Constructions 15.2 (2023)
doi 10.24338/cons-549

A construction morphology approach to Yoruba numerals

Matthew Ajibade[1]
[1] Indiana University Bloomington

Abstract

This paper provides a form-meaning account of Yoruba numerals, demonstrating how they can be analyzed with morphological schemas, as developed in (Booij 2010). The discussions on Yoruba numerals, described as one of the most complicated numeral systems of the world (Hurford 1975), have centered on the array of arithmetic operations involved in their computation. This complex computation results in phonological variations for some numbers in isolation and when such numbers are stems for the derivation of complex numerals, thereby constituting a challenge to the identification of numerical meanings from phonological forms. In a more recent approach to the analysis of complex words, Geert Booij, through a series of works (Booij 2005, 2007, 2009, 2010, 2018) developed Construction Morphology (CxM) to provide a systematic pair of linguistic forms and meanings of complex words. This approach uses constructional schema to account for the holistic properties of complex words and how existing templates can be used to create new words (Booij 2010, 2018; Jackendoff 2022). Following the CxM approach, this paper, through a systematic examination of each group of numerals with a similar arithmetic operation, shows that numbers with similar mathematical operations have feature commonalities that set them apart from other groups of numerals. These commonalities show how Yoruba users make form-meaning representations of numbers that are represented with different allomorphs, whether phonologically governed or idiosyncratic. The paper concludes that presenting holistic properties of Yoruba numerals through constructional schemas shows a direct way to identify their mathematical properties from phonological forms.

Keywords: Construction Morphology, Schema, Numerals, Allomorphy, Yoruba

1 Introduction

Numerals are intrinsically complex linguistic expressions whose configurations follow different patterns across languages. In Yoruba, a language with "probably the most unusual and complicated of any of the world's natural language numerical systems" (Hurford 1975:211), the basic numbers are 1-10, and isolated numbers 20, 30, 200, 300, and 400. Other numerals from 11 are complex linguistic expressions whose compositions consist of an array of numerical operations. As exemplified in Table 1, some subsets of Yoruba numerals within this

Corresponding author(s): Matthew Ajibade, Indiana University Bloomington, maajibad@iu.edu

complex system are a product of simple subtraction, addition, or multiplication as in (f), (g), and (h), respectively, while the derivation of other subgroups as exemplified with 563 in (e) and 500 in (i) involves a concatenation of numerical operations[1]. In addition to these eclectic arithmetic operations, the complexity of Yoruba numerals is attributed to the numerous subgroups to which each numeral belongs and allomorphy within the system. Examples of allomorphy include the appearance of 10 as *ẹ̀wá* in isolation in (b) but as *àá* in 13 in (g), 200 as *igba* in (d) but as *ẹgb-* in (h & i), and 100 as *ọgọ́rùn-ún* in (c) but as *ẹ̀ẹ́* in (i) [2].

	Numerical Form	Computation		Numerical Form	Computation
a.	ẹ̀ta '3'	3	f.	ẹ̀ta-dín-l-ógún 3-minus-PREP-20 '17'	20 – 3
b.	ẹ̀wá '10'	10	g.	ẹ̀tà-l-àá 3-plus-10 '13'	3 + 10
c.	ọg-ọ́rùnún twenty-five '100'	20 x 5	h.	ẹgb-ẹ̀ta two hundred-three '600'	200 x 3
d.	igba '200'	200	i.	ẹ̀ẹ́-d-ẹ́gb-ẹ̀ta 100-minus-200-3 '500'	200 x 3 – 100
e.	ọ̀tà-lé-l-ẹ́ẹ̀-d-ẹ́gb-ẹ̀ta ó lé m-ẹ́ta 60-plus-PREP-100-minus-200-3 RP plus ADJ-3 '563'				((200 x 3 – 100) + (60) + (3))

Table 1: A Representative Sample of Yoruba Numerals

Much of the previous work has focused on providing an input-output account of Yoruba numerical forms. In this traditional word formation rule approach, which presents the (morpho)phonological processes involved in the derivation of each numeral, two proposals have been made in accounting for the derivation of 100 (and other multiples of 20 in Table 4), for example. The first (Awobuluyi 1967, 1992, 2008) argues that *ọgọ́rùnun* derives through compounding *ogún+ẹrún*. Compounding leads to vowel coalescence [ũ] + [ɛ]= [ɔ] as a result of a constraint that disallows vowel hiatus in Yoruba, creating an environment for cross-consonant vowel harmony. An alternate approach (Ajolore 1972; Bamgbose 1986, 1990; Owolabi 2011) argues that hiatus is resolved through vowel deletion *og+ẹ̀rùnún*, vowel harmony *ọg+ẹ́rùnún*, and cross-consonant vowel assimilation *ọgọ́rùnún*. Similarly, this account considers *ẹgbẹ̀ta* '600' to derive through compounding *igba-ẹ̀ta* '200-3', which creates an environment for vowel deletion *igbẹ̀ta* and cross-consonant vowel harmony *ẹgbẹ̀ta*.

Although this approach offers an insight into each constituent of Yoruba complex numbers, two main issues are worth consideration: (i) it does not provide a direct phonology-semantic or form-meaning interpretation, and, most importantly, (ii) it fails to account for allomorphy

1 Examples presented in this paper will use Yoruba orthography. Vowels ẹ and ọ represent [ɛ] and [ɔ], respectively. *n* after a vowel represents nasality (un = [ũ]. The consonant *p* is [k͡p], *gb* is [g͡b]. The acute sign represents high tone, the grave sign represents low tone, and mid tone is unmarked.

2 These examples and others presented in this study are the basic counting pattern of Standard Yoruba, which differs from other numerical functions such as the long (e.g., eéjì), the cardinals (e.g., méjì), and the ordinals (e.g., kejì).

that seems to not follow regular phonology of the language, such as the realization of 100 as *ọgọ́rùn-ún* in isolation in (c) but as *ẹ̀ẹ́* in 500 in (i) (see Table 8 for multiples of 100 indivisible by 200). These variations, in addition to the complex mathematical operations involved, pose a challenge to the learnability of Yoruba numerals for native speakers and non-native speakers alike, and have contributed to the endangerment of the numeral system, described as "the aspect of the language that is most susceptible to falling out of use" (Eludiora 2017:862). These difficulties call for an approach that makes a direct connection between Yoruba numerical forms and their meaning.

More recently, Booij, through a series of work (Booij 2005, 2007, 2009, 2010, 2018), developed Construction Morphology (CxM) to account for the holistic properties of complex words by making a connection between the (output) form of words and their meaning. CxM represents complex words as constructions and accounts for their holistic properties through the use of a constructional schema. Constructional schemas show a form-meaning mapping of complex words by representing their common and predictable aspects and account for how language users can use existing templates to generate new structures. In doing this, CxM shows that both unproductive patterns, existing word forms with no generative function, and productive patterns, existing templates that speakers can use to generate new forms, can be represented in the same way. The only difference in the representation of both patterns is that in the representation of the former, the legal instances have to be listed in the lexicon, whereas the schema representation of the latter can, in addition to representing existing words, generate new word forms that meet the required constraint (Booij 2018; Booij & Audring 2018; Jackendoff 2022).

As morphology research increasingly considers the application of CxM to different word categories, research focus has extended to provide a form-meaning account of numerals and to account for the complex network of systems involved in their derivation (Appah 2019; Booij 2009, 2010). In pursuit of these concerns, this study presents numeral data from Yoruba. The main goal of this paper is to provide a form-meaning account of Yoruba numerals, demonstrating how they can be analyzed with morphological schemas, as developed in (Booij 2010). In providing a schema representation, the paper categorizes Yoruba numerals into different subgroups according to their commonalities and the mathematic operations involved. Through this categorization, I wish to illustrate that the Yoruba numeral system forms a complex network of systems at different layers, some of which are more productive than others. I conclude by providing potential learning implications for a form-meaning account of Yoruba numerals.

After this introduction, section 2 provides a brief background of Yoruba, describes the sources of data presented in this paper, and provides a brief overview of Construction Morphology. Section 3 offers a detailed account of Yoruba numerical groups, presenting schema representation for each group. Section 4 discusses the issue of idiosyncrasy and regular phonology as related to Yoruba numerals. Section 5 concludes by relating the contribution of Yoruba numeral data to construction theory and by providing directions for future research.

2 Background

Data for this study are from Yoruba, a language spoken in Western Africa, primarily in the Southwestern part of Nigeria, Benin, and Togo. Yoruba belongs to the Yoruboid group of languages within the Volta-Niger branch of the Niger-Congo family (Williamson & Blench 2000). Yoruba is a dialect continuum consisting of several dialects, which Adetugbo (1982) classifies into North-West Yoruba, North-East Yoruba, Central Yoruba, South-West Yoruba, and South-East Yoruba. Due to the influence of slave trade, some dialects of Yoruba also exist in countries like Brazil, Cuba, Trinidad and Tobago, and Haiti. In addition to these dialects, there is a Standard Yoruba (SY) that speakers of all the various dialects understand and that forms a strong integration element in the formation of a common Yoruba identity (Adetugbọ 1973; Falola & Akínyẹmí 2016). This standard variety is the variety learned in school in all Yoruba-speaking communities, the language of the media, and the variety represented in major dictionaries and grammars of the language prepared over the past two centuries and in most pedagogical materials prepared for learners of the language.

Standard Yoruba has a phonological system that distinguishes it from other dialects of the language. This system includes seven oral vowels [a], [e], [ɛ], [i], [o], [ɔ], [u], five nasal vowels [ã], [ɛ̃], [ĩ], [ɔ̃], and [ũ], three tones High, Mid, and Low, and an ATR vowel harmony system. Orthographically, vowels ẹ and ọ represent [ɛ] and [ɔ], respectively. *n* after a vowel represents nasality as in *an* for [ã]. The acute sign [é] represents high tone, the grave sign [è] stands for low tone, and the mid tone is unmarked. The ATR harmony system requires mid vowels [e], [o], [ɛ], and [ɔ] to agree in [ATR] within a phonological word.

The author of the paper is a native speaker of Yoruba. The data presented, which are from Standard Yoruba, and the analysis developed in section 3 are informed by my native speaker's intuition and existing written works on Yoruba numerals such as Abraham (1958), Armstrong (1962), De Gaye & Beecroft (1923), Ekundayo (1977), and Hurford (1975). From the entire numeral system, I selected and grouped number sets according to their mathematical composition (see 3 for the description of each group).

2.1 Construction Morphology

Construction Morphology (CxM) is a theory of linguistic morphology that accounts for the properties of complex words as systematic pairs of linguistic form and meaning. CxM as an offshoot of Construction Grammar applies to the analysis of words, where complex words are analyzed as constructions. The notion "construction" accounts for the morphological structure of existing words and how speakers use existing templates to create new words. A central assumption of CxM is the use of constructional schemas to account for and represent systemic relationships between words' forms and their meaning (Booij 2010, 2018). Booij (2010) presents an example of a systematic form-meaning relation in the analysis of Dutch numeral sets 3-9 and 10 in (a) of Table 2 and the set of numerals 13-19 in (b).

(a) Numerals 3-9			(b) Numerals 13-19		
(i)	drie	'3'	(i)	der-tien	'13'
(ii)	vier	'4'	(ii)	veer-tien	'14'
(iii)	vijf	'5'	(iii)	vijf-tien	'15
(iv)	zes	'6'	(iv)	zes-tien	'16'
(v)	zeven	'7'	(v)	zeven-tien	'17'
(vi)	acht	'8'	(vi)	acht-tien	'18'
(vii)	negen	'9'	(vii)	negen-tien	'19'
(viii)	tien	'10'			

Table 2: Dutch Numerals 1-10 and 13-19 (Booij 2010:195)

Booij represents the common and predictable properties of the closed set of Dutch numerals 13-19 with the schema in (1). He explains that the schema shows the form-meaning relationship between a particular set of numerals with *tien* '10' where dig denotes numerals

3-9. The meaning of this numerical construction is the addition of the number *tien* '10' and any digit between 3-9. Thus, the schema represents how Dutch speakers store this closed subset of numerals in the lexicon.

(1) *Schema for Dutch Numerals 13-19* (Booij 2010:196)

$[[x]^{i}Num, [dig]\ tien]^{j}Num \longleftrightarrow [NUM^{i} + 10]^{j}$

Constructional schemas help to express generalizations over sets of words. This generalization can make predictions of how language users can derive new words through existing word structures (productive pattern) or only express the common and predictable features of a group of words (unproductive pattern), such as Dutch numbers 13-19 in Table 2. Schemas show that the form and meaning of every subset of numerals, as in other word categories, are not completely arbitrary, but that a relationship exists between their output forms and meaning. This form-meaning mapping of numerals is established not as a concatenation of morphemes, but as "independent meaningful units within which certain subcomponents (morphemes) may be distinguished on the basis of the paradigmatic relations with other words" (Booij 2018:4-5).

In the following section, I will present a form-meaning account of Yoruba numerals. Through the use of constructional schemas, I will show how phonological properties and arithmetic patterns distinguish a numerical group from another.

3 Yoruba Numerals

The Yoruba numeral system is intricate and unique as a result of the complex patterns involved in its derivation. Hurford (1975), notes that, in contrast to other languages, Yoruba numerals pose a serious challenge to giving an adequate theoretically motivated analysis. Other theory-driven works such as Ajiboye (2016), Awobuluyi (2016), Ekundayo (1972, 1977), and descriptive works such as Babarinde (2013), Bamgbose (1966, 1986) , De Gaye & Beecroft (1923), Mann (1887), and Olubode-Sawe (2016) allude to the complexity of the Yoruba numerals, which (Bamgbose 1986:35) refers to as "the problems of Yoruba numerals". According to Armstrong (1962:5), the "Yoruba numeral system is a fascinating chapter in the history of mathematics and of the development of human thought". Awobuluyi (1994:33) refers to the derivations of these numerals as done in a "very cumbersome and complicated manner involving multiplication, addition, and subtraction". Some numerals are derived through simple addition, subtraction, or multiplication. Others are derived through a concatention of numerical operations.

3.1 Basic Layer of Structures

The most basic uninflected numerical forms in the Yoruba system, presented in (a) of Table 3, are 1-10, and isolated numbers 20, 30, 200, and 400. These uninflected numbers form the blocks for the computation of the group of numerals in (b) and other groups presented in this section.

(a) Basic Numerals		Gloss	(b) Numbers 11-14		Computation
(i)	ọ̀kan	'1'	(i)	ọ̀kàn-l-àá	
(ii)	èjì	'2'		1-plus-10	1 + 10
(iii)	ẹ̀ta	'3'		'11'	
(iv)	ẹ̀rin	'4'	(ii)	èjì-l-àá	
(v)	ẹ̀rún/àrún	'5'		2-plus-10	2 + 10
(vi)	ẹ̀fà	'6'		'12'	
(vii)	èje	'7'	(iii)	ẹ̀tà-l-àá	
(viii)	èjọ	'8'		3-plus-10	3 + 10
(ix)	ẹ̀sán	'9'		'13'	
(x)	ẹ̀wá	'10'	(iv)	ẹ̀rìn-l-àá	
(xi)	ogún	'20'		4-plus-10	4 + 10
(xii)	ọgbọ̀n	'30'		'14'	
(xiii)	igba	'200'			
(xiv)	ọ̀ọ́dúnrún	300			
(xv)	irínwó	400			

Table 3: Basic Numerals and numbers 11-14

The group of numerals (b) in the table above is a clear case of complex words involving addition, where the addends precede the number to which they are added. These numbers have features that set them apart from other subgroups in the system. The common part is the form *l-àá* "plus-ten", which in addition to any digit between 1-4 forms numbers 11-14. Consistent with early interpretation as *lé ní ẹ̀wá* "increase by 10" (Armstrong 1962; Bamgbose 1966, 1986; Mann 1887), we can interpret *l-àá* as (+10) by considering *l* as a variant of *lé* '+' and *àá* is an allomorph of *ẹ̀wá*. The realization of *ẹ̀wá* as *àá* in this context results from a regular phonological requirement in Yoruba that deletes intervocalic [w], thereby creating an environment for regressive assimilation. Regardless of the variation between *ẹ̀wá* and *àá*, Yoruba speakers do not need to revert to the form *lé ní ẹ̀wá* to get the meaning of the group of numerals in (b). Instead, they represent the meaning of each numeral in *-làá* as the addition of the numerical digits 1-4 and 10. The morphological schema in (2) expresses the commonalities of the close group of numeral set 11-14 and represents their holistic properties as stored in speakers' mental representation.

(2) *Schema for 11-14*

a. $[[\text{Num}]_{[\text{dig}1\text{-}4]}{}^{i}\text{l}^{j}\text{àá}^{k}]^{l}_{\text{Num}}$ ↔ $[\text{NUM}^{i} +^{j} 10^{k}]^{l}$

b. $[[\text{ẹ̀tà}^{i}]\text{l}^{j}\text{àá}^{k}]^{l}_{\text{Num}}$ ↔ $[3^{i} +^{j} 10^{k}]^{l}$

The schema in (2a) represents the connection between a single form and meaning. The meaning of the whole construct is the addition of the numerical value of the closed digital numbers 1-4, represented with [Num][dig1-4], and *àá* '10'. The superscripts *i*, *j*, and *k*, which are variables for lexical indexes, establish a correspondence- the arithmetic value- between the form and the meaning. Num with the index *i* and the subscript [dig1-4], which stands for digital numerals 1-4, on the left side corresponds with NUM with the index *i* on the right, -l- with the index *j* corresponds with +, and *àá* with the index *k* corresponds with the numerical value 10. Num and the superscript *l* at the end of the schema indicate that the entire construction is a numeral. (3a) will express a number like *ẹ̀tàlàá* '13' in (b). The form *ẹ̀tà* on the left side of the schema corresponds with 3 on the right side, *l-* corresponds with +, and *àá* corresponds with 10. The same schema can extend to other numerals within the 11-14 system. Replacing *ẹ̀tà* '3' with *ọ̀kan* '1', *èjì* '2', or *ẹ̀rin* '4' will express 11, 12, and

14, respectively. Note that the schema above only expresses the common and predictable properties of numbers 11-14 and serves no derivation function because this is a closed set of numerals with an unproductive word formation pattern.

3.1.1 Vigesimal System

The next sets of multiples of 20, 200, and 2,000 in Tables 4, 5, and 6 show the preponderant use of vigesimal system and multiplication in the Yoruba numerical system. For multiples of 20 such as 40, 80, 100...200[3] in Table 4, 20 is multiplied by any number between 2-10. Notably, we see allomorphy for 20 and 2-10 in isolation and when these are constituents of multiples of 20. 20 is *ogún* in isolation in (a) but *og/ọg* as parts of multiples of 20 in (b-j) and *ọ* appears invariantly as the first vowel of 2-10 (note that the initial vowel of 2-10 in isolation in Table 2 is either *e/ẹ*). This allomorphy results from the Yoruba phonological constraints that disallow vowel hiatus and require vowel harmony, the co-occurrence of mid vowels with identical [ATR] values, within a word (Ajiboye 2016; Archangeli & Pulleyblank 1989; Bamgbose 1965; Ola Orie & Pulleyblank 2002; Olubode-Sawe 2016; Owolabi 1989; Oyelaran & Bamgbose 1972; Pulleyblank 1988; Seidl 2000). This phenomenon is common in the Yoruba grammar as evidenced in "owner-of-X/ possessor-of-X/ seller-of-X" construction, where *alájá* "owner of dog/seller of dog" and *ọlọ́gbọ́n* "owner of wisdom/ a wise person" are each constituent of *oní* "owner" and *ajá* "dog" and *ọgbọ́n* "wisdom" (Akinlabi 2022). Despite the variation between *ogún* and *og/ọg* for 20, there is a consistent use of the latter for 20 within the group of numerals that are multiples of 20. The choice between *og* and *ọg* resolves to vowel harmony requirements, where the former surfaces when all vowels are (+ATR) as in *ogóje* '40' in (g) and the latter selects all (-ATR) vowels as in *ọgọ́jọ* '160' in (h).

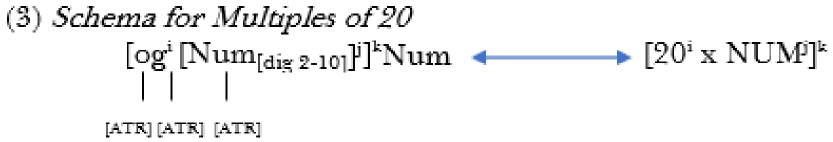
	Numerical Form	Computation		Numerical Form	Computation
(a)	ogún '20'		(f)	ọg-ọ́fà twenty-six '120'	(20 x 6)
(b)	og-óji twenty-two '40'	(20 x 2)	(g)	og-óje twenty-seven '140'	(20 x 7)
(c)	ọg-ọ́ta twenty-three '60'	(20 x 3)	(h)	ọg-ọ́jọ twenty-eight '160'	(20 x 8)
(d)	ọg-ọ́rin twenty-four '80'	(20 x 4)	(i)	ọg-ọ́sànán twenty-nine '180'	(20 x 9)
(e)	ọg-ọ́rùnun twenty-five '100'	(20 x 5)	(j)	ọg-ọ́wàá twenty-ten '200'	(20 x 10)

Table 4: Multiples of 20

The regularities across the group of numerals above, despite the variation for 20, show that Yoruba users do not need to revert to the form *ogún* to interpret the phonological form of

[3] A variant of *ọgọ́wàá* (20 x 10) "200" that is commonly used within the numeral system, especially for the computation of multiples of 200, is *igba*.

numbers that are multiples of 20. Rather, they make meaning connection with *og* or *ọg* as 20. We can, therefore, represent this group of numerals with the schema in (3). As with (2), the schema does not have a generative power but only expresses the common and predictable properties of multiples of 20.



The schema above shows the relationship between the output form of multiples of 20 and their meaning. The left side represents the phonological form, while the right side expresses their semantic interpretation. *og* with the index *i* on the left side of the schema coindexes with 20 on the right side, to which we multiply numbers 2-10, represented with Num[dig 2-10] with the index *j*. The variable [ATR] indicates that all vowels within this numerical form must be identical in ATR value. The constraint, for example, permits *ọgọ́jọ* but disallows **ogójọ* for 160. Replacing the index Num[dig 2-10] with 2, 3, or 4, will express *ogóјì* '40', *ọgọ́t*a '60', and *ọgọ́rin* '80', respectively.

The set of multiples of 200 in Table 5 employs the same arithmetic operation as multiples of 20. In this group, 200 serves as the base through which we multiply digits 2-10. Similar to the observation in multiples of 20, 200 appears as *igba* in isolation in (a) but as *ẹgb/egb* in multiples of 200 in (b-j). As with 20 in multiples of 20, this variation results from the phonological constraints of Yoruba that disallow vowel hiatus and require vowels with identical [ATR] values to co-occur within a word.

	Numerical Form	Computation		Numerical Form	Computation
(a)	igba	200	(f)	ẹgbẹ̀-fà two hundred-six '1,200'	200 x 6
(b)	egbè-jì two hundred-two '400'	200 x 2	(g)	egbè-je two hundred-seven '1,400'	200 x 7
(c)	ẹgbẹ̀-ta two hundred-three '600'	200 x 3	(h)	ẹgbẹ̀-jọ two hundred-eight '1,600'	200 x 8
(d)	ẹgbẹ̀-rin two hundred-four '800'	200 x 4	(i)	ẹgbẹ̀-sán two hundred-nine '1,800'	200 x 9
(e)	ẹgbẹ̀-rún two hundred-five '1,000'	200 x 5	(j)	ẹgbẹ̀-wá two hundred-ten '2,000'	200 x 10

Table 5: Multiples of 200

We can represent the group of numerals above with the schema in (4). In the schema, the phonological form *egb* on the left side coindexes with 200 on the right side. Num[dig 2-10] with the index *j* denotes the set of digital numbers 2-10 with which we multiply 200. The variable [ATR] expresses the requirement for the co-occurrence of vowels with identical ATR values within a numerical form. This constraint explains why we have *egb* in *egbèje* '1,400' but *ẹgb* in *ẹgbẹ̀jọ* '1,600'. Replacing the index NUM[dig2-10] with 2, 3, or 4 will express *egb-èjì* (200 x 2) '400', *ẹgb-ẹ̀ta* (200 x 3) '600', and *ẹgb-ẹ̀rin* (200 x 4) '800', respectively.

(4) *Schema for Multiples of 200*

[egb^{i} [Num$_{[dig\ 2\text{-}10]}$]^{j}]^{k}Num ⟷ [200^{i} x NUM^{j}]^{k}

[ATR] [ATR] [ATR]

The number *ẹgbàá* '2000', which in itself is an allomorph of *ẹgbẹ̀wá*, serves as the base for the computation of multiples of 2,000 in Table 6. As with multiples of 20 and 200, 2,000 is multiplied with the digits 2-10 to derive 4,000, 6,000, 8,000, up to 20,000. Based on this computation, we can present a form-meaning representation of this group of numbers with the schemas in (5). Since this is a closed set of numerals that speakers have to store in their mental lexicon, the schema, as with previous ones, does not have a generative function. Rather, it shows the common and predictable properties of numbers in Table 6 and expresses how Yoruba speakers form a connection between the phonological forms of this group of numerals and semantic interpretations.

	Numerical Form	Computation		Numerical Form	Computation
(a)	ẹgbàá	2,000	(f)	ẹgbàá-fà two thousand-six '12,000'	2,000 x 6
(b)	ẹgbàá-jì two thousand-two '4,000'	2,000 x 2	(g)	ẹgbàá-je two thousand-seven '14,000'	2,000 x 7
(c)	ẹgbàá-ta two thousand-three '6,000'	2,000 x 3	(h)	ẹgbàá-jọ two thousand-eight '16,000'	2,000 x 8
(d)	ẹgbàá-rin two thousand-four '8,000'	2,000 x 4	(i)	ẹgbàá-sànán two thousand-nine '18,000'	2,000 x 9
(e)	ẹgbàá-rùnún two thousand-five '10,000'	2,000 x 5	(j)	ẹgbàá-wàá two thousand-ten '20,000'	2,000 x 10

Table 6: Multiples of 2,000

The schema in (5) represents the meaning of each number in *ẹgbàá* as the multiplication of any of the digital numerals 2-10 and 2,000. *ẹgbàá* on the left side of the schema co-indexes with 2,000 on the right. NUM[dig2-10] with the index *j* denotes the set of numerals 2-10. If we replace the index Num[dig2-10] with a specific number such as 3, we will express the meaning of *ẹgbààta* as (2,000 x 3) '6,000'.

(5) *Schema for Multiples of 2,000*

[ẹ̀gbàá^{i} [Num$_{[dig\ 2\text{-}10]}$]^{j}]^{k}Num ⟷ [2000^{i} x NUM^{j}]^{k}

The groups of numerals presented in this section use multiplication. Multiplication is not morphologically marked in Yoruba but inferred in complex numerals with bases that are of vigesimal value (multiples of 20, 200, and 2,000) and digital numerals 2-10. The next set of numerals, which operate within the vigesimal system, uses subtraction. Unlike numerals derived through multiplication, the meaning of subtraction is expressed with the morpheme *dín/d*.

3.1.2 Subtraction within Vigesimal

Numeral sets presented in this section show how derived complex words can serve as the base for morphological operations. The derived multiples of 20, 200, and 2,000 in Tables 4, 5, and 6 above serve as bases for multiples of 10 that are indivisible by 20 (50, 70, 90...190), multiples of 100 that are undividable by 200 (500, 700, 900…1,900), and multiples of 1,000 that are indivisible by 2,000 (5,000, 7,000, 9,000… 19,000).

For multiples of 10 indivisible by 20 in Table 7, we remove 10 from the next multiple of 20 (60, 80, 100...200) such that *àádọ́ta* '50' is (60-10). Notably, 10 which is *ẹ̀wá* in isolation in (a) and appears as *àá* in numbers 11-14 in Table 3 resurfaces here as *àá* in (b-i); "minus" appears as *d* but as *dín* in isolation (see Table 10); the multiples-of-20 template to which 10 is subtracted is the last CVC of the base in isolation in Table 4. Despite these variations, we see the regular use of *àá* for 10, *d* for minus, and the VCV template for multiples of 20.

	Phonological Form	Computation		Phonological Form	Computation
(a)	ẹ̀wá '10'	10			
(b)	àá-d-ọ́ta 10-minus-60 '50'	60-10	(f)	àá-d-óje 10-minus-140 '130'	140-10
(c)	àá-d-ọ́rin 10-minus-80 '70'	80-10	(g)	àá-d-ọ́jọ 10-minus-160 '150'	160-10
(d)	àá-d-ọ́rùnún 10-minus-100 '90'	100-10	(h)	àá-d-ọ́sànán 10-minus-180 '170'	180-10
(e)	àá-d-ọ́fà 10-minus-120 '110'	120-10	(i)	àá-d-ọ́wàá 10-minus-200 '190'	200-10

Table 7: Multiples of 10 indivisible by 20

The consistent use of *àá* '10' as an allomorph of *ẹ̀wá* and *d* 'minus' as a variant of *dín* shows how multiple phonological forms can represent a single numerical meaning. This is consistent with other languages such as Dutch and English, where Dutch speakers represent the digit 4 as *vier* /vi:r/ in isolation but as *veer* /ve:r/ as a stem for the formation of 14 (Booij 2010) and English speakers represent the number 10 as *ten* in isolation but as *-teen* in numbers 13-19 and as *-ty* in multiples of 10 (Armstrong 1962). For Yoruba speakers, 10 is represented as *àá* within numbers 11-14 and multiples of 10 that are indivisible by 20 but as *ẹ̀wá* in isolation. Thus, we can express the holistic properties of the data above with the schema below.

(6) *Schema for multiples of 10 indivisible by 20*

$$[\text{àá}^{i}\text{d}^{j}\text{Num}_{[x20]}{}^{k}]^{l}\text{Num} \longleftrightarrow [\text{NUM}^{k} \;{}^{-j}\, 10^{i}]$$

The phonological form *àá* on the left side of the schema coindexes with 10 on the right side; *d*, 'reduce/minus' coindexes with (-); the subscript [x20] denotes that Num with the index *k* represents multiples of 20 to which we subtract 10. Notably, we see a positional variability between the phonological representation on the left side of the schema, where the subtrahend precedes the minuend, and the semantic interpretation on the right side, where the computation requires the minuend to precede the subtrahend. This form-computation order mismatch is adequately captured and annotated with the use of superscripts: Num[x20], indexically marked with *k*, on the left side of the schema corresponds with NUM on the right side, with [x20] denoting it as a multiple of 20; *d* marked with the index *j* corresponds with (-); *àá*, marked with the index *i*, which appears at the beginning on the left side of the schema corresponds with 10 at the end on the right side.

Multiples of 100 that are indivisible by 200 in Table 8 rely on multiples of 200 for their computation. For this set of numerals, we remove 100 from the next multiple of 200 such that *ẹ̀ẹ́dẹ́gbẹ̀ta* '500' in (b), for example, is 600-100. Unlike the phonological explanation posited to have accounted for the realization of *ẹ̀wà* '10' as *àá*, the source of *ẹ̀ẹ́* in this subgroup is not phonologically governed (since 100 is *ọgọ́rùn-ún*), but rather appears to be idiosyncratic. Regardless of the source, however, there is a direct form-meaning association for *ẹ̀ẹ́* as, in relation with multiples of 200, expressing 100. Phonologically, *ẹ̀ẹ́* can appear as either *èé* as in (f) or as *ẹ̀ẹ́*, depending on the [ATR] value of other vowels in the number. When other vowels are [-ATR], *ẹ̀ẹ́* surfaces, whereas *èé* appears with a number whose other vowels are [+ATR].

	Numerical Form	Computation		Numerical Form	Computation
(a)	ọgọ́rùn-ún '100'	100			
(b)	ẹ̀ẹ́-d-ẹ́gbẹ̀ta 100-minus-600 '500'	600-100	(f)	èé-d-égbèje 100-minus-1,400 '1,300'	1,400-100
(c)	ẹ̀ẹ́-d-ẹ́gbẹ̀rin 100-minus-800 '700'	800-100	(g)	ẹ̀ẹ́-d-ẹ́gbẹ̀jọ 100-minus-1,600 '1,500'	1,600-100
(d)	ẹ̀ẹ́-d-ẹ́gbẹ̀rún 100-minus-1,000 '900'	1,000-100	(h)	ẹ̀ẹ́-d-ẹ́gbẹ̀sán 100-minus-1,800 '1,700'	1,800-100
(e)	ẹ̀ẹ́-d-ẹ́gbẹ̀fà 100-minus-1,200 '1,100'	1,200-100	(i)	ẹ̀ẹ́-d-ẹ́gbẹ̀wá 100-minus-2,000 '1,900'	2,000-100

Table 8: Multiples of 100 indivisible by 200

We can extend the schema in (6) to (7) to express the holistic properties of the data above. Num with the subscript [x200] stands for multiples of 200. The subscript [x200] stipulates the constraint governing the number that can appear in this position as multiples of 200. The meaning of *ẹ̀ẹ́* is expressed through its co-indexation with 100 on the right side of the schema. *d* co-indexes with "minus". We can replace Num[x200] with any multiple of 200 such as *ẹgbẹ̀t*a '600' to express *ẹ̀ẹ́-d-ẹ́gbẹ̀ta* (600-100) '500'. Notably, the same structural-semantic juxtaposition that exists in multiples of 10, where the subtrahend precedes the minued in the phonological form, also exists in multiples of 100. Again, co-indexation adequately captures this order mismatch (*ẹ̀ẹ́* coindexes with 100; *d* coindexes with "-"; Num[x200] coindexes with NUM]. As with the previous schemas, (7) does not serve a generative function, but rather expresses the common and predictable properties of the group of numerals above.

(7) *Schema for Multiples of 100 Indivisible by 200*

$$[\text{ẹ̀ẹ́}^{i}\text{d}^{j}\text{Num}_{[x200]}{}^{k}]^{1} \longleftrightarrow [\text{NUM}^{k}\ {-}^{j}100^{i}]^{1}$$

There is a multi-meaning use of *ẹ̀ẹ́* in multiples of 100 in Table 8 and in multiples of 1,000 in Table 9. The source of *ẹ̀ẹ́* within the overall Yoruba numeral system is generally unknown since 100 is *ọgọ́rùn-ún* and 1,000 is *ẹgbẹ̀rún*. Although we can postulate *ẹ̀ẹ́* to have phonologically developed from *ẹgbẹ̀rún*, its development from *ọgọ́rùn-ún* '100' raises more questions than answers. Nonetheless, its meaning can be expressed when taken holistically within the entire numerical construction: *ẹ̀ẹ́* in relation to multiples of 200 denotes multiples of 100 that are indivisible by 200, whereas it indicates 100 within multiples of 2,000 that are indivisible by 2,000.

	Numerical Form	Computation		Numerical Form	Computation
(a)	ẹgbẹ̀rún '1,000'	1,000			
(b)	ẹ̀ẹ́-d-ẹ́gbàáta 1,000-minus-6,000 '5,000'	6,000-1,000	(f)	ẹ̀ẹ́-d-ẹ́gbàájè 1,000-minus-14,000 '13,000'	14,000-1,000
(c)	ẹ̀ẹ́-d-ẹ́gbàárin 1,000-minus-8,000 '7,000'	8,000-1,000	(g)	ẹ̀ẹ́-d-ẹ́gbàájọ 1,000-minus-16,000 '15,000'	16,000-1,000
(d)	ẹ̀ẹ́-d-ẹ́gbàárùnún 1,000-minus-10,000 '9,000'	10,000-1,000	(h)	ẹ̀ẹ́-d-ẹ́gbàásànán 1,000-minus-18,000 '17,000'	18,000-1,000
(e)	ẹ̀ẹ́-d-ẹ́gbàáfà 1,000-minus-12,000 '11,000'	12,000-1,000	(i)	ẹ̀ẹ́-d-ẹ́gbàáwàá 1,000-minus-20,000 '19,000'	20,000-1,000

Table 9: Multiples of 1,000 indivisible by 2,000

The morphological schema in (8) shows the common and predictable parts of multiples of 1,000 Table 9. In this schema, *ẹ̀ẹ́* as a common feature translates to 1,000 on the right side, Num with the subscript [x2000] denotes multiples of 2,000 from which 1,000 is removed. *d* with the index *j* co-indexes with subtraction. As such, we can replace Num[x2,000] with any multiple of 2,000 such as *ẹgbàáta* '6,000', *ẹgbàárin* '8,000', or *ẹgbàárún* '10,000' to express 5,000 as *ẹ̀ẹ́-d-ẹ́gbàáta* (6,000-,000), 7,000 as *ẹ̀ẹ́-d-ẹ́gbàárin* (8,000-,000), and 9,000 as *ẹ̀ẹ́-d-ẹ́gbàárùnún* (10,000-1000), respectively.

(8) *Schema for Multiples of 1,000 Indivisible by 2,000*

$$[\text{ẹ̀ẹ́}^{i}\text{d}^{j}\text{Num}_{[x2,000]}{}^{k}]^{l} \longleftrightarrow [\text{NUM}^{k} - {}^{j}1000^{i}]^{l}$$

The data presented in 3.1 shows that the Yoruba numerical system is vigesimal based, built upon multiples of 20, 200, and 2,000 and the digits 2-10. Members of each group are closed and cannot be freely generated. Thus, a schema representation of each group is to show common and predictable properties of each subgroup of numerals. The next section discusses a more productive pattern of numeral formation in Yoruba.

3.2 Complex Layer of Structures

The next sets of numerals operate at a more complex layer and have more productive patterns than those discussed in the previous section. These groups of numerals, which rely on the already derived numbers in section 3.1 for their formation, present evidence that complex lexical units can serve as the basis for morphological operations.

For the next subset of numerals in Table 10[4], Yoruba employs addition to and subtraction from higher decades. Armstrong (1962) describes a decade in Yoruba as a set of ten numbers such that each of the number groups 21-30, 31-40, and 41-50 constitutes a decade. This system has more productive patterns, as long as the numerical form observes the constraints governing its realization. Within this decade system, we can add any digit from 1 to 4 to any multiple of 10 such as 20, 30, 40, 50, 120, 300, 1000, 5000, and so on, thereby creating groups 21-24, 31-34, 41-44, 51-54, as exemplified with the data in (a).

(a) Whole numbers increase by 1 to 4			(b) Whole numbers decrease by 5 to 1		
	Numerical Form	Computation		Numerical Form	Computation
i.	ọ̀kan-lé-l-ógún 1-increase-PREP-20 '21'	(20 + 1)	i.	ọ̀kan-dín-l-ógún 1- remove-PREP-20 '19'	(20 - 1)
ii.	èjì-lé-l-ọgọ́ta 2-increase-PREP-60 '62'	(60 + 2)	ii.	èjì-dín-l-ọgọ́ta 2-minus-PREP-60 '58'	(60 - 2)
iii.	ẹ̀ta-lé-l-ăàdọ́rùnún 3-increase-PREP-90 '93'	(90 + 3)	iii.	ẹ̀ta-dín-l-ăàdọ́rùnún 3- remove-PREP-90 '87'	(90 - 3)
iv.	ẹ̀rin-lé-l-ọ́gọ́fà 4-increase-PREP-120 '124	(120 + 4)	iv.	ẹ̀rin-dín-l-ọ́gọ́fà 4-remove-PREP-120 '116'	(120 - 4)
			v.	àrún-dín-l-ẹ́gbẹ̀ta 5-remove-PREP-600 '595'	(600 - 5)

Table 10: Representative Sample of Non-Whole Numbers

Conversely, we can subtract any of the digits 1-5 from the next multiple of 10 to create late members of a decade, that is, numbers in the groups 25-29, 35-39, 45-49, as evident from the data in (b). Note here that '-' appears as *dín* as opposed to *d* in multiples of 10, 100, and 1,000 in Tables 7, 8, 9 and '+' appears as *lé* as against -*l*- in Table 3. These variations show that the numeral subgroup selects a specific allomorph for addition and subtraction in Yoruba. Speakers make form-meaning connections through the holistic properties of numerals, not just a subsect of a word. They represent subtraction as *dín* in the group above but as *d* in multiples of 10, 100, and 100, and addition as *l* in Table 3 but as *lé* in (b) above.

We can represent the addition pattern in (a) above with the schema in (9a) and that of subtraction in (b) with the schema in (9b). (a) represents the form-meaning association of the addition of 1-4 and multiples of 10. The morpheme that expresses addition, lé, on the left side co-indexes with + on the right side. Num[dig1-4] stands for digital numerals, with the subscript [1-4] showing its limitation to 1-4. NUM[x10] with the index *k* represents any multiple of 10. Thus, we can substitute NUM[x10] with numbers like *ogún* '20', *ọgbọ̀*n '30', *ogóјì* '40, *ẹgbẹ̀t*a '600', *ẹgbàáta* '6,000'. Unlike previous schemas that serve a representative function, the schema in (9a) has a generative power, in addition to representing common and predictable properties of numerals in this group. Any numeral that satisfies the constraints, which requires addends to be any multiple of 10 and any digit between 1 and 4, can fit into the Yoruba numerical system.

4 Only a representative sample of this construction is presented here because members of this group are unlimited.

(9) *Schema for the addition of 1-4/Subtraction of 1-5*

(a) Addition: $[\text{Num}_{[\text{dig1-4}]}{}^{i}\text{lé}^{j}\text{Num}_{[\text{x10}]}{}^{k}] \longleftrightarrow [\text{NUM}^{k} +^{j} \text{NUM}^{i}]$
(b) Subtraction: $[\text{Num}_{[\text{dig1-5}]}{}^{i}\text{dín}^{j}\text{Num}_{[\text{x10}]}{}^{k}] \longleftrightarrow [\text{NUM}^{k} -^{j} \text{NUM}^{i}]$

The difference between the schemas in (9a and b) is the arithmetic operation involved and the constraints guiding the generation of each set. (9a) uses *lé* 'addition' and requires the addend to be any digit between 1-4, whereas (9b) uses *dín* 'subtraction' and requires the subtrahends to be digits 1-5. Similar to (9a), the schema in (9b) serves a generative function. Any numeral that satisfies the constraint which requires the subtrahend to be 1-5 and the minuend to be a multiple of 10 can fit into the Yoruba numeral system. For example, replacing Num[dig1-5] with *èjì* '2' and Num[x10] with *ọgóta* '60' will express *èjì-dín-lọ́gọ́ta* (60-2) '58'.

The use of *lé* 'increase/plus' also extends to multiples of 20 between 20 and 80 (*okòó*, a variant of 20, *ójì*, a variant of *ogójì* '40', *ọ̀ta*, a variant of *ọgọ́ta* '60', and *ọ̀rin*, a variant of *ọ̀gọ́rin* '80') and multiples of 100 (200, 300, 400, 500, etc). These data sets are in blocks of (220, 240, 260, 280), (320, 340, 360, 380), (420, 440, 460, 480) and so on. Multiples of 100 in this subgroup are already-derived numbers that can be divided by 100 such as 200, 300, 400, 500, 1000, 1900, 2,000, etc. Although the meaning of the phonological forms for 40, 60, and 80 in this subset of numerals reflects easily from their variants in isolation, the word *okòó* for twenty seems to have derived outside the numeral system from *okò owó* 'bundle of money' (Bamgbose 1966; Olubode-Sawe 2016). The source does not pose any challenge to its interpretation; we can easily interpret *okòó* when taken holistically within a numeral form. The addition of 20 to a number uses the form *okòó*; adding other numbers to 20 uses *ogún*. Table 11 presents examples of numerals within this construction.

	Phonological Form	Computation
a.	okòó-lé-n-ígba 20-increase-PREP-200 '220'	(200 + 20)
b.	òjì-lé-l-ẹ́ẹ̀dẹ́gbẹ̀ta 40-increase-PREP-500 '540'	(500 + 40)
c.	ọ̀tà-lé-l-ẹ́gbẹ̀rin 60-increase-PREP-800 '860'	(800 + 60)
d.	ọ̀rìn-lé-l-ẹ́ẹ̀dẹ́gbẹ̀wá 80-increase-PREP-1900 '1980'	(1900 + 80)

Table 11: Complex Numerals (Multiples of 20 and Multiples of 100)

The computation of the set of numerals above is similar to those in (a) of Table 10 in that they both involve the addition of two groups of numerals. They differ, however, in the arithmetic constraints that govern them. The addends in Table 10 are digital numerals 1-4 and multiples of 10, whereas those in Table 11 are multiples of 20 (from 20-80) and multiples of 100. We can, therefore, extend the schema in (9a) to (10) by modifying its constraint.

(10) *Schema for Multiples of 20 and 100*

$[\text{Num}_{[\text{x20}]}{}^{i}\text{lé}^{j}\text{Num}_{[\text{x100}]}{}^{k}]^{l} \longleftrightarrow [\text{NUM}^{k} +^{j} \text{NUM}^{i}]^{l}$

Num[x20] with the index *i* stands for multiples of 20 such as *okòó* '20' *òjì* '40', *ọ̀ta* '60', and *ọ̀rin* ‘80’. Num with the index *k* and on both sides of the schema represents multiples of 100, such as *igba* ‘200’, *ọ̀ọ́dúnrún* ‘300’, *irínwo* ‘400’, *ẹ̀ẹ́dẹ́gbẹ̀ta* ‘500’, *ẹgbẹ̀ta* ‘600’, *ẹ̀ẹ́dẹ́gbẹ̀rin* ‘700’, *ẹgbẹ̀rin* ‘800’ *ẹ̀ẹ́dẹ́gbẹ̀rún* ‘900’, *ẹgbẹ̀rún* ‘1,000’, and so on. The subscripts 20 and 100 indicate the numerals that can occupy each position in the schema as multiples of 20 and multiples of 100. For example, if we replace Num[x20] with *ọ̀tà* ‘60’ and NUM[x100] with *ẹgbẹ̀rin* ‘800’, we will express *ọ̀tàlélẹ́gbẹ̀rin* [800 + 60] ‘860’ in (c) of Table 11.

A different pattern exists for the sets of numerals that are a combination of multiples of 10 indivisible by 20 (such as 30, 50, 70) and multiples of 100 (such as 100, 200, 300, 400). These numerals are in groups of (210, 230, 250, 270), (310, 330, 350, 370), (410, 430, 450, 470), (510, 530, 550, 570), and so on, and are formed from the combination of multiples of 100 and multiples 20 in Table 12. More specifically, we subtract 10 from the next multiples of 20 and 100 such as 210 *òkòóléníg bá ó dín m-ẹ́wàá* ‘220 RP reduce in ADJ-10’ is (220 - 10); 230 *ọ̀jìléníg bá ó dín m-ẹ́wàá* ‘240 RP reduce in ADJ-10’ is (240 – 10); 250 *ọ̀tàléníg bá ó dín m-ẹ́wàá* ‘260 RP reduce in ADJ-10’ is (260 – 10); 270 *ọ̀rìnléníg bá ó dín m-ẹ́wàá* ‘280 RP reduce in ADJ-10’ (280 – 10). We can replace igba ‘200’ with *ọ̀ọ́dọ́rún* ‘300’, *irínwó* ‘400’ *ẹ̀ẹ́dẹ́gbẹ̀ta* ‘500’, *ẹgbẹ̀ta* ‘600’, or any multiple of 100.

	Numeral Form	Computation
a	okòóléḷẹ́gbẹ̀rin ó dín m-ẹ́wàá 820 RP minus ADJ-10 ‘810’	(820 – 10)
b	òjìlélẹ́gbẹ̀rin ó dín m-ẹ́wàá 840 RP minus ADJ-10 ‘830’	(840 – 10)
b	òjìlélẹ̀ẹ́dẹ́gbẹ̀rin ó dín m-ẹ́sànán 740 RP minus ADJ-9 ‘731’	(740 – 9)
c	ọ̀tàlélẹ́gbẹ̀ta ó dín m-ẹ́jọ 660 RP minus ADJ-8 ‘652’	(660 – 8)
d	ọ̀rìnlélẹ̀ẹ́dẹ́gbẹ̀ta ó dín m-éjì 580 RP minus ADJ-2 ‘578’	(580 – 2)

Table 12: Representative Sample of Multiples of 20 and 100 Minus (10-1)

In principle, this system covers numbers in the groups 331-339, 351-359, 371-379, and so on[5]. Technically, we can remove any digit between 1-10 from the next multiple of 20 and 100 in Table 10. Thus, we can replace 700, in 731 in (b) above, with any multiple of 100, 30 with any multiple of 10 that is indivisible by 20, and 1 with any digit between 1-10. This group of numerals can be represented with the schema in (11):

(11) *Schema for Multiples of 20 and 100 Minus (10-1)*

$[\text{Num}_{[[x100]+[x20]]}{}^{i}\ \text{ó}^{i}\ \text{dín}^{j}\ \text{Num}_{[dig1\text{-}10]}{}^{k}]^{l} \longleftrightarrow [\text{NUM}^{i} -^{j} \text{NUM}^{k}]$

Num[[x100]+[x20]]i stands for multiples of 100 and 20 such as (220, 340, 460, 580). The co-indexation of [NUM[[x100]+[x20]] and *ó* on the left side of the schema shows the interconnectedness of the two: *ó*, a resumptive pronoun, stands for multiples 100 and 20 and connects the phrase with the rest of the construction. *-dín* with the index *j* on the left side of

5 we can replace 300 with any multiple of 100 in this system to give us (231-239, 251-259, 271-279), (431-439, 451-459), (471-479), (531-539, 551-559, 571-579), (631-639, 651-659, 671-679), etc.

the schema coindexes with minus on the right side. Num[dig1-10] with the index *k* on the left of the schema stands for any digit between 1-10. We can replace Num[[x100]+[x20]] with any multiples of 20 and 200 such as *okòólélẹ́gbẹ̀rin* '820' and NUM[dig1-10] with any digit between 1-10 such as *mẹ̀wàá* '10' to get *okòólélẹ́gbẹ̀rin ó dín m-ẹ́wàá* '810'. Replacing NUM[[x100]+[x20] with *ọ̀jìlélẹ́ẹ̀dẹ́ẹ̀gbrin* '740' and Num[dig1-10] with *mẹ́sànán* '9' will express *ọ̀jìlélẹ́ẹ̀dẹ́gbẹ̀rin ó dín m-ẹ́sànán* '731'. Similar to other schemas in this section, this schema has a generative function, as long as the constraint governing its formation is observed.

The subgroup of numerals in Table 12, represented with the schema in (11), is structurally different from other subgroups in the Yoruba numeral system by having the appearance of a phrase. Phrasal appearance is shown through orthographic space between words, a phenomenon not observed in other subgroups, and the use of a resumptive pronoun *ó*. In addition, unlike in other uses of subtraction where there is a positional variability between the phonological forms and mathematical operations, subtrahend and minuend on the left side of the data in Table 12 appear in the same position as the computation on the right side. Therefore, the next step for the analysis of Yoruba numerals should be to explore how these structural differences can offer insight into how each numerical group fits into the architecture of grammar.

4 Idiosyncrasy or Regular Phonology?

The form-meaning account presented in this work shows that various allomorphs exist within the Yoruba numerical system. Examples of allomorphy include the realization of 10 as *ẹ̀wá* or *àá*, 20 as *ogún*, *ọg/og*, or *okòó*, 200 as *igba* or *egb/ẹgb*, 100 as *ọgọ́rùn-ún* or *ẹ̀ẹ́*, and variation for arithmetic operations evidenced in the realization of subtraction as *dín* or *d* and addition as *lé* or *l*. Research on the interaction between phonology and morphology is robust. A notable account of this interface shows that the choice of allomorphs may be governed by phonological conditions (Booij 1998; Rubach & Booij 2001). Certain allomorphs within the Yoruba numeral system fall under those categories whose variations are governed by regular phonology of the language. For example, we can explain the reduction of *ẹ̀wá* to *àá* as a result of the phonological constraint that deletes intervocalic [w], thereby creating an environment for vowel assimilation. Similarly, *ogún-ọg* variation has been posited to result from hiatus resolution and vowel coalescence (Awobuluyi 1967, 1992, 2008) or vowel deletion and cross-consonant vowel harmony (Akinlabi 2022; Ajolore 1972; Bamgbose 1986, 1990). Even though these variations are phonologically governed, we see regularities within a particular group of numerals. This suggests that Yoruba speakers have to associate meaning with a particular form within a group of numerals. For example, speakers assign *ẹ̀wá* to 10 in isolation but represent 10 as *àá* within 11-14 and multiples of 10 that are indivisible by 20. Similarly, they represent 20 as *ogún* in isolation or when it is a base to which we subtract or add 1-5 and as *og/ọg* when 20 is multiplied by any number between 2-10. The form-meaning association is also made for 200, which speakers represent as *igba* in isolation but as *ẹgb/egb* among multiples of 200.

In the same vein, speakers make form-meaning associations for construction-specific numerical forms that seem not to have developed through regular phonology of the language. An Example of such forms include the realization of 20 as *okòó* or *ogún* and 100 as *ọgọ́rùn-ún* or *ẹ̀ẹ́*. This type of variation is construction-specific for which speakers associate each form within a subgroup of numerals. For example, they associate the form *ọgọ́rùn-ún* to 100 in isolation and the phonological form *ẹ̀ẹ́* to the same number among multiples of 100 that are indivisible by 200. Speakers also make a form-meaning mapping with *ẹ̀ẹ́* as 1,000 among multiples of 1,000 that are indivisible by 2,000 and make the meaning distinction through the holistic properties of the whole numeral construction. Yoruba speakers also assign the form *okòó* to 20 when it is added to multiples of 100 but attribute different phonological forms in other numerical contexts.

5 Conclusion

This study contributes to constructional theory by addressing the challenge that the traditional word formation rule has in accounting for Yoruba numerals. Yoruba numerals are unique and intricate in that their formation involves an array of arithmetic operations and relies on a vigesimal system. In these operations, subtraction plays an important role as well as addition and multiplication. The product of these complicated mathematical operations is the creation of groups of complex numerals whose bases have different phonological forms from their corresponding number in isolation. These variations, especially those that seem not to have developed through the regular phonology of Yoruba such as the realization of 100 as either *ọgọ́rùn-ún* or *ẹ̀ẹ́*, present a challenge for an input-output account. To address this challenge, the present study provided a Construction Morphology (CxM) approach to Yoruba numerals. This approach accounts for the holistic properties of each numeral group and shows that the meaning of Yoruba numerals is not taken in isolation. Rather, speakers of the language make form-meaning associations with all numerals that belong to the same group.

The present study, which shows that some subsets of Yoruba numerals have limited members, while other groups have an infinite number of numerals, also strengthens the notion that productive and unproductive word formation patterns can be represented in the same way under the Construction Morphology approach (Booij 2010, 2018; Jackendoff 2022). The only difference is that schemas representing productive patterns also make predictions on how Yoruba speakers can form a long list of numerals, as long as the constraints governing their formation are obeyed. The implication of the productive/unproductive distinction and allomorphy within numerals is that young Yoruba users who have difficulties learning the Yoruba numerical system (Eludiora 2017) and L2 learners of Yoruba may need to learn the system in groups. The group of numerals with less productive power should be learned as lexical items. On the other hand, learning complex layers with more generative patterns should include learning their structures and the specific constraints governing their generation. Future studies could assess the pedagogical effects of this approach.

NOTES

The abbreviations used in this paper are as follows:

RP: Resumptive Pronoun

PREP: Preposition

ADJ: Adjective

ATR: Advanced Tongue Root

References

Abraham, Roy C. 1958. *Dictionary of Modern Yoruba.* London: University of London Press.

Adetugbo, Abiodun. 1982. Towards a Yoruba dialectology. In Adebisi Afolayan (ed.), *Yoruba Language and Literature*, 207–224. Ibadan: University Press Limited.

Adetugbọ, Abiọdun. 1973. The Yoruba language in Yoruba history. In Saburi Biobaku (ed.), *Sources of Yoruba history*, 176–204. Oxford: Clarendon Press.

Ajiboye, Oladiipo. 2016. The Yorùbá numeral system. In Ndimele Ozo-mekuri & S.L. Chan (eds.), *The numeral systems of Nigerian languages*, 189–201. Oxford: African Books Collectives.

Ajolore, Olusola. 1972. When vowel clusters occur in Yoruba., Indiana University Bloomington.

Akinlabi, Akinbiyi. 2022. Cross consonantal vowel assimilation in Yoruba: A review. In Moses E Ekpenyong & Imelda I. Udoh (eds.), *Current issues in descriptive linguistics and digital humanities: A Festschrift in honor of Professor Eno-Abasi Essien Urua*, 3–19. Singapore: Springer.

Appah, Clement K. I. 2019. Ordinal numeral constructions in Akan. *Constructions* 1. 1–12.

Archangeli, Diana & Douglas Pulleyblank. 1989. Yoruba vowel harmony. *Linguistic Inquiry* 20(2). 173–217. http://www.jstor.org/stable/4178624.

Armstrong, Robert G. 1962. *Yoruba Numerals*. Oxford: Oxford University Press.

Awobuluyi, Oladele. 1967. Vowel and consonant harmony in Yoruba. *Journal of African Languages* 6(1). 1–8.

Awobuluyi, Oladele. 1992. Aspects of contemporary standard Yoruba in dialectological perspective. In Ìṣọlá Akínwùmí (ed.), *New findings in Yoruba studies: F. Ọdunjọ Memorial Lectures Organising Committee* (J.F. Odunjo Memorial Lecture 3), 12–35.

Awobuluyi, Oladele. 1994. The development of modern Yoruba. In István Fodor & Claude Hagège (eds.), *Language reform VI: History and future*, 25–42. Hamburg: Helmut Buske Verlag.

Awobuluyi, Oladele. 2008. *Eko Iseda-Oro Yoruba*. Akure: Montem Paperbacks.

Awobuluyi, Oladele. 2016. *Eko Iseda-Oro Yorùbá*. Ibadan: Kingdom Arts Publishing.

Babarinde, Olusanmi. 2013. Linguistic analysis of the structure of Yoruba numerals. *Studies in the Languages of Africa* 45(1). 127–147. //core.ac.uk/download/pdf/234692872.pdf.

Bamgbose, Ayo. 1965. Assimilation and contraction in Yoruba. *Journal of West African Languages* 2(1). 21–27.

Bamgbose, Ayo. 1966. *A Grammar of Yoruba*. Cambridge: Cambridge University Press.

Bamgbose, Ayo. 1986. *Yoruba: A Language in Transition* (J.F. Odunjọ Memorial Lecture 1). Ibadan: Molukom.

Bamgbose, Ayo. 1990. *Fonoloji ati Girama Yorùbá*. Ibadan: University Press Limited.

Booij, Geert. 1998. Phonological output constraints in morphology. In W. Kehrein & R. Wiese (eds.), *Phonology and morphology of the Germanic languages*, 143–163. Tubingen: Max Niemeyer Verlag.

Booij, Geert. 2005. Construction-Dependent Morphology. *Lingue e linguaggio* (2). 163–178. doi:10.1418/20719. https://doi.org/10.1418/20719.

Booij, Geert. 2007. Construction morphology and the lexicon, 34–44. Somerville: Cascadilla Proceedings Project.

Booij, Geert. 2009. Constructions and lexical units, an analysis of Dutch numerals. In Susan Olsen (ed.), *New impulses in word formation* (Linguisitsche Berichte, sonderheft 11), 81–100. Hamburg: Helmut Buske Verlag.

Booij, Geert. 2010. *Construction morphology*. Oxford ; New York: Oxford University Press.

Booij, Geert. 2018. The construction of words: Introduction and overview. In Geert Booij (ed.), *The construction of words: Advances in construction morphology*, 3–16. Cham: Springer.

Booij, Geert & Jenny Audring. 2018. Partial motivation, multiple motivation: The role of output schemas in morphology. *The construction of words: Advances in Construction Morphology* 59–80.

De Gaye, Jules A. & W.S. Beecroft. 1923. *Yoruba Grammar*. London & Lagos: London: Kegan Paul, Trench, Trubner & Co., LTD. & CMS Bookshop. 2nd edn.

Ekundayo, Ayotunde. 1972. *Aspects of underlying representation in the Yoruba noun phrases*. Edinburgh: University of Edinburgh dissertation.

Ekundayo, Samuel Ayotunde. 1977. Vigesimal numeral derivational morphology: Yoruba grammatical competence epitomized. *Anthropological Linguistics* 19(9). 436–453.

Eludiora, Safiriyu. 2017. Development of a Yorùbá arithmetic multimedia learning system. *Universal Journal of Educational Research* 5(5). 862–873. doi:10.13189/ujer.2017.050518. http://www.hrpub.org/journals/article_info.php?aid=5959.

Falola, Toyin & Akíntúndé Akínyẹmí (eds.). 2016. *Encyclopedia of the Yoruba*. Bloomington: Indiana University Press.

Hurford, James R. 1975. *The linguistic theory of numerals*. Cambridge: Cambridge University Press.

Jackendoff, Ray. 2022. Alternative theories of morphology in the Parallel Architecture: A reply to Benavides 2022. *Isogloss. Open Journal of Romance Linguistics* 8(1). 1–10. doi:10.5565/rev/isogloss.250. https://revistes.uab.cat/isogloss/article/view/v8-n1-jackendoff.

Mann, Adolphus. 1887. Notes on the numeral system of the Yoruba nation. *The Journal of the Anthropological Institute of Great Britain and Ireland* 16. 59–64. doi:10.2307/2841738. https://www.jstor.org/stable/2841738?origin=crossref.

Ola Orie, Olanikẹ & Douglas Pulleyblank. 2002. Yoruba vowel elision: Minimality effects. *Natural Language & Linguistic Theory* 20(1). 101–156. https://doi.org/10.1023/A:1014266228375.

Olubode-Sawe, Funmi. 2016. Sources of complexity in the Yorùbá numeral system. In Ndimele Ozo-mekuri & S.L. Chan (eds.), *The numeral systems of Nigerian languages*, 189–201. Oxford: African Books Collectives.

Owolabi, Kola. 1989. *Ìjìnlẹ̀ Ìtúpalẹ̀ Èdè Yoruba: Fónẹ́tíìkì. àti Fonọ́lọ́jì*. Ibadan: Onibonoje Press.

Owolabi, Kola. 2011. *Ìjìnlẹ̀ Ìtúpalẹ̀ Èdè Yoruba: Fónẹ́tíìkì. àti Fonọ́lọ́jì*. Ibadan: Akada Press.

Oyelaran, Olasope Oyediji & Ayo Bamgbose. 1972. Some hachneyed aspects of the phonology of the Yoruba verb phrase. In *The Yoruba verb phrase: papers of the seminar on the Yorba verb phrase, Ibadan, 1-2 April, 1971*, Ibadan: Ibadan Univ. Press; Inst. of African Studies (IAS), Univ. of Ibadan.

Pulleyblank, Douglas. 1988. Vowel deletion in Yoruba. *Journal of African Languages and Linguistics* 10(2). 117–136. https://www.degruyter.com/document/doi/10.1515/jall.1988.10.2.117/html.

Rubach, Jerzy & Geert Booij. 2001. Allomorphy in Optimality Theory: Polish iotation. *Language* 77(1). 26–60. http://muse.jhu.edu/content/crossref/journals/language/v077/77.1rubach.pdf.

Seidl, Amanda. 2000. Yoruba vowel elision and compounding. *University of Pennsylvania Working Papers in Linguistics* 6(3). https://repository.upenn.edu/pwpl/vol6/iss3/15.

Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 11–42. Cambridge: Cambridge University Press.