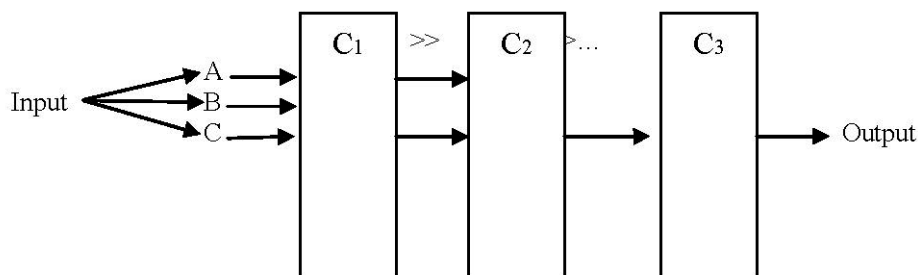
Failure to employ vowel epenthesis or consonant deletion in these cases is well-justified as all three syllable well-formedness conditions in Sengwato are met, namely (19a, b) and (24). This suggests that repair strategies in Sengwato apply minimally and only to resolve a violation of a syllable structure condition.

### 2.2.1 OT Analysis of Loanword Syllable Structure

In Optimality Theory (Prince & Smolensky, 1993), surface phonological forms are arrived at via input-output correspondence relations. The basic idea is that the (*Gen*)erator draws on underlying forms encoded in the *Lexicon* of a language to generate an infinite number of outputs for a specific input. Output forms are assessed through the function of *Eval(uator)* and whichever candidate incurs the least serious violation(s) of a set of violable constraints is the *optimal* candidate which gets to be the



actual output. Candidates are first evaluated on the highest ranked constraint, and any candidate that violates that constraint is eliminated. The same process applies for the rest of the hierarchally ranked constraints until one candidate survives by either satisfying all the constraints or violating the least marked one. A schematic representation of how the assessment and elimination process proceeds in OT is given below in Figure 3:



**Figure 3. Input to Output Mapping in OT Grammar (Kager, 1999, p. 8)**

There are two types of constraints: Markedness and Faithfulness constraints. Markedness constraints impose some criterion of structural well-formedness on outputs while Faithfulness constraints require that output candidates be faithful to their lexical input by preserving their input's properties (Kager, 1999, pp. 8-9).

In order to construct an OT-based analysis of Sengwato loanwords' syllable structure, the use of the following faithfulness and markedness constraints is necessary (Kager, 1999):

*Faithfulness constraints*

(37) **DEP-IO( $\mu$ )**

(38) **MAX-IO**

*Markedness constraints*

(39) **NO-CODA**

(40) **\*COMPLEX**

(41) **\*r[+ATR]#**

The faithfulness constraint in (37) stipulates that the number of mora(s) in the output match the number of mora(s) in the input while (38) prohibits any deletion in the output by preserving the input segments. The markedness constraint in (39) disallows closed syllables by banning codas while (40) is prohibitive of onset and coda clusters; in other words, it requires onsets and codas to be simple. The markedness constraint in (41) penalizes sequences of rhotic /r/ followed by [+ATR] vowels word finally.

Drawing back on the discussion in 2.2 and more specifically on the mora count requirement stated in (19b), an additional faithfulness constraint that gives preference to bi-moraic CV words over mono-moraic ones is needed in the analysis of Sengwato's prosodic adaption of loanwords:

(42) **MIN- $\mu\mu$** 

The constraint in (42) stipulates that every word must minimally consist of two moras; that is if faced with two possible output words, the one with the maximum number of moras will be preferred as illustrated in Table 1:

**Table 1. Preference of Bi-moraic vs. Mono-moraic Outputs in the Sengwato Word “Tea”**

<b>ti:</b>	<b>MIN-<math>\mu\mu</math></b>
ti	*!
→ ti:	

Crucially, the max mora count constraint in (42) has to be ranked above the no-deletion faithfulness constraint stated in (37): **MIN- $\mu\mu$  >> DEP-IO( $\mu$ )**; this ranking assures the elimination of mono-moraic outputs:

**Table 2. Ranking of MIN- $\mu\mu$  over DEP-IO( $\mu$ )**

<b>ti:</b>	<b>MIN-<math>\mu\mu</math></b>	<b>DEP-IO(<math>\mu</math>)</b>
ti	*!	
→ ti:		*

Furthermore, to ensure that the open-syllable stipulation in (19b) is respected, **DEP-IO( $\mu$ )** has to be ranked lower than the markedness constraint **NO-CODA**:

**Table 3. Ranking of NO-CODA over DEP-IO( $\mu$ )**

<b>tafəl</b>	<b>NO-CODA</b>	<b>DEP-IO(<math>\mu</math>)</b>
ta.fəl	*!	
→ ta.fo.le		*

The markedness constraint **\*COMPLEX** which mitigates against onset and coda clusters needs to be ranked above **DEP-IO( $\mu$ )** as well to prohibit the occurrence of words with clusters. This is demonstrated in Table 4:

**Table 4. Ranking of \*COMPLEX over DEP-IO( $\mu$ )**

<b>xlas</b>	<b>*COMPLEX</b>	<b>DEP-IO(<math>\mu</math>)</b>
xlas	*!	
→ xa.la.sa		**

Next is the markedness constraint  $*r[+ATR]\#$  which disallows sequences of /r/ followed by [+ATR] vowels word finally. While  $*r[+ATR]\#$  can be freely ranked with **MIN- $\mu\mu$**  and **NO-CODA**, it has to dominate **\*COMPLEX**, **DEP-IO( $\mu$ )** as well as **MAX-IO** in order to derive the right output form:

**Table 5. Ranking for the Word “Mister” in Sengwato**

	<b>mister</b>	$*r[+ATR]\#$	<b>MAX-IO</b>	<b>*COMPLEX</b>	<b>DEP-IO(<math>\mu</math>)</b>
→	mi.sta		*	*	
	mi.ste.re	*!		*	*

The relative ranking of the constraints in Table 5 favors the output with the r-deletion, namely *mi.sta* over the other output *mi.ste.re* which incurs a violation of the higher-ranked  $*r[+ATR]\#$  and is, therefore, eliminated. Note that in Table 5 the ranking between **\*COMPLEX** and **DEP-IO( $\mu$ )** is irrelevant as changing the ranking between these two constraints would still yield *mi.sta* as the winner. To sum up, the overall ranking of the constraints needed to account for the syllable structure of loanwords in Sengwato can be stated as follows:

(43) **MIN- $\mu\mu$ , NO-CODA,  $*r[+ATR]\#$  >> MAX-IO, \*COMPLEX >> DEP-IO( $\mu$ )**

The ranking here has to be maintained in order to arrive at the desired optimal output; the following Table 6 and Table 7 show how the proposed hierarchy above in (43) derives the right forms for the words /mi.sta/ > English “mister” and /se.ta.ra.ta/ > Afrikaans Dutch /stra:t/ “street”:

**Table 6. Ranking for the Word /mi.sta/ > English “mister”**

	<b>mister</b>	<b>MIN-<math>\mu\mu</math></b>	<b>NO-CODA</b>	$*r[+ATR]\#$	<b>MAX-IO</b>	<b>*COMPLEX</b>	<b>DEP-IO(<math>\mu</math>)</b>
a.	mis.ter	*!	**				
→ b.	mi.sta				*	*	
c.	mis.ta		*!		*		
d.	mi.ste.re			*!		*	*

**Table 7. Ranking for the Word /se.ta.ra.ta/ > Afrikaans Dutch /stra:t/ “street”**

	<b>stra:t</b>	<b>MIN-<math>\mu\mu</math></b>	<b>NO-CODA</b>	$*r[+ATR]\#$	<b>MAX-IO</b>	<b>*COMPLEX</b>	<b>DEP-IO(<math>\mu</math>)</b>
a.	stra:t	*!	*			*	
→ b.	se.ta.ra.ta						***
c.	set.ra.ta		*!				**
d.	se.tra.ta					*!	**
e.	se.ta.ra			*!	*		**
f.	ste.ra.ta					*!	**

Note that in Tables 6 and 7, the hierarchical relationship between the different constraints is basically the same: the constraints **MIN- $\mu$** , **NO-CODA** and **\*r[+ATR]#** have to be ranked above **MAX-IO,\*COMPLEX** and **DEP-IO( $\mu$ )**. Crucial also is the ranking of **MAX-IO,\*COMPLEX** above **DEP-IO( $\mu$ )**. However, the ranking amongst **MIN- $\mu$** , **NO-CODA**, **\*r[+ATR]#** and amongst **MAX-IO,\*COMPLEX** is irrelevant.

### 3. Morphosyntax in Sengwato loanwords

This section focuses on a particular characteristic of Sengwato morphosyntax: the utilization of gender or noun classes to divide its lexicon. Gender in Sengwato is present in the form of prefixes (although some internal changes do occur) in the singular and plural, as well as in adjectival and verbal agreement.

#### 3.1 Gender Classes

A preliminary analysis of lexical items from the Swadesh list elicited with their plural forms revealed at least fourteen classes, seven in the singular form and their corresponding forms in the plural. The following illustrate these noun classes as well as an attempt to give a general description for each class:

##### A. Class 1: mo- > ba-

This class seems to encompass mostly human beings, including corpses:

- |                       |   |                   |
|-----------------------|---|-------------------|
| (44) mo-sadi “woman”  | > | ba-sadi “women”   |
| (45) mo-simanje “boy” | > | ba-simanje “boys” |
| (46) mo-to “person”   | > | ba-to “people”    |

##### B. Class 2: se- > di-

The informant’s intuition defined this class as mostly inanimate nouns, with a few exceptions:

- |                        |   |                    |
|------------------------|---|--------------------|
| (47) se-tunja “flower” | > | di-tunja “flowers” |
| (48) se-boko “worm”    | > | di-boko “worms”    |
| (49) se-tari “tree”    | > | di-tari “trees”    |

##### C. Class 3: bo- > ma-

This class contains abstract concepts, as the following examples illustrate:

- |                          |   |                      |
|--------------------------|---|----------------------|
| (50) bo-wpi “flour”      | > | ma-wpi “flours”      |
| (51) bo-lwetsi “disease” | > | ma-lwetsi “diseases” |
| (52) bo-tsello “life”    | > | ma-tsello “lives”    |

##### D. Class 4: Ø- > di-

This class seems to be the plural-forming default class. It constituted the highest amount of lexicon both in the Swadesh list and in the lexical borrowings list. Interestingly, it has no singular prefix:

- |                        |   |                       |
|------------------------|---|-----------------------|
| (53) poloxolo “animal” | > | di-poloxolo “animals” |
| (54) tsweni “monkey”   | > | di-tsweni “monkeys”   |
| (55) noka “river”      | > | di-noka “rivers”      |

**E. Class 5: le- > ma-**

In the examination of the words that belonged to this category, it was not possible to narrow down the semantic characteristics for this class. However, the notion of “foreign” seems to play an important role in some contexts, as will be expounded on in the following sections:

- (56) le-tsoho “hand” > ma-tsoho “hands”  
 (57) le-hoko “word” > ma-hoko “words”  
 (58) le-ru “cloud” > ma-ru “clouds”

**F. Class 6: mo- > me-/-eb-**

This class presents the lowest number of words both in the Swadesh list and in the lexical borrowings list. This class also seems to be very irregular, as some plural forms appear to undergo some internal change. No common semantic criteria can be found in the lexicon accounted in this class:

- (59) mo-riri “hair” > me-riri “hairs”  
 (60) mo-bu/mmu “soil” > me-bu “soils”  
 (61) m-muta “rabbit” > mi-bu:ta “rabbits”

**G. Class 7: Ø- > bo-**

This class contains personal names and their plurals, as well as creatures anthropomorphically modified such as in a story or fable. Similar to Class 4, this class becomes evident in the plural as it has no singular prefix:

- (62) marija “Maria” > bo-marija “Maria and company, Marias”  
 (63) tsweni “monkey” > bo-tsweni “monkeys in a fable”  
 (64) tawu “lion” > bo-tawu “lions in a fable”

*3.2 Noun Classes and Lexical Borrowings*

Following Corbett’s argument that meaning-based gender assignment must function as the core of form-based gender systems (Corbett 1991, p.30), it is suggested that the Sengwato gender assignment system is initially based on meaning. Section 3.1 attempted to identify some of these semantic criteria: humans for Class 1, inanimate nouns for Class 2, abstractions for Class 3, and persons and personified characters for Class 7.

Thus, it seems that form-based gender follows into the semantically-based gender categories. Thus, Sengwato nouns are being incorporated into each of the seven singular-plural classes in 3.1 depending on their form. Both meaning and form-based incorporation into Sengwato’s noun classification system are attested in the lexical borrowing data at hand. As presented in the previous sections, loanwords from English and Afrikaans Dutch undergo certain phonological and syllabic changes as they are incorporated into Sengwato. These morphosyntactic changes as observed are not random, but rather reflect how the Sengwato phonology and syllable structure operate when incorporating loanwords into the native lexicon. Corbett (1991) proposes that once loanwords have been phonologically and, in the particular case of Sengwato, syllabically, incorporated into the borrowing language, their gender

assignment follows that of other nouns in the language.

The following is a short analysis of the inclusion of lexical borrowings into the different Sengwato word classes. Focus is given to incorporation based on semantic content and on reinterpretation of loanwords' forms.

### 3.3 Incorporation due to Semantic Content

Perhaps the most straightforward approach to Sengwato's loanword incorporation into a noun class is determined by its meaning. Thus, nouns depicting human referents have two evident choices: the *mo-/ba-* class (humans, divinities, and corpses), and the  $\emptyset$ -/bo- class (personal names and personified creatures). Both choices are present in the data at hand. Consider the following examples from the *mo-/ba-* class in (65) and from the  $\emptyset$ -/bo- class in (66-68):

- |                  |               |   |                |
|------------------|---------------|---|----------------|
| (65) mo-porofiti | > ba-porofiti | < | Eng. "prophet" |
| (66) mista       | > bo-mista    | < | Eng. "mister"  |
| (67) misisi      | > bo-misisi   | < | Eng. "Mrs."    |
| (68) motlara     | > bo-motlara  | < | Eng. "model"   |

Notice, however, that (68) *motlara* "model" is not incorporated into the *mo-/ba-* class by reinterpretation of its initial *mo-* syllable. There seem to be several underlying cultural perceptions that play an important role in determining classes as will be discussed below. There is yet a third class that could involve humans, the *le-/ma-* class:

- |                   |   |           |          |
|-------------------|---|-----------|----------|
| (69) le-podisiele | > | ma-podisi | "police" |
|-------------------|---|-----------|----------|

It is interesting to note that initially most nationalities were included in this *le-/ma-* "stranger" category, for example:

- |               |   |          |   |                                       |
|---------------|---|----------|---|---------------------------------------|
| (70) le-fora  | > | ma-fora  | < | Afrikaans Dutch <i>Frans</i> "French" |
| (71) le-dache | > | ma-dache | < | Eng. "Dutch"                          |

Even though "French" and "Dutch" have human referents, they did not acquire the predicted *mo-/ba-* class prefixes, but instead were incorporated into the *le-/ma-* class. Even nationalities that could had been reinterpreted as having an inherent *mo-* prefix were modified in order to fit this class, hence (72) below:

- |                 |                          |   |            |                          |
|-----------------|--------------------------|---|------------|--------------------------|
| (72) le-zambiki | "person from Mozambique" | > | ma-zambiki | "people from Mozambique" |
|-----------------|--------------------------|---|------------|--------------------------|

The Sengwato informant in this study revealed that this mechanism is the result of the Sengwato perception of how the *mo-/ba-* class encompasses only Sengwato humans and not foreigners. This may also explain why (68) *motlara* "model" above was not interpreted as a *mo-/ba-* class noun due to its foreign provenance. The informant also added that in modern "politically-correct" Sengwato all nationalities would acquire the *mo-/ba-* prefixes and thus be included in the "human" category:

- |               |          |   |          |                 |
|---------------|----------|---|----------|-----------------|
| (73) mo-fora  | "French" | > | ba-fora  | "French people" |
| (74) mo-dache | "Dutch"  | > | ba-dache | "Dutch people"  |

This cultural perception is embedded deeply in the language, but it may also function backwards, that is, it may modify nationalities so that it includes them with the Sengwato people *mo-/ba-* class. Observe (75) for example:

(75) mo-sotho “a Lesotho person” > ba-sotho “Lesotho people”

Notice that even though “Lesotho” refers to a foreign nationality and it begins with the *le-* syllable, thus increasing its chance to be included in the *le-/ma-* class, it has been modified to fit the Sengwato notion of the familiar *mo-/ba-* class. The Sengwato informant explained that this particular incorporation is arguably due to the strong cultural ties between the Lesotho and the Setswana, which most residents in the region deem to be one people.

Because it was very hard to circumvent specific semantic criteria for most of the other noun classes, it was difficult to appreciate other incorporations due to semantic content in other nouns not depicting humans. There was, however, an instance of full incorporation into the *se-/di-* inanimate class:

(76) se-kamo > di-kamo < Eng. “comb”

However, these types of incorporation were very rare in the data list where most lexical borrowings seem to follow a reinterpretation pattern.

### 3.4 Reinterpretation of Form

The form of lexical borrowings after phonological and syllabic modifications was simply reinterpreted as containing a prefix which determined its noun class. This mechanism further reinforces Corbett’s (1991) assumption that modified loanwords function just like the native lexicon in the language.

Consider the following examples that were included in the *se-/di-* class:

(77) se-pune > di-pune < Eng. “spoon”

(78) se-tofo > di-tofo < Eng. “stove”

(79) se-pekere > di-pekere < Afrikaans Dutch *spijker* “nail”

(80) se-tarata > di-tarata < Afrikaans Dutch *straat* “street”

As observed in section 2.2, initial [strident+stop] clusters can include an epenthetic vowel in order to simplify complex onsets, and so the onset cluster in the word *spoon* is incorporated as *sepune*, with an epenthetic vowel /e/. The resulting *se-* syllable is therefore reinterpreted as the prefix *se-* of the *se-/di-* class. However, there is also the argument that perhaps *spoon*, *stove*, *nail*, and *street* were perceived as inanimate and therefore categorized in this class. One way to examine how reinterpretation occurs is by analyzing loanwords that do not share the general semantic criteria of their noun classes; for example, consider the following from the *bo-/ma-* class:

(81) bo-rotho > ma-rotho < Afrikaans Dutch *brood* “bread”

(82) bo-roxo > ma-roxo < Afrikaans Dutch *brug* “bridge”

(83) bo-rokwe > ma-rokwe < Afrikaans Dutch *broek* “trousers”

Note that the examples above all contain onset clusters which when incorporated into Sengwato were simplified via epenthesis of the vowel /o/. As such the initial prefix *bo-* which resulted from

resyllabification was reinterpreted as the *bo-* prefix of the *bo-/ma-* class for abstract concepts. Notice that the fact that *bread*, *bridge*, and *trousers* have no relation with abstract concepts strengthens their inclusion by reinterpretation and not by semantic content. The same mechanism may be present in (84) from class 6 below:

(84) mmishini > mebishini < Eng. “mission”

The assumption that initial syllables are reinterpreted as noun class prefixes is therefore well attested. However, there are other instances observed in the data list where it is not the apparent presence of a prefix what triggers noun classification, but the absence of it.

### 3.5 Reinterpretation vs. Incorporation

The  $\emptyset$ /*di-* class (class 4) contains the most numerous lexical items in both the Swadesh list and the lexical borrowings data. Therefore, the definition of this class as a default or residue class seems to best fit its criteria (Corbett, 1991, p. 35). Because this class encompasses a wide gamma of concepts and because it requires no initial prefix (or the reinterpretation of such), it is difficult to assert whether lexical borrowing inclusion in this class is due to the absence of prefixes or because their semantic content does not fit into any of the semantic criteria of the other classes. The following variety of examples illustrate this wide range of concepts:

(85) apole > di-apole < Eng. “apple”

(86) buka > di-buka < Eng. “book”

(87) xalasi > di-xalasi < Afrikaans Dutch *glas* /xlas/ “glass”

(88) poso > di-poso < Afrikaans Dutch *post* “mail”

Furthermore, there seems to be no systematic assignment of which lexical items will be incorporated in terms of their semantic content and which will be reinterpreted as containing noun class prefixes. Observe the minimal pairs (89) and (90) below:

(89) boroto > di-boro < Eng. “board”

(90) bo-rotho > ma-rotho < Afrikaans Dutch *brood* “bread”

The lexical borrowing for *board*, even though it underwent resyllabification and therefore it contains the initial *bo-* syllable, is not assigned into the *bo-/ma-* class like the word for *bread*. These words have identical structures, and there is no apparent reason to differentiate between them. In this particular case, the system for deciding integration by semantic content or by lexical form reinterpretation appears to be random.

### 3.6 Other Relevant Noun Class Features

The noun classes of Sengwato also seem to display a general vs. specific semantic component. That is, the generalization or specification of a noun can be determined by different noun suffixes. For example, let us return to the *board* vs. *bread* example discussed earlier.

As observed, *boroto* “board” was not reinterpreted as \**ma-roto* in the plural like the *bo-rotho* > *ma-rotho* “bread/breads” pattern. The *ma-* prefix, however, can replace the *di-* prefix in the plural of



*boroto* “board”, and yield a different semantic context as such:

(91) *boroto* > **di-***boroto* “some general boards”

(92) *boroto* > **ma-***boroto* “specifically blackboards”

This general vs. specific association with prefixes is not extended to all nouns. For example, (93) *karata* > *di-karata* below is also from Class 4 ( $\emptyset$ -/*di*-) like *boroto*, but the prefixes do not work similarly:

(93) *karata* > **di-***karata* “cards”

(94) *karata* > \***ma-***karata* “cards” (ungrammatical)

Other general vs. specific prefix relations seem to involve class-crossing. Recall that the semantic criteria of Class 7 ( $\emptyset$ -/*bo*-) was composed mainly of people, proper names, and personified creatures.

(95) *porofese* > **bo-***porofese* “professor/s”

However, the native Sengwato informant in this study pointed out that the use of the plural *bo-porofese* was very specific, as in a professor specified with a name. In order to generalize “professor”, the prefix *di-* is attached instead:

(96) *porofese* > **di-***porofese* “professor in general”

Interestingly, class-crossing from 7 ( $\emptyset$ -/*bo*-) to Class 4 ( $\emptyset$ -/*di*-) implies a less degree of specification. However, like the *boroto* vs. *karata* Class 4 examples above, the same prefix relations are not followed. For example, (97) *titfara* “teacher”, also from Class 7 like *porofese*, has a different combination of prefixes:

(97) *titfara* > **bo-***titfara* “teacher/s”

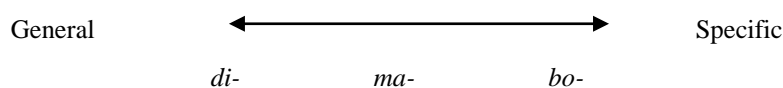
Like *bo-porofese*, the plural of “teachers” *bo-titfara* was perceived by the informant as being too specific. However, the generalized form does not take the *di-* prefix like *bo-porofese*, but it takes the *ma-* prefix:

(98) *titfara* > **bo-***titfara* “specific teachers”

(99) *titfara* > \***di-***titfara* (ungrammatical)

(100) *titfara* > **ma-***titfara* “very general teacher”

Note that this pattern seems to be a crossing from Class 7 ( $\emptyset$ -/*bo*-) to Class 5 (*le*-/*ma*-). Most interestingly, there seems to be a hierarchy of prefixes and class-crossing. For example, the change of *boroto* to *ma-boroto* yields a more specific definition, while the change of *bo-titfara* to *ma-titfara* yields a more general definition. Thus, general vs. specific not only depends on the prefixes, but also on the original noun class:



**Figure 4. Hierarchy of Prefixes in the Incorporation of Loans in Sengwato**

The  $\emptyset$ -/bo- noun class has two ways to be generalized in the plural, it can either take the **ma-** prefix such as *titfara* “teacher” > *bo-titfara* “specific teachers” > *ma-titfara* “general teachers” or it can take the **di-** prefix such as *porofese* “professor” > *bo-porofese* “specific professors” > *di-porofese* “general professors”. However **ma-** vs. **di-** is not optional and it is noun-specific. The failure to assign the right prefix would yield ungrammaticality as in (99) *\*di-titfara*.

The  $\emptyset$ -/di- class has an inherently generalized plural, but it can change to more specific by acquiring the **ma-** prefix, which according to the hierarchy in Figure 4 is more specific than **di-**, but more general than **bo-**. This scale can be illustrated by the following forms: *boroto* “board” > *di-boroto* “general boards” > *ma-boroto* “specific boards, such as blackboards”. Apparently, in this class the prefix choice is not free, but rather noun-specific, as the ungrammatically of *\*ma-karata* in (94) demonstrates.

#### 4. Conclusion

This short paper examined a well-known dialect of the Setswana language, Sengwato. The modification of lexical borrowings, largely from English and Afrikaans Dutch, as incorporated into Sengwato reveals many of the language’s native phonological, syllabic, and morphosyntactic characteristics. Phonological adaptation unravels a number of sound change processes which are best captured in Sengwato by the ruled-based and Optimality-Theoretic framework of analysis. Lexical adaptation illustrates the different steps of lexical incorporation proposed by Corbett (1991), in which loanwords are first phonologically (and syllabically) modified before they interact with the language’s morphosyntax. Although this short sample of the Sengwato morphosyntax provides an insight on how lexical borrowings are fully introduced into the language by acquiring noun classes and utilizing their specific vs. general spectrum, more data are needed to visualize the entire morphosyntactic system of Sengwato and how it interacts with the seven singular and seven plural classes.

#### References

- Barlow, J. (2001). The structure of /s/-sequences: Evidence from a disordered system. *Journal of Child Language*, 28, 291-324. <https://doi.org/10.1017/S0305000901004652>
- Borden, G., Katherine, H., & Lawrence, R. (2003). *Speech science primer: Physiology, acoustics and perception of speech*. Lippincott Williams and Wilkins.
- Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, 7, 35-59. <https://doi.org/10.1177/026765839100700102>
- Broselow, E. (1993). Transfer and universals in second language epenthesis. In S. Gass, & L. Selinker (Eds.), *Language Transfer in Language Learning* (pp.71-86). Amsterdam: John Benjamins.
- Cairns, C., & Feinstein, M. (1982). Markedness and the theory of syllable structure. *Linguistic Inquiry*, 13, 193-225.

- Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston, & M. Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 283-333). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627736.017>
- Clements, G., & Keyser, S. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
- Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University. <https://doi.org/10.1017/CBO9781139166119>
- Giegerich, H. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139166126>
- Greenberg, J. (1978). Some generalization concerning initial and final consonant clusters. In J. Greenberg (Ed.), *Universals of Human Language: Phonology* (pp. 243-279). Stanford, CA: Stanford University Press.
- Ito and Mester. (2003). *Systematic markedness and faithfulness*. Retrieved November 11, 2018, from [http://people.ucsc.edu/~mester/PAPERS/ito\\_mester\\_cls39.pdf](http://people.ucsc.edu/~mester/PAPERS/ito_mester_cls39.pdf)
- Kager, R. (1999). *Optimality theory*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511812408>
- Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford, UK: Blackwell.
- Prince, A., & Smolensky, P. (1993) *Optimality theory: Constraint interaction in generative grammar*. The Rutgers Optimality Archive, ROA-537-0802. Retrieved November 13, 2017, from <http://roa.rutgers.edu>
- Roach, P. (2000). *English phonetics and phonology*. Cambridge University Press.
- Rubach, J. (2002). An overview of palatalization-*i*\*. *Studies in Phonetics, Phonology and Morphology*, 8(2), 169-186. Retrieved October 20, 2018, from [http://society.kordic.re.kr/~pmc/work/vol\\_08\\_2/01RubachJ.pdf](http://society.kordic.re.kr/~pmc/work/vol_08_2/01RubachJ.pdf)
- Selkirk, E. (1982). The syllable. In H. Van der Hulst, & N. Smith (Eds.), *The Structure of Phonological Representation* (pp. 337-384). Dordrecht: Foris Publications.
- Spencer, A. (1996). *Phonology: Theory and description*. Oxford: Blackwell Publishers Ltd.
- Trask, R. L. (1996). *Historical linguistics*. London: Edward Arnold.
- Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, & D. Steriade (Eds.), *Phonetically Based Phonology* (pp. 34-57). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486401.002>