Gĩkũyũ Orthography and the Social Media

Peter MBURU KURIA

Kenyatta University, Department of Literature, Linguistics and Foreign Languages

https ://doi.org/10.17184/eac.7851

Résumé : Cet article vise à étudier comment l'orthographe du Gĩkũyũ se manifeste dans les médias sociaux et si les médias sociaux ont une influence sur l'orthographe. L'étude s'articule autour des prémisses selon lesquelles les auteurs de Gĩkũyũ sur les médias sociaux ne respectent pas l'orthographe de la langue tout en communiquant leur message. Par conséquent, nous avons étudié les graphèmes utilisés par les rédacteurs et les avons comparés à ceux figurant dans l'inventaire graphémique de Gĩkũyũ. Nous avons analysé les graphèmes à l'aide de la sous-théorie de la correspondance de la théorie de l'optimalité. Cette théorie est phonologique, mais elle a été utilisée dans des études sur l'orthographe. Pour obtenir nos données, nous avons pris des échantillons de fils de conversation en Gĩkũyũ à partir des pages Facebook des stations de médias en Gĩkũyũ, y compris Inooro TV, Kameme TV, et Njatatv Official, ainsi que d'autres pages interactives en Gĩkũyũ, telles que Ithaga Rĩene. À partir des fils de discussion, nous avons recueilli vingt échantillons de chaque graphème Gĩkũyũ. Comme les graphèmes représentent des phonèmes, nous avons vérifié si les significations qu'ils portaient étaient celles voulues par les écrivains et comprises comme telles par les lecteurs, telles qu'observées dans leur interaction. Nous avons supposé que les participants étaient des locuteurs natifs de Gĩkũyũ sur la base du fait que, parmi une pléthore de langues, ils ont choisi de communiquer dans cette langue. Le sexe, l'âge et le niveau d'éducation des interlocuteurs n'étaient pas faciles à déterminer à partir des fils de conversation et, par conséquent, n'étaient pas des variables dans l'étude. Nos conclusions sont que les écrivains utilisent divers symboles pour un graphème donné, mais qu'ils communiquent néanmoins leur message. La théorie de l'optimalité s'est également avérée applicable à l'analyse des graphèmes utilisés.

Mots-clés : Orthographe, graphème, inventaire graphémique, théorie de l'optimalité, médias sociaux

Abstract : This paper sought to investigate how the Gĩkũyũ orthography manifests in the social media and whether social media has influence on the orthography. The study was hinged on the premises that writers of Gĩkũyũ on social media do not observe the orthography of the language yet they communicate their message. Consequently, we investigated the graphemes that the writers used and compared them to what is given in the Gĩkũyũ graphemic inventory. We analyzed the graphemes using the Correspondence sub-theory of the Optimality Theory. This theory is phonological but it has been used in orthography studies. To get our data, we took samples of conversation threads in Gĩkũyũ from the Facebook pages of Gĩkũyũ media stations, including Inooro TV, Kameme TV, and Njatatv Official, as well as other interactive pages in Gĩkũyũ, such as Ithaga Rĩene. From the threads, we collected twenty samples of each Gĩkũyũ grapheme. Since graphemes represent phonemes, we checked to see if the meanings they carried were the ones intended by the writers and understood as such by the readers, as observed in their interaction. We made the assumption that the participants were native speakers of Gĩkũyũ based on the fact that from a plethora of languages, they chose to communicate in the language. The gender, age and level of education of the interlocutors were not easy to ascertain from the conversation threads and, therefore, were not variables in the study. Our findings are that writers use diverse symbols for a given grapheme, but they, nevertheless, communicate their message. The Optimality Theory was also found applicable in the analysis of the graphemes used.

Keywords : Orthography,Grapheme,Graphemic Inventory,Optimality Theory,Social Media

Introduction

The term orthography refers to a system of representing a language in written form. The system uses graphemes, word breaks, punctuation, diacritics, capitalization, and other features (Cahill& Karan, 2008). Additionally, the orthographic system is language specific since languages differ.

Gĩkũyũ is a Bantu language spoken in Kenya by 8,148,668 speakers (Kenya National Bureau of Statistics, 2019). Guthrie (1967) classifies it as E.51 in the Central branch of the Bantu subgroup of the Niger-Congo language family. The language has five major dialects (Mwangi, Njoroge and Mose, 2013) identified as Gĩcũgũ, Mathĩra, Ndia, Northern Gĩkũyũ and Southern Gĩkũyũ.

An orthography for the language was developed by the early Christian missionaries and there was concerted effort in recommending the standard one in 1967 by Armstrong (Armstrong, 1967). However, discrepancies and inconsistencies in the graphemic representation of some phonemes have been observed and studied (Kuria, 2021 ; Kioko, Njoroge and Kuria, 2012) and a standard orthography recommended.

Given the fact that there is increased communication in social media, we sought to find out if these discrepancies and inconsistencies were also a feature in the social media. To do this, we purposefully selected one social media forum, Facebook. The choice was based on the number of the interlocuters on the pages written in Gĩkũyũ, which ensured a richness of data. These included Ithaga Riene iNooro TV (about 183,400 members), Kameme TV (2,000 members), NjataTV (30,488 followers) and Inooro TV (2,700 followers). From Ithaga Riene iNooro TV, we picked ten conversations that had more than twenty interlocutors, and from each of the other stations, we picked five. We followed the conversations to ascertain whether the meanings were the ones intended by the writers, in spite of the graphemes they might have used. We then collected twenty samples of each grapheme from the conversations. The topics on which writers contributed were of a wide range that included politics, sex, security issues, education, and other social issues such as death. They provided adequate data on which we based our study. Gender, age and the educational level of the interlocutors were not variables in our study since there was no way of accurately establishing them.

We examined both consonantal and vocalic phonemes, and their respective graphemes. To analyse the phenomenon of the discrepancies and inconsistencies, we used the Optimality Theory (Smolenski, 1991), specifically the Correspondence Theory.

Gĩkũyũ phonemes and graphemes

Gĩkũyũ has 18 consonantal phonemes (Kuria, 2021). These include two voiceless plosives, four prenasalised voiced plosives, four fricatives, the alveolar tap, four nasals, the voiced affricate, and two approximants. This is shown in Table 1 below.

Each consonantal phoneme is represented by one grapheme. This information is presented in Table 2.

Gĩkũyũ language has seven vocalic phonemes and each phoneme is represented by one grapheme. Figure 1 below presents the phonemes.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

	bilabial	dental	alveolar	palatal	velar	glottal
Stop	ᵐb		t ⁿd		k ᵑg	
Nasal	m		n	ɲ	ŋ	
Tap			ɾ			
Fricative	β	ð		ʃ	ɣ	h
Affricate				ᶮʄ		
Approximant	w			j		

Tab. 1 : Table 1 : Gĩkũyũ Consonant Phoneme Inventory
Adopted from Kuria, 2021

	bilabial	dental	alveolar	palatal	velar	glottal
Stop	mb		t nd		k ng	
Nasal	m		n	ny	ng'	
Tap			r			
Fricative	β	th		c	g	h
Affricate				nj		
Approximant	w			y		

Tab. 2 : Table 2 : Gĩkũyũ Consonant Grapheme Inventory
Adopted from Kuria, 2021

	Front		Back	
[+ATR]	/i/		/u/	High
[+ATR]	/e/		/o/	Mid
[-ATR]		/ɛ/	/ɔ/	Mid
[-ATR]		/a/		[Low]

Fig. 1 : Gĩkũyũ Vowels
Adopted from Kuria, 2021

Each vocalic phoneme is represented with a single grapheme. The figure below presents the graphemes alongside the phonemes they represent.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

[+ATR]	/i/ <i>	/u/ <u>	High
[+ATR]	/e/ <ĩ>	/o/ <ũ>	Mid
[-ATR]	/ɛ/ <e>	/ɔ/ <o>	Mid
[-ATR]	/a/ <a>		[Low]

Fig. 2 : Gĩkũyũ Vocalic Graphemic Inventory
Source : Kuria, 2021

1 Findings

From the data we collected, our observation was that, ten consonantal phonemes and five vocalic graphemes were represented with more than one grapheme. The exceptions were the consonants /m/, /n/, /t/, /k/, /ð/, /h/, /w/ and /j/ and vowels /i/ and /a/. The findings are summarized in Table 3.

Phoneme	Grapheme	Other graphemes			
ᵐb	mb	b			
ⁿd	nd	d			
ᵑg	ng	g	gh		
ᶮj	nj	j			
ɲ	ny	nj			
ŋ	ng'	ng			
β	b	p	v	f	
ʃ	c	sh	s	ch	x
ɣ	g	ng	kh		
ɾ	r	l			
e	Ĩ	i			
ɛ	e	ĩ			
ɔ	o	ũ			
o	ũ	u	o		
u	u	o			

Tab. 3 : Graphemic Representations of Phonemes

We also observed that graphemes meant to represent one phoneme are used to represent different phonemes, some of which are non-existent in the language. This resulted in discrepancies. Table 4 summarizes these findings.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

Grapheme	Respective phoneme	Other phonemes	
mb	/ᵐb/	/β/	
nd	/ⁿd/	*/d/	
ng	/ᵑg/		
nj	/ᶮɟ/	*/j/	
b	/β/	/ᵐb/	*/b/
g	/ɣ/	/ᵑg/	
ĩ	/e/	/i/	
e	/ɛ/	/e/	/i/
o	/ɔ/	/o/	
u	/u/	/o/	
h	/h/	/ᵐb/	

Tab. 4 : Phonemes represented by individual Graphemes. The *(asterisk) shows phonemes not found in the language

2 Discussion

i) Phonemes are represented by more than one grapheme leading to overrepresentation or underrepresentation. Overrepresentation occurs when there are more graphemes than there are phonemes (Kuria, 2021). Such could arise if a phoneme has allophones and each one of them is represented. For example, in the representation of the voiceless palatal fricative/ʃ/ which is realised through the allophones[ʃ], [s] and [tʃ] across the Gĩkũyũ dialects, the graphemes <c>, <s>, <sh> and <ch> are all used in place of the conventional one, <c>. Notably, the last three graphemes are not in the phonetic inventory of Gĩkũyũ. They are the graphemes used in English and Kiswahili for /s/, /ʃ/ and /tʃ/, respectively. Apparently, the writers/ speakers have borrowed these graphemes which is another cause of overrepresentation. Other graphemes that have been used by the writers, yet they do not belong to the graphemic inventory of Gĩkũyũ, include <d>, <kh>, <gh>, <f>, <v>, <p>, <l> and <x>.

ii) Graphemes represent more than one phoneme, including phonemes not found in the phonemic inventory of the language. This phenomenon is referred to as underrepresentation or underdifferentiation. According to Cahill & Karan (2008), underrepresentation occurs when there are fewer graphemes than phonemes. As is apparent in this study, some graphemes are used to represent more than one phoneme. An

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

example is the use of <g> to represent both /γ/ and / ᵑg/. In another example, <e> represents /ɛ/, /e/ and /i/.

iii) We examined both the overrepresentation and underrepresentation against the Optimality Theory (Smolensky and Kiparsky, 1991), specifically the Correspondence Theory, in an endeavour to establish the graphemes that should have been used.

According to Prince and Smolensky (1993), OT has a set of three interrelated modules : GEN, H-EVAL and Lexical Optimization. GEN and H-EVAL together comprise the grammar proper. Lexicon Optimization stores all the lexical forms of a language and provides the input specifications that are input to GEN (Kager, 1999). GEN is the operational component. For any given input (Ink), it generates a set of Candidate Output forms which may vary from the input in several ways. It applies its operations freely, optionally and sometimes, repeatedly. The component is universal within all languages, and it anticipates all options that any language could transform the input into. In the current study, the inputs are the Gĩkũyũ phonemes and the candidate output forms are the graphemes.

H-EVAL takes the Output Candidate forms and evaluates them against a set of constraints (CONS). The constraints are universal and violable. H-EVAL selects the most harmonic, most optimal as the output of that language. The constraints are of types : those that evaluate surface forms with respect to well-formedness (i.e., markedness), and those that deal with faithfulness. Markedness constraints are sensitive to the well-formedness of candidate form, banning disprefered or marked structures or those requiring felicitous or unmarked properties. Faithfulness requires that the input (in our case, the phonemes) should be identical to the output (the correspondent graphemes) of the respective input. Constraint interaction consists in the application of all the constraints, which are put into an order of dominance (the constraint hierarchy), to the inputs. Constraints have conflicting demands on a form. This is recognizable by making constraints violable though they can be violated only if this leads to the non-violation of a higher-ranking constraint. Their ranking varies between languages thus making languages different.

The Correspondence Theory is the standard theory of faithfulness within Optimality Theory (van Oostendorp, 2004). It observes that there is a correspondence relation ($\mathcal{R}$) between elements of each input and output level. Correspondence is a relation between two structures, for example the base and reduplicant, or between the input and output. As McCarthy and Prince (1995, p. 14) note, "Given two strings S1 and S2, correspondence is a relation $\mathcal{R}$ from the elements of S1 to those of S2. Elements $\alpha \in$S1 and $\beta \in$S2 are referred to as correspondents of one another when $\alpha \mathcal{R} \beta$." (p 23)

Correspondence is presented as a relation rather than a function, to allow for one-to-many relations.

The reduplicants or outputs are evaluated together with the base or input. Each candidate pair (S1, S2) comes from Gen equipped with a correspondence relation between S1 and S2 that expresses the relation between them. There is a correspondence relation for each (I, O) candidate-pair. EVAL then considers each candidate pair with its

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

associated correspondence relations, assessing the completeness of correspondence in S1 or S2, the featural identity of correspondent elements in S1 and S2, and so on.

Constraints assess correspondence and identity of correspondent elements. For every correspondence relation, we have separate constraints. The separateness makes the constraints rankable.

We took the both markedness and faithfulness constraints as identified from the criteria applied in setting up the orthography of a language (Kuria, 2021) and used them to analyse our data. The constraints included the following :

Faithfulness Constraints

1) Constraints that mitigate against deletion of a segment from the input

   MAX – (1PHONEME, 1GRAPHEME) = MAX-(1P, 1G)

2) Constraints which require that the relationship between a grapheme and a phoneme should be regular, thus mitigating against inconsistency and unpredictability

   MAX-(1GRAPHEME, 1 PHONEME) = MAX – (1G, 1P)

3) Constraints which require that all segments in the output must be present in the input.

   DEP – (PHONEME, GRAPHEME) = DEP – (P, G)

4) Constraints which require that allophones be represented.

   IDENT- (FP)

Markedness Constraints

5) Constraints which demand that all voiced plosives and affricates must be preceded by a nasal.

   NC- (NASAL, PLOSIVE) = NC – (N, P)

6) Constraints with the requirement that salient features beyond the segment must be represented.

   DIACRITICS – (FEATURES) = DIA-(FT)

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

7) Constraints that ban the writing of graphemes that do not convey meaning, or convey a different meaning

   CM – (CONVEY MEANING) = CM – (MEAN)

8) Constraints that demand no change in the graphemic inventory

   INVENTORY-(GRAPHEME) = INV-(G)

9) Constraints which require that familiar symbols should not be used unfamiliarly

   IDENT-(FAMILIAR, UNFAMILIAR) = IDENT- (F, U)

10) Constraints that advocate for simplicity

    *COMP

11) Constraints which advocate for writing in continuous flow

    DIACRITICS = DIA

12) Constraints which require that salient features be represented

    VIS-(F)

13) Constraints that demand the visibility of vowel length

    VIS – (V-L)

For the aforementioned constraints, we establish a hierarchy which leads to the choice of the correct orthographic output for Gĩkũyũ, in line with arguments by other researchers (Wiese & Song, 2010 ; see also McCarthy, 2007, p. 8). We analyzed each of the 10 consonantal and the 5 vocalic phonemes, together with their graphemes, against the relevant constraints using OT Tableaus as presented below :

1) The prenasalised voiced bilabial plosive /^{m}b/

The relevant faithfulness constraints for this phoneme are ID- (1G, 1P) and MAX- (1P, 1G), while the markedness constraints are *COMP and NC- (N, P). We established the ranking :

NC- (N, P) » *COMP » MAX- (1P,1G) » ID- (1G, 1P)

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

/ᵐb/<mb>	NC- (N, P)	*COMP	MAX- (1P,1G)	IDENT- (1G,1P)
a. ☞mb		*		
b. b	*!		*	*

The analysis is as presented in the tableaux below.

Grapheme (b) violates MAX- (1P, 1G) because it is not the specified grapheme for the phoneme, and violates ID- (1G, 1P) because it represents another phoneme in Gĩkũyũ. These are lowly ranked constraints in the language. The two relevant markedness constraints here are NC- (N, P) and COMP, where the first one ranks higher in this case. Grapheme <mb> violates COMP because it is complex, but does not violate the higher ranked constraint. Grapheme <b> fatally violates NC-(N, P) because it does not represent the pre-nasal segment in the phoneme. Thus, it loses as a representation of the phoneme which leaves Candidate a) as the optimal one.

2) The prenasalised voiced alveolar plosive /ⁿd/

The relevant faithfulness constraints are MAX- (1P, 1G) while the markedness ones are *COMP, NC – (N, P) and INV – (G). They were ranked as :

INV- (G) » NC – (N, P) » *COMP » MAX- (1P, 1G)

The analysis is as shown below.

/ⁿd/ <nd>	INV-(G)	NC-(N, P)	*COMP	MAX- (1P, 1G)
a). ☞nd			*	
b). d	!*	*		*

The faithfulness constraint in this case is MAX- (1P, 1G), a low ranking constraint which Candidate (b) violates because the input does not match the output for /ⁿd/. There are three markedness constraints, INV- (G), NC- (N, P) and COMP, ranked in that order. Candidate (a) violates COMP because it is complex, but it does not violate the higher ranked constraints. Grapheme <b> violates NC – (N, P) since it leaves out representation of the pre-nasal segment. It fatally violates INV-(G) since it is not in the graphemic inventory of Gĩkũyũ.

3) The prenasalised voiced velar plosive /ᵑg/

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

The faithfulness constraints that were applicable here were MAX- (1P, 1G), IDENT- (1G, 1P) and IDENT- (F P) while the markedness ones were NC- (N, P) and *COMP. They were ranked as :

NC- (N, P) » *COMP » MAX- (1P, 1G) » IDENT- (1G-1P) » IDENT – (FP).

The analysis is presented below.

/ᵑg/<ng>	NC- (N, P)	*COMP	MAX- (1P, 1G)	IDENT- (1G, 1P)	IDENT- (FP)
a). ☞ng		*			
b). g	!*		*	*	*

There are five constraints in consideration here, with the three faithfulness ones ranking lower. Candidate (b) is a grapheme that represents another phoneme and, so, violates both IDENT-(1G, 1P), and of IDENT- (FP). It is also not the representation of the given input which violates MAX-(1P, 1G). Though the grapheme is easy to write and, therefore, does not violate *COMP, it does not represent the prenasal segment and so violates NC- (N, P). This is a fatal violation.

Candidate (a) is the most optimal candidate. It does not violate any of the faithfulness constraints. However, it violates *COMP since it is made up of two segments. Nevertheless, it does not violate the other higher ranked constraints in that it conveys the correct meaning and also represents the prenasal segment that is present in the phoneme.

4) The prenasalised voiced affricate

The faithfulness constraints here are MAX – (1P, 1G) and IDENT- (F, P). The markedness ones are INV- (G), NC- (N, P) and *COMP. They are ranked as :

INV- (G) » NC- (N, P) » *COMP » MAX – (1P, 1G) » IDENT- (F, P).

The analysis is given below.

The ranking for the constraints is much like for the prenasalised phonemes discussed above. Candidate (b) violates IDENT-(FP) and MAX- (1P, 1G) since it represents another phoneme. Though it does not violate *COMP in that it is a simple grapheme, it violates NC- (N, P) because it does not represent the prenasal segment of the affricate and fatally violates INV – (G) because it is not in the graphemic inventory of Gĩkũyũ. Candidate (a) wins by observing the higher ranked constraints.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

/ᵐɟ/<nj>	INV- (G)	NC- (N, P)	*COMP	MAX- (1P, 1G)	IDENT- (FP)
a). ☞nj			*		
b). j	!*	*		*	*

5) The palatal nasal /ɲ/

The relevant faithfulness constraints are MAX – (1G, 1P), IDENT- (1G, 1P) and DEP- (P, G) while the markedness ones are CM- (MEAN). They were ranked as follows :

CM- (MEAN) » MAX – (1G, 1P) » IDENT- (1G, 1P) » DEP- (P, G)

The analysis of the palatal nasal is presented below.

/ɲ/ <ny>	CM- (MEAN)	MAX- (1G, 1P)	IDENT- (1G, 1P)	DEP (P,G)
a). ☞ny				
b). nj	!*	*	*	*

Candidate (b) violates all the lower ranked faithfulness constraints. It has no correspondence with the input which violates DEP-(P, G), and it introduces the nasal component in its representation, thereby violating IDENT-(1G, 1P). In addition, it is not the representation of /ɲ/, which makes it violate MAX-(1G, 1P). Its violation of CM-(MEAN), a higher ranked constraint is, however, fatal.

6) The velar nasal /ŋ/

There were three relevant faithfulness and two markedness ones. These were faithfulness MAX- (1P, 1G), MAX-(1G, 1P) and DEP- (P, G), and markedness VIS-(F) and *DIA.

They were ranked as :

VIS-(F) »*DIA » MAX- (1P, 1G) » MAX-(1G, 1P) » DEP- (P, G)

The analysis is shown in the tableau below.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

<ng'>	VIS-(F)	*DIA	MAX-(1P, 1G)	MAX-(1G, 1P)	DEP-(P, G)
a). ☞ng'		*			
b). ng	!*		*	*	*

Candidate (b) violates all the lower ranked faithfulness constraints. Since there is no correspondence between it and the input which is a velar nasal, it flouts DEP – (P, G). It flouts MAX-(1G, 1P) because it is the grapheme for another phoneme. It is also not the grapheme for the input which violates MAX-(1P, 1G). However, the three constraints are not decisive in the identification of the more harmonic candidate. The candidate observes the higher ranked *DIA-(F) because, between the two candidates, it is the easier one to write ; the other one has a diacritic. However, it fatally violates VIS-(F) which requires that salient features be represented. In this case, nasality is the salient feature.

Candidate (a) is the winner. It does not violate the low ranked constraints but violates *DIA-(F). However, this is not decisive because, though diacritics make writing difficult, they may, nevertheless, carry prominent features. This is catered for by VIS-(F) constraints which are ranked higher. Since the candidate does not violate these constraints, it emerges as the more harmonious of the two.

7) The voiced bilabial fricative /β/

Two markedness constraints, INV- (G) and *COMP, and four faithfulness constraints, MAX-(1P, 1G), MAX-(1G, 1P), DEP- (P, G) and IDENT-(F, U), were identified for this phoneme. Their ranking is given below.

INV- (G) » *COMP » MAX-(1P, 1G) » MAX-(1G, 1P) » DEP- (P, G) » IDENT-(F, U)

The analysis is presented in the following tableau.

/β/ <b>	INV-(G)	*COMP	MAX-(1P, 1G)	MAX-(1G, 1P)	DEP-(P, G)	IDENT-(F, U)
a). ☞b						
b). mb		!*	*	*		*
c). p	!*		*			*
d). v	!*				*	
e). f	!*		*		*	

Candidate (e) violates DEP- (P, G) by inserting the feature [+ dental] which is not present in the input. It violates MAX-(1P, 1G) by deleting the feature [+voice] that

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

is present in the input. It loses by fatally violating INV-(G) because it is not in the inventory of the language. Candidate (d) also inserts the feature [+dental] thereby violating DEP- (P, G) and is eliminated for violating INV-(G). Candidate (c) violates IDENT-(F, U) because no respondent produced the voiceless bilabial plosive /p/. The deletion of the [+voice] feature makes it violate MAX-(1P, 1G), but it is the violation of INV-(G) which proves fatal to it. Candidate (b) violates the lowest ranked constraint. In place of the given input, respondents do not produce the phone that the candidate represents. The candidate also violates *COMP which knocks it out of contention. Candidate (a) proves the most harmonic, and so the winner by not violating any of the constraints.

8) The voiceless palatal fricative /ʃ/

To analyse the phoneme, two faithfulness constraints, IDENT- (F, U) and IDENT (F,T), as well as markedness constraints INV- (G) and *COMP, were identified. They were ranked as follows :

INV- (G) » *COMP » IDENT- (F, U) » IDENT (F,T)

We present the analysis in the tableau below.

/ʃ/ <c>	INV-(G)	*COMP	IDENT-(F, U)	IDENT-(F,T)
a). ☞c			*	*
b). s	!*			
c). ch	!*	*		

Candidate (c) observes IDENT-(FT) because some speakers produce [tʃ]. It also observes IDENT-(F, U) because some speakers use a familiar symbol though in an unfamiliar way. In addition, it is complex and so violates *COMP. However, it is the violation of INV-(G) that renders it disharmonic and so it loses out.

Candidate (b) does not win either. Though it violates none of the faithfulness constraints as well as *COMP, it nevertheless, violates INV-(G), the highest ranked and decisive constraint because it is not in the inventory of Gĩkũyũ. Candidate (a) wins though it violates the lower ranked faithfulness constraints.

9) The voiced velar fricative /ɣ/

The faithfulness constraints relevant for the phoneme were MAX- (1P, 1G), IDENT- (1G, 1P) and DEP- (P, G). There was one relevant markedness constraint, CM- (MEAN). The ranking was :

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

CM- (MEAN) » MAX- (1P, 1G) » IDENT-(1G, 1P) » DEP- (P, G)

The tableau below shows the analysis.

/ɣ/ <g>	CM- (MEAN)	MAX- (1P, 1G)	IDENT- (1G, 1P)	DEP- (P, G)
a). ☞g				
b). k	!*	!*	*	*

There is only one markedness constraint in this case. Candidate (b) violates DEP-(P, G) by inserting a new feature, [+plosive]. It represents the grapheme for another phoneme and so violates IDENT-(1G, 1P). The candidate also deletes the features [+fricative, +voice] from the input and so violates MAX-(1P, 1G). The violation of CM-(MEAN) knocks it out of contention against the other candidate. Candidate (a) is optimal. It wins because it violates none of the constraints under consideration.

10) The Tap /ɾ /

The relevant faithfulness constraints for the phoneme are MAX- (1P, 1G) and ACCEPT. There is one relevant markedness constraint, INV- (G). They are ranked as follows :

INV- (G) » MAX- (1P, 1G) » ACCEPT

The tableau below shows the analysis.

/ɾ/ <r>	INV - (G)	MAX- (1P, 1G)	ACCEPT
a). ☞r			
b). l	!*	*	*

Candidate (b) is not accepted by some respondents and so violates the low ranked ACCEPT constraint (Githiora, 2003). Since it interferes with the One Phoneme-One grapheme accuracy, it also violates MAX-(1P, 1G). However, its violation of INV-(G) is fatal and eliminates it. Candidate (a) wins. It does not violate any of the constraints and so is the optimal candidate.

11) The Mid [+ATR] Vowel /e/

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

Faithfulness constraints ACCEPT, MAX-(1P, 1G), MAX-(1G, 1P) and IDENT-(FP) were found relevant. The relevant markedness constraints were CM- (MEAN), DIA and COMP. We ranked them as follows :

CM- (MEAN) » *DIA* » COMP » ACCEPT » MAX-(1P, 1G) » MAX-(1G, 1P) » IDENT-(FP)

The analysis is presented in the tableau below.

/e/<î>	CM- (MEAN)	*DIA	*COMP	ACCEPT	MAX- (1P, 1G)	MAX- (1G, 1P)	IDENT- (FP)
a).☞î		*	*	*			
b). i	!*				*	*	*

Candidate (b) violates most of the faithfulness constraints except ACCEPT. This is because this is the grapheme that most speakers are familiar with. It is easy to write and read and also has no diacritic. Thus, the candidate does not violate both COMP and DIA. However, this candidate carries an unintended meaning and so violates CM-(MEAN), the highest ranked constraint. This makes it a loser. Candidate (a) is the winner. Though it is rejected by some candidates and so violates ACCEPT, it is also not easy to write, which violates COMP. In addition, it has a diacritic which makes it violate DIA. Nevertheless, it carries the intended meaning and so does not violate CM-(MEAN). Naturally then, it is the optimal candidate.

12) The Front [-ATR] Vowel /ɛ/

MAX-(1P, 1G), IDENT-(1G,1P) and IDENT (FP) were the relevant faithfulness constraints. The markedness constraints were CM-(MEAN). They were ranked as follows :

CM-(MEAN) » MAX-(1P, 1G) » IDENT-(1G,1P) » IDENT (FP)

Analysis done for the phoneme is presented in the tableau below.

/ɛ/ <ɛ>	CM- (MEAN)	MAX- (1P, 1G)	IDENT- (1G,1P)	IDENT (FP)
a). ☞e				*
b). a	!*	*	*	*

Candidate (b) is the loser. Though it does not violate IDENT-(FP) since the phone produced by speakers in this word position is a long one and length is ordinarily

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

represented with double graphemes, it is not the graphemic representation of /ɛ :/ and so violates MAX-(1P, 1G). Its violation of VIS-(V-L), which mitigates against representation of length in the word initial position, proves fatal and eliminates it. Candidate (a) is the winner. It violates IDENT-(FP) because it does not represent the phone for this particular phoneme. It also represents another phoneme, /e/ and so violates IDENT-(1G, 1P). These constraints are both ranked low. The candidate wins because it does not violate the high ranked markedness constraint, VIS-(V-L).

13) The Back High [+ATR] Vowel /u/

There faithfulness constraints for the analysis of the phoneme were MAX- (1P, 1G), IDENT-(1G, 1P), IDENT-(1G,1P) and IDENT-(FP). The markedness constraints were CM- (MEAN). Their ranking was :

CM- (MEAN) » MAX- (1P, 1G) » IDENT-(1G, 1P) » IDENT-(1G,1P) » IDENT-(FP)

The analysis is presented in the following tableau.

/u/ <u>	CM- (MEAN)	MAX- (1P, 1G)	IDENT- (1G, 1P)	IDENT- (1G,1P)	IDENT-(FP)
a). ☞u					
b). ũ	!*	*	*	*	*

Candidate (e) violates all the faithfulness constraints. The combination of two symbols renders it complex and so violates COMP. It conveys an unintended meaning which makes it fatally violate CM-(MEAN). Candidate (d) fares much like Candidate (e) though it does not violate COMP. It is eliminated when it fatality violates CM-(MEAN). Candidate (c) violates all the low ranking faithfulness constraints as well as COMP. Its violation of CM-(MEAN) is, however, fatal. Candidate (b) violates MAX-(1P, 1G) because it is not the grapheme for the phoneme in the word initial position. However, this is a low ranked constraint in the language. It is knocked out of contention when it violates the VIS-(V-L) constraint which mitigates against representing vowel length in the word initial position. Candidate (a) is the most harmonic and, therefore, the winner. It does not violate any of the constraints.

14) The Mid [+ATR] Vowel /o/

The identified faithfulness constraints are FAM, MAX-(1P, 1G), IDENT-(1G,1P) and IDENT-(FP). The markedness ones are CM- (MEAN) and *DIA. They were ranked as follows :

CM- (MEAN) » *DIA » FAM » MAX-(1P, 1G) » IDENT-(1G,1P) » IDENT-(FP)

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

The analysis is presented in the tableau below.

/o/ <ũ>	CM-(MEAN)	*DIA	FAM	MAX-(1P, 1G)	IDENT-(1G,1P)	IDENT-(FP)
a). ☞ũ		*	*			
b). u	!*			*	*	*
c). o	!*			*	*	*

Candidate (c) and (b) violate IDENT-(FT) because they represent phones that are not articulated for the phoneme. Candidate (c) deletes the property [+ATR] from the phoneme and inserts [-ATR] which violates the faithfulness constraints. Candidate (b) deletes the property [MID] and inserts [HIGH] which also violates the faithfulness constraints. Both candidates also violate CM-(MEAN) because they convey different meanings from what phoneme /o/ conveys. This is a fatal violation for the two. Candidate (a) is the winner. Though it violates FAM because it is a familiar symbol used in an unfamiliar way and also violates *DIA because of the diacritic, it nevertheless, wins because it conveys the correct meaning carried by phoneme /o/.

15) The Back [-ATR] Vowel /ɔ/

The relevant faithfulness constraints were MAX-(1P, 1G), MAX-(1G,1P), DEP-(P, G), and IDENT-(FP). The markedness constraints were CM- (MEAN). Their ranking was :

CM- (MEAN) » MAX-(1P, 1G) » MAX-(1G,1P) »DEP-(P, G) » IDENT-(FP).

The analysis is as presented in the tableau below.

/ɔ/ <o>	CM-(MEAN)	MAX-(1P, 1G)	MAX-(1G,1P)	DEP-(P, G)	IDENT-(FP)
a). ☞o					
b). u	!*	*	*	*	*

Candidate (b) violates IDENT-(FT) because it does not represent the phone for the phoneme. It also violates DEP-(P, G) by inserting the feature [+ATR] and violates MAX-(1P, 1G) by deleting the feature [-ATR]. Fatally, it violates CM-(MEAN) because it conveys an unintended meaning. Candidate (a) is the most harmonic since it violates none of the constraints.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

Conclusion

Both overrepresentation and underrepresentation cause discrepancies and inconsistencies. Consequently, since there is a mismatch between the phoneme and the respective grapheme, or between the grapheme and the respective phoneme, we expect a breakdown in communication between the writers. For example, the minimal pair /ɣea/ "gĩa" (acquire) and / ᵑgea/ "ngĩa" (poor person) would mean the same, and so would /ᵐbaᵐba/ "mbamba" (cotton wool) and /βaβa/ "baba" (father). Other words would have no meaning to readers. Such include "khai', and "ghai" as used in reference to "Ngai" (God), or even "sai" as used in reference to "cai" (tea). Interestingly, from the threads of conversations studied, this has not been the case ; the intended meanings were conveyed.

Given the above, we hypothesise that both the writers and readers cognitively have the correct orthographical forms and are aware of the orthographical constraints in Gĩkũyũ. However, since languages are lazy, the interlocutors prefer forms that are easier to use, forms that use minimum effort for maximum communication and that are available either in the graphemic inventory of the language or languages that Gĩkũyũ is in contact with. We also hypothesise that other features of conversation, more so the context, are utilised in the conversations, over and above the orthography.

As a recommendation, we propose that vibrant efforts towards the standardization of the Gĩkũyũ orthography should be done. The process and the outcome of such efforts should be made highly visible in social media platforms. Additionally, there should be concerted efforts at teaching of Gĩkũyũ orthography in schools since an orthography is learnt rather than acquired.

References

Armstrong, L.E. *The Phonetic and Tonal Structure of Kikuyu.* London : Dawsons of Pall Mall, 1967.

Banda, F. *Language across Border.* Ed., Cape Town : CASAS, 2002.

Baroni, A. 2016, *Constraint Interaction and Writing Systems Typology.* Dossiers d'hel, SHESL, 2016, Ecriture(s) et representations du langage et des langues, 9, p. 290-303. linguist.univparisdiderot.fr/hel/-dossiers/numero9.hal013053

Boersma, P. 2014. *Functional phonology : Formalizing the Interaction between Articulatory and Perceptual drives. Downloaded from UvA-DARE, the Institutional Repository of the University of Amsterdam (UvA) dare.uva.nl/cgi/arno/show.cgi*

Bradley, D. *Issues in Orthography Development and Reform.* La Trobe University, Australia, 2003.

Bradley, D. Language attitudes : The key factor in language maintenance. *Language endangerment and language maintenance*, 2002.

Cahill, M. and Karan, E. *Factors in designing effective orthographies for unwritten languages.* SIL Electronic Working Paper 2008-001. www.sil.org/silewp/2008/silewp2008-001.pdf, 2008.

Clements, G.N. *Vowel Height Assimilation in Bantu Languages. Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society : Special Session on African Language Structures, p. 25-64,* 1991.

Coulmas, F. *The Writing Systems of the World.* Blackwell : Oxford, 1989.

Flemming, E. *Auditory Representations in Phonology.* New York : Routledge, 2002.

Frost, R. Orthography and Phonology : The Psychological Reality of Orthographic Depth. *Haskins Laboratories Status Report on Speech Research 1989,* S-R-99/100, 162-171, 1989.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

Githiora, C. Gĩkũyũ orthography : Past and Future Horizons. Githiora, C., Littlefield, H. & Manfredi, V. (Eds). *Kinyira njira ! Step Firmly on the Pathway ! Trends in African Linguistics 5* (85-95, Lawrenceville, New Jersey_ :_ Africa World Press, 2003.

Githiora, C. *Gĩkũyũ*. London : Elsevier, 2016.

Guthrie, E. *Classification of Bantu Languages*. London : Dawsons of Pall Mall, 1967.

Hasselbring, S. A. *Cross Dialectal Acceptance of Written Standards*. Unpublished PhD Thesis, University of South Africa, 2006.

Iribemwangi, P. I. A Case for the Harmonization of Kikuyu, Kiembu and Kimbeere Phonology and Orthography. Ogechi, N. A., Oduor, J. A. N.& Iribemwangi (éds.) *The Harmonization and Standardization of Kenyan Languages* : *Orthography and Other Aspects,* p. 20-38. Cape Town : CASAS, 2012.

Jeon, L., Li, J., Mauney, S., Navarro, A. & Wittke, J. A Basic Sketch Grammar of Gĩkũyũ. Englebretson, R. (ed). *A Special Issue of Rice Working Papers in Linguistics Volume VI,* 2015.

Kager, R. *Optimality Theory*. New York : Cambridge University Press, 1999.

Karan, E. *Writing System Development and Reform : A Process*. Unpublished MA Thesis, University of North Dakota, 2006.

Kerr, H. *The Cognitive Psychology of Literacy Teaching : Reading, Writing, Spelling, Dyslexia (& a bit besides)*. www.hugokerr.info, 1998.

Kioko, A. N. A Case for Improved Orthography. In Owino, F. R. (éd.) *Speaking African : African Languages for Education and Development* (p. 231-241). Cape Town : CASAS, 2002.

Kioko, A. N., Njoroge, M. C., Kuria, P. M. (2012). Harmonizing the Orthography of Gĩkũyũ and Kĩkamba. In Ogechi, N. A., Oduor, J. A. N. & Iribemwangi (éds.) *The Harmonization of and Standardization of Kenyan Languages* : *Orthography and Other Aspects* (p. 39-63). Cape Town : CASAS.

Kioko, A. N., Mbori, B., Kitetu, C., Ongarora, D., Buliba, F. A., Mworia, F., Chai, F., Basuba, H., Mwaliwa, H., Chacha, L., Saidi, M. V., Njoroge, M., Lonyangapuo, M., Njoka, M., Ndung'u, R. & Ogechi, N. (2012). *A United Orthography for Bantu Languages of Kenya*. Cape Town : CASAS.

Kuria, P.M. *An Examination of the Discrepancies between the Graphemic and the Phonological Structure of Gĩkũyũ*. Unpublished M.A. Thesis, Kenyatta University, 2005.

Kuria, P. M. *A Graphonological Study of Gĩkũyũ : An Optimality Approach*. Unpublished PhD Thesis, Kenyatta University, 2021.

Marinelli, C.V., Zoccolotti, P. and Romani, C. The ability to learn new written words is modulated by language orthographic consistency. In *PLoS ONE* 15(2) : e0228129. doi.org/10.1371/journal.pone.0228129 2020.

Massamba, D. P. M. *Phonological Theory : History and Development*. Dar-es-Salaam : Dar-es Salaam University Press, 1996.

McCarthy, J. and Prince A. "Faithfulness and Reduplicative Identity." *University of Massachusetts Occasional Papers in Linguistics 18 : Papers in Optimality Theory*. GLSA, University of Massachusetts, Amherst. [ROA-60.], 1995.

McCarthy, J. J. *What is Optimality Theory ?* University of Massachusetts, Amherst, 2007.

Mutahi, E. K. *Sound Change and the Classification of the Dialects of Southern Mt. Kenya*. Unpublished PhD Thesis, University of Nairobi, 1977.

Mwihaki, A. N. *Loanword Nativization : A Generative View of the Phonological Adaptation of Gĩkũyũ Loanwords*. Unpublished PhD Thesis, Kenyatta University, 1998.

Peng, L. Kikuyu Vowel Harmony. *South African Journal of African Languages, 20(4),* 2000_._

Perry, C., Ziegler, J. C., & Colheart, M. How Predictable is Spelling ? Developing and Testing Metrics of Phoneme – Grapheme Contingency. *The Quarterly Journal of Experimental Psychology,* 2002, 55A(3) (p. 897-915), 2002.

Prince, A.S. and Smolensky, P. *Optimality Theory : Constraint Interaction in Generative Grammar*. Ms., Rutgers University, Rutgers, N.J. and University of Colorado at Boulder. Boulder, Col., 1993.

Republic of Kenya. *Basic Education Curriculum Framework*. Nairobi : Kenya Institute of Curriculum Development, 2017.

Ruddell, S. *An acoustic study of the Japanese voiceless bilabial fricative*. San Francisco State University, 2013.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406

Sampson, G. *Writing Systems*. London : Hutchison, 1985.

Schroeder, L. Bantu Orthography Manual. *SIL International* : SIL e-books, 2008.

Smalley, W. A. How shall I Write this Language ? In *Orthography Studies : New Articles on New Writing System.* London : Oxford University Press, 1964.

Smith, E.J. M. Using Optimality Theory to Reconstruct Elamite Phonology. *LIN 1290 Forum Paper,* 2004.

Starwalt, C. G. A. *The Acoustic Correlates of ATR Harmony in Seven- And Nine Vowel African Languages : A Phonetic Inquiry into Phonological Structure*. Unpublished PhD Thesis, University of Texas at Arlington, 2008.

van Oostendorp, Marc. *The Theory of Faithfulness*. Amsterdam : Meertens Institute, 2004.

Venezky, R. L. In Search of the Perfect Orthography. *Written Language and Literacy 7 :2* (139-163). John Benjamins Publishing Co., 2004.

Wanjau, G. *Mwandikire wa GiGĩkũyũ Karin'ga : Ibuku ria Keeri*. Karatina : Gakaara Press Ltd., 1991.

Published in : WERE, Vincent Otaba ; ZOUOGBO, Jean-Philippe (dir.), *Développement durable : Amplifier les langues. Valoriser les cultures. Impliquer les populations. Sustainable development : Promoting languages. Leveraging cultures. Working with populations*, Editions des archives contemporaines, ISBN : 9782813004406, doi : 10.17184/eac.9782813004406