

ON THE ORTHOGRAPHY OF URHOBO-ENGLISH LOANWORDS

Philip Oghenesuowho Ekiugbo

National Institute for Nigerian Languages, Aba, Nigeria
oghenesuowho@gmail.com

Abstract

This study examines the orthographic representation of Urhobo-English loanwords in the literature. Following Hosken (2003), one of the hardest issues any orthography has to deal with is how to spell words that are imported from another language: should the words keep their original spelling, or should they be spelled according to how they sound when spoken in the target language? This issue is not unconnected with the fact that loanwords are often modified, especially when they do not conform to the phonology of the borrowing language. Additionally, sounds that are similar in both source and target languages may be represented with different graphemes in the two languages. This study, therefore, seeks to examine the various modifications that loanwords undergo in Urhobo and how these can be best represented in the orthography, as well as how similar sounds that are represented by different graphemes are handled in the borrowing language. The data employed in the study were extracted from religious and literacy materials in Urhobo and evaluated by the researcher, relying on native speaker competence. The data analysis focused on the phonological nuances of loanwords, which have implications for Urhobo orthography, guided by the principle of consistency. The study made a case for the representation of the adapted forms in Urhobo. This is to ensure fluency, word recognition, and disambiguation in the orthographic reading of these words.

Keywords: Loanword orthography, Consistency, Reading, Phonological recoding

1. Introduction

This study examines the orthographic representation of Urhobo-English loanwords in the literature. The earliest known orthographic representation of the Urhobo language includes Kilham (1828), Norris (1841), Clarke (1848), Koelle (1854), and Latham (1862). These works were mainly wordlists of several African languages or dialects “as spoken by freed slaves in Freetown, Sierra Leone” (Rolle, 2013:282). Following Hair (1992:1-2), Europeans wrote down occasional terms in African languages from their first arrival on the coasts of Black Africa, and from the sixteenth century onward, short lists of useful words and phrases in a few of the many West African languages found their way into print. However, the orthographies utilised at that period were primarily based on the authors' various European languages. Subsequent writings in Urhobo, such as Thomas (1910), relied on Lepsius' (1863) orthography, which was one of the most harmonised and commonly used writing systems in Africa and Asia during the pre-colonial and early-colonial eras. The Lepsius system was supplanted, first, by the practical orthography designed by James (1929), and then by the orthographies established for the Urhobo language (see Basu, 2020).

As noted by Ivworin (2018), missionaries were the first to make an attempt to construct a uniform orthography and write the Urhobo language in the 1910s. However, the standard dialects targeted in the orthographies produced by missionary groups differ. The Agbarho form was adopted by the Christian Missionary Society. According to Nabofa (1997), its adoption stems from a meeting held in Effurun in 1914 to translate Christian literature from English to Urhobo and to settle the

controversy over the various spoken forms or dialectal renderings of the Lord's Prayer, which was largely due to a lack of standardised translation at the time. The Agbon dialect was adopted by the Baptist mission and other groups (see Kelly, 1969a, b; Welmers, 1969). It was perhaps the formation of the joint consultative Urhobo Bible translation committee, which included members of various missionary groups, and the Urhobo Progress Union's formation of an Urhobo language committee that resulted in agreement on the standard dialect that written Urhobo literature should target. The Urhobo Progress Union attempted to develop orthography for the Urhobo language by forming a language committee in 1948 as part of an effort to promote a standard style of literacy in Urhobo writing and reading. The committee, which was revived in 1952, contributed immensely to the standardisation of the Urhobo orthography.

Although Umukoro (1999:7) notes that “most books in the Urhobo language are written using the Olomu dialect while news broadcasts and translations in Urhobo language are relayed using the Agbarho dialect”, there is no evidence to substantiate this at the moment. Iweh (1987) also notes that the Urhobo Language Orthography Committee chose Agbarho and Agbon dialects as the standard for the formal writing of the Urhobo language. According to Iweh (1987:58), “the reason given by the Urhobo Language Orthography Committee for specifying the two groups- instead of Agbarho alone, which was the variety normally regarded as the standard or central language- was the desire to underscore the fact that the Agbon variety does serve as a source of vocabulary items which may be found wanting or inappropriate in Agbarho”. He further states as an example of inappropriation in Agbarho, that “the Agbarho word “áǵáǵá” “matchet” is undesirably similar to “àǵàǵà” “crotch”, in which case, the Agbon word for matchet “òǵà” is therefore preferred.

The Urhobo orthography is phonemic. Thus, it involves representing the

significant sounds of the spoken language with written symbols in a consistent and systematic way. Phonemic orthography is based on the hypothesis that a writing system should ideally have a one-to-one correspondence between sounds (phonemes) and the symbols used to represent them (graphemes). This principle aims to create a unique, straightforward, and intuitive relationship between spoken language and its written representation. Accordingly, given that the Urhobo language has seven vowels, each with oral and nasal counterparts (Rolle 2013; Aziza, 2002), and thirty consonants (Kelly 1969a), the graphemes in the current orthography of the language are as shown in Example (1) below.

	Phoneme	Grapheme		Phoneme	Grapheme
a.	/p/	<p>	w.	/m/	<m>
b.	/b/		x.	/n/	<n>
c.	/t/	<t>	y.	/ɲ/	<ny>
d.	/d/	<d>	z.	/ŋm/	<mw>
e.	/c/	<ch>	aa.	/r/	<r>
f.	/ɟ/	<dj>	bb.	/ɾ/	<rh>
g.	/k/	<k>	cc.	/j/	<y>
h.	/g/	<g>	dd.	/w/	<w>
I.	/kp̄/	<kp>	ee.	/i/	<i>
j.	/gb̄/	<gb>	ff.	/e/	<e>
k.	/φ/	<ph>	gg.	/ε/	<ē>
l.	/v/	<vw>	hh.	/a/	<a>
m.	/f/	<f>	ii.	/ɔ/	<ɔ̄>
n.	/v/	<v>	jj.	/o/	<o>
o.	/s/	<s>	kk.	/u/	<u>
p.	/z/	<z>	ll.	/ĩ/	<in>
q.	/ʃ/	<sh>	mm.	/ẽ/	<en>
r.	/ʒ/	<j>	nn.	/ɛ̃/	<ɛ̄n>
s.	/x/	<h>	hh.	/ã/	<an>
t.	/ɣ/	<gh>	oo.	/õ/	<on>
u.	/x ^w /	<hw>	pp.	/õ/	<on>
v.	/ɣ ^w /	<ghw>	qq.	/ũ/	<un>

Iweh's (1987) study shows the different modifications that the Urhobo orthography has undergone. One of the overarching principles and guiding forces in the modifications that the Urhobo orthography has undergone is the need for consistency. This notwithstanding, Elugbe (1990) identified some instances of inconsistency in the Urhobo orthography. Elugbe's study focused attention on the vowels of Urhobo both in terms of analysis and in terms of orthographic practice. He notes that phonetically, there are seven oral and seven nasalized vowels in Urhobo. He notes that phonetically, there are seven oral and seven nasalized vowels in Urhobo. He further demonstrated that the vowels /e, o/ have dual behaviour, that is, they behave as high vowels in some instances, which he termed e_1/o_1 and as non-high vowels in other instances, which he termed e_2/o_2 (See also Aziza, 1994, 2008; Cf. Ekiugbo and Ugorji, 2019); but that native speakers are found to favour and be at home with the current orthography, which is based on a seven vowel system. Attempts to be accurate in Okpe, a system similar to Urhobo, proved difficult and unpopular with native speakers. Thus, the two instances of [e, o] are under-differentiated, that is, both instances are represented with a single grapheme.

A preliminary observation of the orthographic representation of Urhobo-English loanwords also shows some inconsistency in the orthographic representation of loanwords in Urhobo. The challenges of representing loanwords orthographically have been observed in some languages (Tian, 2012; Hannas, 1997; Edward, 1994). The challenges often arise when loanwords from one language are incorporated into a language that has different phonemic and/or orthographic conventions. However, these challenges have not been given much attention in the literature. The present study seeks to examine the case in Urhobo. The lexicon of Urhobo is replete with loanwords, mainly from the English language but also from Portuguese, Yoruba, Edo, Igbo, etc. (Ojaide and Aziza, 2014; Darah and Ekiugbo, 2014). The present study is concerned with how loanwords of English origin are represented in the literature for three reasons. Firstly, most of the loanwords in Urhobo are of English origin. Secondly, the

loanwords from Portuguese have been fully nativized both in pronunciation and spelling, as evident in such words as oro (Ouro) 'gold', ukujere (colher) 'spoon', etc. Lastly, most loanwords from Nigerian indigenous languages need no modification, as these languages share many similar sounds and structures, as evident in examples such as akpu (Igbo), agbero (Yoruba), etc.

2. Literature Review

Borrowing is a cross-linguistic phenomenon. This is because all languages incorporate words from other languages into their vocabularies. However, the fact that languages differ in their phonologies opens up the possibility that when words are borrowed, they may 'violate' the phonology of the borrowing language; that is, some of the array of segments, suprasegments, and structures of loanwords may be illicit to the borrowing language. Thus, the differences in phonology between languages have implications for loanwords. According to Alquarni (2021), given that phonological principles and constraints differ cross-linguistically, a loanword with ill-formed segmental, phonotactic, suprasegmental, or morphophonological content may result in violations. He also noted that these violations require that the borrowing language make repairs and modifications to the non-native segment or structure to be in compliance with its phonology, either by mapping the ill-formed segment onto the same or onto the closest sound in its inventory. Cross-linguistic studies continue to demonstrate that loanwords typically go through adaptation processes to adhere to the phonological constraints of the borrowing language (Alquari, 2021; Al-Athwary, 2017; Ugorji, 2013).

Adaptation affects all aspects of the phonological structure of the loanwords, including the segmental, phonotactic, and prosodic. A number

of studies have examined the phonology of Urhobo-English loanwords. Aziza and Utulu's (2006) study on the strategies for adapting sound segments and syllables of English loanwords in Urhobo (and Yoruba) shows that the adaption of segments does not occur in the two languages if the sounds in the loanwords are phonetically and phonologically similar to the sounds in the native phonology, but where they are not, the English sounds are modified to the approximate perceptually closest native sounds. Accordingly, the English consonants /tʃ, θ, ð, l, ŋ/ are adapted: /tʃ/ is realized as [ʃ]; /θ/ as [t] in word initial position and [f] in word final position; /ð/ as [d], clear /l/ as [n] in initial position, as [r] when it follows a labial/velar consonant or its interrupted through vowel insertion in CC context; dark /l/ as Ø, vocalized to [o] or modified to [r]; and /ŋ/ as Ø, with the nasal feature realized on the preceding vowel. With regard to vowels, the study notes that only /ɛ/ is not adapted in Urhobo, /ɪ, i:/ are realized as [i], /a:, æ/ as [a], /ɜ:/ as /ɛ/, /ɔ:, ʌ, ɒ/ as [ɔ], /u:, ʊ/ as [u], /iə, eə/ as [iɛ], while the other diphthongs are changed to monophthongs because they lack phonetic similarity in the phonological grammar of Urhobo.

Aziza and Utulu (2006) also note that deviant structures of English loanwords in Urhobo are modified using consonant deletion, vowel insertion, or a combination of both. The repairs are motivated by the need to obligatorily fulfil the two constraints that all syllables entering into the Urhobo language (and Yoruba) must observe: (i) no (illicit) complex onset, and (ii) no coda. Accordingly, all deviant structures are repaired through vowel insertion or consonant elision: a complex onset is simplified either by deleting one of the consonant or more commonly by inserting vowel /i/ between the two consonant; a noun beginning with a C element is repaired by the use of a prothetic vowel, given that the language requires that nouns should begin with vowels; while coda are repaired by

either elision or epenthesis. They further argued that the prothetic vowel is /i/ while the epenthetic vowels are /I, u/.

Similar propositions are presented in Ugorji (2013), Bolaji (2016), Ukeagbu, Odeh and Nwosu (2022), Ekiugbo and Eme (2023), etc. Although there are variations in the arguments and analyses in these studies, the central position is that illicit sounds or structures are modified in line with the phonology of Urhobo. While the conformation of loanwords to the phonological demands of target languages is well established in the literature, much is not known about the systematic representation of loanwords in the orthography of borrowing languages. Two broad patterns may be noted in the orthographic representation of loanwords. These are adoption and adaptation (Koffi, 2010). The former involves borrowing the spelling of the loanword from the source language with little or no modification, whereas the latter involves adapting loanwords in line with the phonological modification of the borrowing language. Coulmas (1989:174) notes that “the principle of spelling loan words in English is to preserve the source language orthography but pronounce the orthographic word as if it were English”, whereas, “the principle of Spanish orthography, by contrast, is to preserve the 'sound' of foreign word in so far as Spanish phonology permits this” (Coulmas, 1989:174).

Jany's (2010) study on the orthography of loanwords of Spanish origin in Chuxnabán-Mixe, a Mixean language spoken in Mexico, also shows that Spanish loans are represented in their original orthography, except for accents as well as Mixean morphological and phonological adaptations. Thus, languages differ in their orthographic rendering of loanwords. As an example, while the classic case in English is to adopt the

loanword as it is in the source language, Spanish adapts its orthographic forms in line with its phonological adaptation, whereas Chuxnabán-Mixe employs both strategies (partial adaptation). For Coulmas (1989:174), “by adopting loan words in their original spelling, orthographies tend to become heterogeneous, incorporating different sets of rules that apply to different parts of the lexicon”. Adoption often results in inconsistencies. For instance, the same letter does not always represent the same sound in English, as some letters can stand for as many as four different sounds (Umera-Okeke, 2008). This will be problematic for phonological recoding in reading.

When reading, individuals decode written words by converting the visual symbols (letters or graphemes) into their corresponding sounds or phonemes, which enables them to effectively recode the text into spoken language. Phonological recoding, therefore, refers to the process of converting written language into its corresponding sounds, or phonemes. It is an essential component of reading (Knoepke et al., 2014), especially in alphabetic writing systems where the relationship between letters and sounds is systematic. Phonological recoding allows readers to decode words by using their knowledge of letter-sound correspondences. Thus, the understanding that letters in written language represent sounds in spoken language is the foundation of phonological recoding. This, therefore, requires a grapheme-to-phoneme correspondence in orthographic design. Where there is a mismatch in grapheme-to-phoneme correspondence, learners may experience difficulties in converting written words into their corresponding sounds or phonemes.

According to Koffi (1990:197), “adoption creates many problems for the orthographic system of the borrowing language”. One of such

problems is the introduction of “spellings that run counter to existing patterns and expectation(s)” into the writing system of a language (Cummings, 1988:20). Koffi (2010) also notes that heterogeneous orthography (which is one of the results of adoption) violates the Simplicity Maxim in orthography design. The simplicity maxim in orthography is a principle that suggests that the writing system of a language should be as simple and straightforward as possible, with a one-to-one correspondence between sounds and letters. Accordingly, Garvin (1972:421) argues that “the spelling of foreign words, particularly common ones, should not follow an orthographic system different from that used for domestic words... Likewise, proper names which have not been nativized can clearly be spelled in line with the original system”. The main position here, which this study also supports, is that “orthographies must pay close attention to loanword phonology” (Koffi, 2010:8).

3. Methodology

The data employed in this study consist of adapted loanwords of English origin extracted from secondary sources. The sources include the Urhobo language Bible (henceforth ULB), the Urhobo prayer book (Society for Promoting Christian Knowledge, 1960; henceforth UPB), the Urhobo language scheme of work (henceforth USW), Ojaide and Aziza (2014), and Ajiboye et. al (in preparation; henceforth ULR). The analysis of the data seeks to examine the extent to which the loanwords follow the orthographic principles and implications for phonological recoding. The principle of consistency in orthography design (Williamson, 1984) requires that a writing system should have consistent and predictable relationships between letters or graphemes and their corresponding sounds. Following Williamson (1984:1), in developing a written form of a language, our aim should be to make the written form agree closely with

the spoken form. If this is well done, we should be able to read the language just as easily as we understand the people speaking it.

4. Data Presentation and Analysis

Urhobo orthography design is based on a set of propositions and guiding factors aimed at ensuring consistency and accuracy in the representation of the language in written form. The main principles are the principles of phonemic representation and consistency. The former implies that each letter or combination of letters represents a specific sound in the language, while the latter requires that the Urhobo orthography is consistent in its representation. This section presents a discussion of how loanwords are represented orthographically in the Urhobo literature in relation to the principles of phonemic representation and consistency. We focus on three critical areas that we think orthography of loanwords should consider. The first relates to how similar sounds represented by different graphemes in the two languages are represented in the target language. The second area relates to how redundant letters are handled in the orthography of the borrowing language, while the third is on how nativised forms are represented.

4.1 Representation of Similar Sounds having Different Graphemes

Most orthographies show a one-to-one correspondence between phonemes and graphemes (Jany, 2010). However, such grapheme representations are language-specific. Thus, a single sound may be represented with different graphemes in different languages or dialects. This is the case with Urhobo loanwords of English origin. For instance, sounds such as /ɛ/, /ɔ/, /k/, /f/, /s/, and /ʒ/ are represented differently in the two languages. In this section, we examine how loanwords with such sounds are represented orthographically in Urhobo.

(2)	a. Corinth	Kọrẹnt	(ULB)
	b. Galician	Galesha	“
	c. Genesis	Jẹnẹsis	“
	d. Amos	Emọs	“
	e. Zephaniah	Zẹfanaya	“
	f. Cherubim	chẹrobim	(UPB)
	g. incense	inces	“

The orthographic representations of these sounds, as shown in example (2) above, indicate that instances of similar sounds represented by different graphemes in the two languages are handled differently in the Urhobo literature, with some retaining the source graphemes and others modifying them. The examples in (2a-e), which are drawn from ULB, show instances in which the loanwords of English origin sharing similar sounds with Urhobo are represented using the Urhobo grapheme for the sounds, whereas the examples in (2f-g), drawn from UPB, retain the English graphemes. Thus, while there is a one-to-one correspondence between letters and sounds in ULB, it is not the same in UPB. Implicitly, both adoption and adaptation are employed in the representation of sounds having different graphemes. Thus, there is an inconsistency, which can result in reading difficulty. Inconsistencies of this kind can result in learners missing pronunciation targets by making errors such as sound additions, deletions, and substitutions (Bassetti, 2024). For instance, representing incense as *inces* may be read as [ikesi] instead of [is̄esi]. Besides the challenge of phonological recoding, adoption may also result in challenges related to the non-integration of loanwords into the language.

In the Table below, the graphemes used in representing some sounds found in English and Urhobo are shown. The first column gives the IPA representation of the sounds. This is followed by the graphemes used in the English language to represent the sounds. The third column shows the equivalent Urhobo graphemes for the sounds, while the last column shows the sound that the English graphemes represent in

Sound (IPA)	English Grapheme	Urhobo Grapheme	Urhobo Sound for English Grapheme
a. /ɛ/	<e>	<ẹ>	/e/
b. /ɔ	<or>, <o>	<ọ>	-, /o/
/c. /k/	<k>, <c>, <ch>	<k>	/k/, -, /c/
d. /g/	<g>, <gh>	<g>	/g/, /ɣ/
e. /f/	<f>, <ph>	<f>	/f/, /ɸ/
f. /s/	<s>, <c>	<s>	/s/, -
g. /ʃ/	<sh>, <si>	<sh	/ʃ/, -
h. /z/	<z>, <s>	><z>	/z/, /s/
I. /ʒ/	<g>, <s>, <j>	<j>	/g/, /s/, /j/

Table 1: Grapheme differences between English and Urhobo

UrhoboTable (1) above indicates that some sounds may be represented differently in the orthographies of English and Urhobo. For instance, the mid-low front vowel /ɛ/ in (a) is represented with <e> in the English language, while the Urhobo language represents the same sound with <ẹ>; whereas the English grapheme for this sound is used to represent the mid-high vowel /e/ in Urhobo. Thus, this sound is represented differently in the orthographies of both languages. Similarly, the voiceless velar

plosive is (c), is represented with three graphemes in English, <k> (as in *kettle*, *kite*, etc) which is the same grapheme used in Urhobo (as in *oka* 'mark', *kare* 'carve', etc), <c> (as in *computer*, *camera*, etc) and <ch> (as in *chemical*, *architect*, etc), which is the grapheme used in Urhobo to represent the voiced alveolar plosive, /c/. In Urhobo, the similar letter-to-grapheme correspondence as in the first graphemes in (c) – (h) above is retained, while the others are adapted, as shown in the data in Example (3) below, taken from Ojaide and Aziza (2014:118).
EnglishUrhoboAdaptation
 3a. Computerikɔputa<c> adapted to <k>
 b. PhoneIfonu<ph> adapted to <f>
 c. RadioIredio<a> adapted to <e>
 d. Morphologyimɔfɔloji<g> adapted to <j>; <y> adapted to <I>

4.2 Redundant Letters

The English language has redundant letters in its orthography, whereas the Urhobo language does not. Redundant letters refer to the use of letters in words that do not contribute to the pronunciation or meaning of the word. In some cases, these letters may have been used historically to indicate a particular pronunciation or derivation of the word. One implication of redundant letters in English is that it makes its spelling complex and inconsistent. These redundant letters can make it challenging for learners to determine the pronunciation of words based solely on their written form (Umera-Okeke, 2008). In this section, we examine how English redundant letters of Urhobo-English loanwords are handled.

(4)	a. John	Jɔn	(ULB)
	b. knit	nit	(ULC)
	c. knight	nait	(UPB)
	d. exhibit	ɛzibiti	(ULM)
	e. foreign	ifɔrini	(ULM)
	f. muscle	imɔzo	(ULM)

As may be inferred from Example (4), all the instances of English loanwords in Urhobo bearing redundant letters in the source language are adapted in the orthography of the target language. Thus, there is consistency in the representation of redundant letters of Urhobo-English loanwords, as they are all omitted. Such omission allows for reading fluency by preventing reading adaptation, which may obscure the meaning of the words. In reading adaptation, loanwords are pronounced based on the spelling of the source word, rather than its actual pronunciation. For example, the French word *cul-de-sac* is adapted into English as [kəldə'sæk], which is a reading adaptation of the spelling rather than the French pronunciation (Vendelin and Peperkamp, 2006).

4.3 Representation of Nativized Forms

This sub-section gives an overview of the areas of repair in loanwords of English origin in Urhobo and how each is represented orthographically in Urhobo. There are at least six areas in English words that are modified when borrowed into Urhobo (see Aziza and Utulu, 2006; Ugorji, 2013; Okolo-Obi, 2015). These are: illicit segment, stress, nouns with initial consonants, illicit consonant cluster, coda, and illicit peak/nucleus. In what follows, these aspects are discussed, albeit with the exception of stress adaptation. Urhobo is a tonal language. Tone plays both lexical and grammatical roles in the language. Thus, all the stressed syllables of English words are adapted with tones in a systematic way (see Okolo-Obi, 2015). However, tones are often not marked in Urhobo language materials; thus, it is difficult to get data on this aspect. Accordingly, this has been left out of this discussion.

4.3.1 Illicit Segments

An illicit element in linguistics refers to any unit that is not attested in a given language. In phonological context, it refers to segmental or suprasegmental units that do not exist in the phonology of a language. Loanwords of English origin are likely to have segments (and suprasegments) that are not attested in the Urhobo language. Such elements often violate the constraints of the borrowing language, which may necessitate repairs. Although this has been noted in the literature (Aziza and Utulu, 2006; Ugorji, 2013), how they are represented orthographically in Urhobo has not been discussed. The English consonants /tʃ, dʒ, θ, ð, ŋ/ and monothong vowels /i:, ɪ, æ, ə, ɜ:, ɔ:, ʌ, ɒ, u:, ʊ/ are not attested in Urhobo. Thus, they are often adapted to phonetically similar sounds, as shown in Table (2) below. The table also shows the source orthography, the grapheme used in representing the sound of the adapted form, and the target language sound represented by the source grapheme. Examples extracted from the literature are also shown in (5) below..

Source Form	Adapted Form (AF)	Source Orth.	Orth. Of AF	Sound represented by source orth.
a. /tʃ/	/ʃ/	<ch>	<sh>	/c/
b. /θ/	/t/	<th>	<t>	-
c. /ð/	/d/	<th>	<d>	-
d. /ɪ/	/I/	<i, ee, y>	<i>	/i, -, -/
e. /a:, æ/	/a/	<a, ar, er>	<a>	/a, -, -/
f. /ə/	/a/	<a>	<a>	/a/
g. /ɜ:/	/ɛ/	<a, e, er, ...>	<ɛ>	/a, e, -/
h. /ɒ/	/ɔ/	<o>	<ɔ>	/o, -, au/
i. /ɔ:/i. /u:, ʊ/	/u/	<or, oa, aw, ...>	<u>	/-, -, -, .../
j. /ʌ/	/ɔ/	<u, oo, u...e>	<ɔ>	/u, -//u/

Table 2: The sounds and graphemes of illicit segments of English loanwords in Urhobo

(5)	a. Church	Ishoṣhi	[O & A]
	b. Ananias	Ananayas	[ULB]
	c. Esther	Esta	“
	d. Thessalonica	Tɛsalonaika	“
	e. Titus	Taitoṣ	“
	f. litany	litani	[UPB]

It is often the case that languages only permit 'native' features in their lexical formatives. The extant literature has shown that when a loanword has an illicit (supra-) segment, it is often substituted with a phonetically similar unit in the borrowing language. The orthographic representations of such forms in Urhobo, as in Example (5) above, show that the words bearing such adapted sounds are represented with the grapheme of the target language in all of the instances in our data corpus. Such orthographic adaptation is necessary, given that it will ease phonological recoding, which will result in ease of reading.

4.3.2 Nouns with Initial Consonants

The morpheme structure condition in Urhobo and other Edoid languages is such that all nouns must begin with a vowel (Aziza and Utulu, 2006). This is, however, not the case with English, in which nouns may begin with consonants or vowels. When nouns beginning with consonants are borrowed, they are often adapted by inserting /i/ at the initial position of the words, as illustrated in Example (6) below. Examples showing how this is represented orthographically in Urhobo-English loanwords are shown in (7).

(6)	a. brɛd	[i.brɛ.di]	'bread'
	b. pɔt	[ipɔtu]	'pot'
	c. ʃaɪn	[ʃai.ni]	'shine'

	d. film	[i.fi.mu]	'film'
	e. bAlb	[ì.bó.bù]	'bulb'
	f. belt	[i.be.ti]	'belt'
(7)	a. Leviticus	Livitikos	[ULB]
	b. Joshua	Joshua	“
	c. Samuel	Samuèl	“
	d. Minister	minista	[UPB]
	e. tailor	itenọ	[O & A]
	f. nurse	inọsu	“
	g. powder	ipọda	“
	h. fuse	ifiuzu	“
	i. phone	ifonu	“
	j. rubber	iroba	[ULR]

In the orthographic representation of nouns beginning with consonants extracted from the Urhobo Bible and some Christian materials as shown in Example (7a-d), the epithetic /i/ is not represented in all instances where loan nouns begin with consonants, whereas it is represented in the examples extracted from literacy materials such as Ojaide and Aziza (2014) and the Urhobo language scheme of work shown in Example (7e-j). This study supports the latter case in which the epithetic vowel is represented in the orthographic representation. This is because the initial onsetless syllable plays a role in Urhobo. For instance, hiatuses in this language are repaired through vowel elision, and this often affects the first vowel (Aziza, 2010; Ekiugbo and Eme, 2024). The implication of this for reading is that this process is embedded in native speakers' phonological awareness. Thus, during reading, elided vowels, though they are part of the written forms, are not pronounced. Given that

elision often targets V_1 in the language, where a loan noun does not occur in initial position in a concatenation, such as that in Example (8) below, extracted from Revelation 1:1, the non-representation of the epithetic vowel in the written form will pose a challenge for readers.

- (8) o de ji amakashe roye **bru Jɔn** odibo roye ra
3sg then send-pst angel his to John servant his go
'Then He sent His angel to John His servant'

In the concatenation *bru#Jɔn* in the above example, the pronunciation ought to read as: [brizɔni]. This is because the possible derivation for this form from its underlying form is: /bru#izɔni/ > brØ#izɔni > [brizɔni]. This derivation is part of the native speakers' phonological knowledge, which is employed in orthographic reading. Thus, omitting the epithetic vowel may result in the expression being read as [bruzɔni] 'cut John', instead of [brizɔni] '(go) to John'. Thus, it opens up the possibility for ambiguity as well as poses problems for word identification as a result of the inability to phonologically recode such a grapheme to its corresponding sound in the Urhobo language.

4.3.3 Illicit Consonant Cluster

Most of the consonant clusters found in English are not permitted in Urhobo. In Urhobo, only CC is the attested consonant cluster, with C1 as labial or velar, and C2 as liquid. In the literature, loanwords bearing illicit consonant clusters are sometimes represented without the modification applied, as shown in (9a-d) and at other times represented with the necessary modifications, as shown in (9e-f).

- (9) a. Esther Ẹsta [ULB]
b. James Jems “
c. window iwindo [USW]

d. drama	idirama	“
e. school	isikuru	“
f. trace	tresi	[ULR]

Thus, there is also inconsistency in the orthographic representation of consonant clusters that are not attested in the Urhobo language, which may affect the reading fluency and ease for language learners. The process of learning to read is aided by consistency in spelling. When there is regular correspondence between letters and sounds, readers can decode words using their understanding of letter-sound relationships. Readers can apply phonics principles and patterns to a wider range of words, allowing for more efficient reading growth.

4.2.4 Coda

The Urhobo syllable structure is of the open type. Thus, coda is not attested in the language. In such loanwords, the final coda is parsed as the onset of a new syllable with an epithetic vowel attached, as shown in example (10a-b), while non-final codas are deleted, as shown in Example (10c-d). Also, where the non-final coda is a nasal, the preceding vowel is realised as a nasal vowel, as shown in Example (10e-f). Data illustrating how coda adaptation is represented orthographically in the Urhobo literature is shown in Example (11).

(10)	a. ʃam	ʃai.ni	'shine'
	b. pɒt	i.pɔ.tu	'pot'
	c. ɪn.spɛk.tə	i.spɛ.tə	'inspector'
	d. sʌb.dʒɛkt	i.sɔ.ʒɛ.ti	'subject (v)'
	e. tʃæm.pjən	ʃã.pjɔ.ni	'champion'
	f. Prɒmpt 'prompt'	prɔ̃.tu	'prompt'

(11)	a. Jude [dʒu:d]	Jud [iʒudu]	(ULB)
	b. Titus [tʰaitəs]	Titos [itaitusu]	“
	c. James [dʒeɪmz]	Jems [iʒemisi]	“
	d. Ruth [rʊθ]	Rutu [rutu]	“
	e. term [tɜ:m]	itɛmu [itɛmu]	(USW)
	f. college [kɒlɪdʒ]	ikɔleji [ikɔlezi]	“
	g. board [bɔ:d]	ibɔdu [ibɔdu]	“
	h. inspector [ɪn.spek.tə]	ispeto [ispetɔ]	(ULR)
	i. subject (v) [sʌb.dʒekt]	sɔʒeti [sɔʒeti]	“
	j. composition	ikɔmpozishɔni	(USW)
	[,kɒmpə'zɪʃən]	[ikɔpoziʃɔni]	
	k. comprehension	ikɔmprehɛshɔni	“
	[kɒmpri'hɛnʃən]	[ikɔprexɛʃɔni]	
	l. window [wɪndəʊ]	iwindo [iwiɔdo]	“

The data in Example (11) shows that while final codas are not adapted in the orthography of religious materials, with the exception of (11d), all instances of final codas are adapted in literacy materials, as shown in (11e-g). Also, non-final codas are adapted in all instances as shown in (11h-i), except where they involve nasals. As already noted, where a non-final coda is involved, the preceding vowel is realised as a nasal vowel, except where the vowel is in an initial onsetless syllable. In the examples in (11j-k) in which the nasal coda /m/ in the first syllables of the words 'composition' and 'comprehension' are realised as nasal features on their preceding vowels, the grapheme <m> is retained in the Urhobo spellings. In Urhobo, contrastive nasal vowel is indicated using the grapheme <n> immediately after the vowel. The present situation, therefore, includes the use of <n> in some instances and <m> in other instances, which is a form of inconsistency in the orthography. Although it

is possible that the use of <m> in those instances may be informed by a possible effect of labialization, subsequent studies will be required to substantiate this.

4.2.5 Illicit Peak

Only vowels function as the nucleus of the syllable in Urhobo. No syllabic consonants are allowed. Thus, while syllabic consonants such as /l, n/ can function as nuclei elements in English, they do not in Urhobo. In the examples below, we show how illicit peaks of Urhobo-English loanwords are represented orthographically.

- | | | | |
|------|-----------------------|---------------------------|-------|
| (12) | a. Bible [baɪ.bɪ] | ibaibo [i.bai.bo] | [ULB] |
| | b. student [stju.dnt] | istudenti [istud̃ɛti] | [ULR] |
| | d. president | ipresidenti [ipresid̃ɛti] | “ |
| | [pre.zɪ.dnt] | | |

As shown in Example (12) above, illicit peak are adapted in the Urhobo literature. Thus, both in phonological and orthographic terms, all instances of syllabic consonants are adapted using equivalent Urhobo sounds and graphemes respectively.

5. Summary and conclusion

The orthographies of most African languages tend to have a more consistent relationship between spelling and pronunciation. Implicitly, the letters used in words correspond more directly to their sounds, making it easier to predict the pronunciation of a word based on its spelling. Thus, orthographic reforms in such languages, where necessary, are targeted at minimising or eliminating any internal inconsistency. Inconsistency in orthography may have a number of negative consequences, such as difficulties in learning how to read and write in the language. One of the

areas in which this variance is most visible is in the orthographic representation of loanwords. Loanwords often introduce inconsistencies into the writing systems of languages. The present paper investigated the inconsistencies that loanwords introduce into the orthography of Urhobo and their implications for phonological recoding. It is an attempt to provide answers to two fundamental questions in loanword orthography: (i) how are loanwords represented orthographically in Urhobo? and (ii) how should we write loanwords in Urhobo for ease of phonological recoding?

The study identified sounds with grapheme differences, nouns with initial onset, illicit consonants, and coda as the areas of inconsistencies in Urhobo-English loanword orthography. These areas are adapted to both source and target orthographies. The study also identified redundant letters, illicit segments, and illicit peaks as areas of consistency and noted that these areas are adapted to target language orthography. The principal position held in the study is that, in order to minimise the inconsistencies that loanwords bring into target languages, such words should be represented in their adapted forms using the graphemes of the target language. This is based on the assumption that orthographic adaptation should be part of the various intentional choices and cumulative efforts involved in loanword integration in the target language. This is in addition to the fact that the source grapheme may represent a different sound from that in the target language.

References

- Ajiboye, Emuobonuvie M., Philip O. Ekiugbo, Clement Onakpoberuo, T. Irikefe Opuidi, Daniel Sanudje & Frank O. Ejime. In preparation. *Irida Urhobo re ikoleji otete*.
- Al-Athwary, Anwar A. H. 2017. The phonotactic adaptation of English loanwords in Arabic. *Arab World English Journal* 8(3): 392-406.
- Alqarni, Muteb. 2021. Arabic loanwords in seven Ethiopia languages. *Brill's Journal of AfroAsiatic Languages and Linguistics* 13: 423-473.
- Aziza, Rose O. 1994. Vowel harmony in Urhobo. *Nigerian Language Studies* 2: 1-7.
- Aziza, Rose O. 2002. Nasality in Urhobo: An autosegmental perspective. *Journal of West African Languages* 26(2): 11-21.
- Aziza, Rose O. 2008. Neutralization of contrast in the vowel system of Urhobo. *Studies in African Linguistics* 37(1): 1-19.
- Aziza, Rose O. 2010. Phonological processes and their effect on lexical tones in Urhobo. In O.-M. Ndimele (Ed.), *Nigeria languages, literatures, culture and reforms: A festschrift for Ayo Bamgbose*, 461-474. Port Harcourt: M & J Grand Orbit Communications.
- Aziza, Rose O. and Don C. Utulu. 2006. Loanword phonology: English in Urhobo and Yoruba. *Journal of West African Languages* 33(2): 3-22.
- Bassetti, Bene. 2024. Orthographic effects in the phonetics and phonology of second language learners and users. In M. Amengual (ed.), *The Cambridge handbook of bilingual phonetics and phonology*, 699-720). Cambridge: Cambridge University Press.
- Basu, Paul. 2020. Experiments in language. Accessed 20/08/2022 at <https://re-entanglements.net/tag/phonetics/>.
- Bolaji, Ezekiel T. 2016. Constraints interaction in the phonology of Urhobo-English loanwords. In A. Odeunmi, A. Osisanwo, H. Bodunde, & S. Ekpe (Eds.), *Grammar, applied linguistics and society: A festschrift for Wale Osisanwo*, 207-221. Ife: Obafemi Awolowo University Press.
- Clarke, John. 1848. *Specimens of dialects: Short vocabularies of languages and notes of countries & customs in Africa*. Berwick-upon-Tweed: Daniel Cameron.
- Coulmas, Florian. 1989. *The writing systems of the world*. MA: Basil Blackwell.

- Cummings, D. W. 1988. *American English spelling*. Baltimore, MD: The John Hopkins University Press.
- Darah, Moses E. and Philip O. Ekiugbo. 2014. Dissecting the phenomenon of borrowing in Urhobo. *Journal of Linguistics, Language and Culture* 1(1): 17-27.
- Edward, Carney. 1994. *A survey of English spelling*. London: Routledge.
- Ekiugbo, Philip O. & Cecilia A. Eme. 2023. Coda adaptation in Urhobo-English loanwords: A constraint-based account. *Journal of Linguistics and Language Teaching* 14(1): 107-120.
- Ekiugbo, Philip O. & Christian U. C. Ugorji. 2019. A descriptive phonology of the vowel system of Uvwie. *Linguistique et Langues Africaines* 5: 89-107.
- Elugbe, Ben O. 1990. The limits of accuracy in the design of orthographies. *Research in African Languages and Linguistics* 1(1): 11-16.
- Eme, Cecilia A. & Philip O. Ekiugbo. 2024. Hiatus resolution in Urhobo: A constraint-based account. To appear in: *An-Najah University Journal for Research – B* 32(2): 403-418.
- Garvin, Paul L. 1972. General principles for the cultivation of good language. In J. Fishman (ed.), *Advances in language planning*, 417-426. Berlin, Boston: De Gruyter Mouton.
- Hair, P. E. H. 1992. *Barbot's West African vocabularies of c. 1680*. Liverpool: Centre for African Studies, University of Liverpool.
- Hannas, W. 1997. *Asia's orthographic dilemma*. Honolulu: University of Hawai'i Press.
- Hosken, M. 2003. Creating an orthography description. Unpublished manuscript.
- Iweh, Odeh. 1987. Urhobo orthography. In Rebecca N. Agheyisi (ed.), *Orthographies of Nigerian languages*, 57-76. Lagos: National Language Centre, Federal Ministry of Education.
- Ivworin, Godwin A. 2018. The development of written literature in Urhobo (1910–2010). *Abraka Humanities Review* 8(1): 256–269.
- James, Lloyd A. 1929. The practical orthography of African languages. *Africa* 1(1): 125-129.
- Jany, Carmen. (2010). Orthography design for Chuxnabán Mixe. *Language Documentation and Conservation* 4: 231-253.
- Kelly, John. 1969a. Urhobo. In E. Dunstan (ed.), *Twelve Nigerian languages*, 153-161. London: Longmans.

- Kelly, John. 1969b. Vowel patterns in the Urhobo noun. *Journal of West African Languages* 6(1): 21-26.
- Kilham, Hannah. 1828. *Specimens of African languages, spoken in the colony of Sierra Leone*. London: P. White for the Committee of the Society of Friends.
- Knoepke, Jukia, Tobias Richter, Maj-Britt Isberner, Johannes Naumann & Yvonne Neeb. 2014. Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Zeitschrift für Erziehungswissenschaft* 17: 447–471.
- Koelle, S. W. 1854. *Polyglotta Africana*. London: Church Missionary Society.
- Koffi, Ettien N. 1990. *The interface between phonology and morpho(phono)logy in the standardization of Anyi orthography*. Ph.D. dissertation, Indiana University.
- Koffi, Ettien N. 2010. Loanword phonology and orthographic metalanguage. *Paper presented at the orthography seminar at Bibles International, Grand Rapids, MI, July 6-9, 2010*.
- Latham, Robert G. 1862. *Elements of comparative philology*. London: Woodfall and Kinder.
- Lepsius, C. R. 1863. *Standard alphabet*. London: Williams & Norgate.
- Nabofa, Michael Y. 1997. *Evolution of the Urhobo bible and some Christian liturgical books*. Ibadan: End-Time Publishing House.
- Norris, Edward. 1841. *Outline of a vocabulary of a few of the principal languages of Western and Central Africa*. West Strand: John W. Parker.
- Ojaide, Tanure & Rose O. Aziza. Urhobo etymology. *Aridon* 1: 103-112.
- Okolo-Obi, Bosco C. 2015. English loanwords in Urhobo: An adaptation of stress to tone. Paper presented at the 28th Conference of the Linguistic Association of Nigeria, Nnamdi Azikiwe University, Awka.
- Rolle, Nicholas. 2013. Phonetics and phonology of Urhobo. *UC Berkeley Phonology Lab Annual Report* 9: 281-326.
- Society for Promoting Christian Knowledge, 1960. *Ọbe r'ẹrhovbo r'ihwo ejobi* (3rd Ed.). London: Society for Promoting Christian Knowledge.
- Thomas, Northcote W. 1910. *Anthropological report on the Edo-speaking peoples of Nigeria. part II: Linguistics*. London: Harrison and Sons.

- Tian, Feiyang. 2012. Orthographic constraints on the integration of English loanwords in Mandarin Chinese. *Theory and Practice in Language Studies* 2(5): 965-971.
- Ugorji, Christian U. C. 2013. Phonology of English loanwords in Urhobo Corpus. In A. T. Akande & R. Taiwo (eds.), *Contact linguistics in Africa and beyond*, 181-196. New York: Nova Publishers.
- Ukaegbu, Nkechi, Bestman Odeh & Ifeanyi Nwosu. 2022. Phonological outcomes of Yoruba and English contact on Urhobo loan words. *Ghana Journal of Linguistics* 11(2): 20–42.
- Umera-Okeke, Nneka. 2008. Spelling and phonetic inconsistencies in English: A problem for learners of English as a foreign/second language. *African Research Review* 2(1): 64-83.
- Umukoro, Irihayoma N. 1999. *A comparative analysis of pluralisation in Urhobo and English*. MA Thesis, Ahmadu Bello University, Zaria.
- Vendelin, Inga & Sharon Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116(7): 996-1007.
- Welters, William E. 1969. Structural notes on Urhobo. *Journal of West African Languages* 6(2): 85-108.
- Williamson, Kay. 1984. *Practical orthography in Nigeria*. Ibadan: Heinemann Educational Books Ltd.