

LEXICO-SEMANTIC ISSUES IN ONTOLOGICAL ANNOTATIONS OF YORÙBÁ NOUN

Akindele A. Aina

Olabisi Onabanjo University, Nigeria
akindeleaina@oouagoiwoye.edu.ng

Abstract

Different criteria are used by scholars to identify the nouns of Yorùbá, a language spoken in southwestern Nigeria. However, complexities and difficulties resulting from these criteria differences have posed serious problems to corpus knowledge extraction useful for ontological annotations. Reconciling the differences with a simple model that dwells on a semantic load of the lexicon is a gap in Yorùbá language scholarship. This study extracts some selected controversial Yorùbá nouns from Awóbùlúyì (1978; 2013), using Qualia structure (QS) of Generative Lexicon - a model containing parameters that focus on the distributed nature of compositionality to analyse the lexical noun, proposed by Pustejovsky (1991a). The paper reveals that the manner nouns [*kiákíá* and *wéréwéré* (quickly)], and polymorphic nouns [*mo* (I), *mi* (me), *o* (you), *ó* (he/she), *irẹ* (you), *irẹ* (him/her), *won* (them)] items classified as a noun by Awóbùlúyì (1978; 2013) are inconsistent and/or incompatible with the QS used and therefore cannot be modeled as a noun, because they lack independently the qualia properties required of nouns. The study therefore concludes that simple descriptive analysis is efficient enough for ontological annotations.

Keywords: Lexical semantics, Yorùbá nouns, Ontological annotation, Generative lexicon

Àṣamọ̀

(Oríṣíríṣi òṣùwòn ni àwọn onímọ̀ n ló láti fi pín ọ̀rọ̀ orúkọ̀ ní èdè Yorùbá tí í ṣe èdè ihà Gúúsù mọ̀ iwọ̀ oorun ilẹ̀ Nàìjíríà. Síbẹ̀, àwọn isokóra àti iṣòro kan n jeyọ̀ láti inú iyàtò tó wà láààrin àwọn òṣùwòn tí a máa n ló wọ̀nyí tí ó fi jẹ̀ okúta idigbòlù àti ipènjà nlá fún iṣẹ̀ iwádíí i- fa -kókó imọ̀ yọ̀ nínú imọ̀ a-so-kókó-imọ̀-móra. Mímú àwọn iyatò yíi dọ̀gba nípasẹ̀ ilò ilàna ẹ̀rọ̀ aṣàmúlò itumọ̀ ajemọ̀ semántiiki jẹ̀ àláfọ̀ gbòògì kan tó wa nínú iṣẹ̀ iwádíí imọ̀ èdà èdè Yorùbá. Iṣẹ̀ yíi ṣe àkọ̀jọ̀ àwọn ọ̀rọ̀ orúkọ̀ tó láríyànṣíyàn nínú Awóbùlúyì (1978, 2013), ó sì lo ihun Qualia (Qualia Structure) tí inú imọ̀ a-lo òfin- idàró -féyọ̀ ọ̀rọ̀ (Generative Lexicon) – ilàna tó ní òdìwòn agbájumọ̀ ifónka itumọ̀ nínú agbekalẹ̀ ẹyọ̀ ọ̀rọ̀ fún tíú ọ̀rọ̀- orúkọ̀- eléyọ̀ ọ̀rọ̀ palẹ̀ tí Pustejovsky (1991a) gbé kalẹ̀. Bébá yíi ṣáfihàn pé àwọn wúnrèn yíi: *kiákíá*, *wéréwéré*, àti àwọn ọ̀rọ̀ bíi *mo*, *mi*, *o*, *ó*, *irẹ*, *irẹ*, *won* tí Awóbùlúyì (1978, 2013) pé ní ọ̀rọ̀ orúkọ̀ kò bá ilàna itupalẹ̀ *Qualia* yíi mu, pàápàá tí a bá fẹ̀ kó wọn jọ̀ sínú asùwòn àká ọ̀rọ̀ tí a fi n rọ̀ ẹ̀rọ̀ a-so- kókó-imọ̀- móra. Iṣẹ̀ iwádíí yíi parí ẹ̀rọ̀ pé ó sà̀n ká máa lo girámà tó rorùn láti fi sàpẹ̀jùwe tíú àwọn ọ̀rọ̀ Yorùbá palẹ̀, nitorí kó lè rorùn fún sisàmúlò wọn nínú imọ̀ a-so-kókó-imọ̀- móra tí Gẹ̀ẹ̀ṣì ǹ pé ní *ontological annotations*.)

Kókó ọ̀rọ̀: Semántiiki eléyọ̀-ọ̀rọ̀, Ọ̀rọ̀-orúkọ̀ Yorùbá, Asòmọ̀ a-so-kókó-imọ̀- móra, Imọ̀ a-lo òfin- idàró -féyọ̀ ọ̀rọ̀)

1. Introduction

Research activities in Yorùbá language studies (a language spoken in South-western Nigeria) commenced over a decade ago in the efforts of employing the advantages of Information and Communication Technology (ICT) in processing Yorùbá corpus. Many models that have been used for the developed European languages are being adapted for developing the resource scarce languages in Africa including Yorùbá language. The core grammar of the language needed to be structured in a format that can be easily accessed, assessed, and shared by machines. This is the core job of ontology annotations. Ontology takes its existence from semantics and alongside with explicitness of a concept and its relation with other elements in the structure of utterance. However, one of the grammatical elements – the nouns and nominal phrases have been given different analyses using different criteria. These criteria are usually based on functions, dialectical perspectives, syntactic behaviour, and derivation from the canonical form as it relates to meaning at the point of utterance. However, these criteria have not started from the point of semantics as it relates to other criteria. This in effect has brought about complexities in the identification and classification of Yorùbá nouns as contained in extant literature. This has in effect created a lot of problems in coding the grammar of Yorùbá language.

This study therefore demonstrates that some items in the analysis may not be compatible with the computational tools adapted from generative lexicon for analysing lexical classes.

2. Statement of the Problem

Two major thoughts about Yorùbá noun identification and classification exist. The first is the traditional view of noun, occupying the subject or object position in a sentence, pre-

posing a phrase with ‘ta’ or ‘ki’- wh phrases to identify nouns (Bamgbose, 1990). The second thought dwells on evidences from dialects especially in Ondo State of Nigeria and its environment to describe some Yorùbá words which the previous school of thought failed to analyse as noun (Awóbùlúyì 1978; 2013). This dichotomy makes ontological annotation for these grammatical concepts difficult. Ontology involves the extraction of concepts from knowledge resources or corpus in a particular domain and process for both human and machine readability. Since there was no complete agreement on the concepts of Yorùbá noun as a domain, encoding it will bring disjunction or overlapping in defining relations of the instances in the domain.

3. Scope and Limitation of the Study

The scope of this paper considers the semantic load of some selected controversial items (Awóbùlúyì 1978; 2013) regarded as nouns based on their function in the structure. This study complements our previous research reported in another academic outlet which as Aina (2019). The previous study reported the design of ontology models for the classification of Yorùbá nouns according to the two pioneer scholars of Yorùbá grammar. The models are three, namely: Yorùbá Noun Ontology according to Awóbùlúyì (YORNOA), Yorùbá Noun Ontology according to Bamgbose (YORNOB) and the hybrid Yorùbá Noun Ontology (YORNO)¹. As a matter of fact, the analysis of the classification and classification of Yorùbá language noun is not the only point of disagreement between the scholars’ viewpoints. The structure of nominal phrases, phrasal verbs, subjectless sentences, and other issues have been considered from different perspectives, based on the functions and the dialectical influences on the standard language by (Awóbùlúyì 1978; 2013). However, this paper limits itself only to the lexico-semantic analysis of those items regarded as nouns in contrast to the other scholars’ perspectives.

4. Research Questions

The lexico-semantics issues and problems arising from the dichotomy in identifying and classifying Yorùbá nouns for ontological models have evolved the following questions:

- i. What are some of the criteria adopted by (Awóbùlúyì 1978; 2013) in identifying Yorùbá nouns?
- ii. What indication has generative lexicon given to the attributes and properties of any item to be regarded as a noun
- iii. How can the lexical semantics and lexical relation augment precisely without ambiguity, the encoding of nouns for ontology annotation?
- iv. How can the Qualia Structure (QS) of Generative Lexicon (GL) schema ensure simplicity and aid compatibility in processing nouns for Natural language processing (NLP systems)?

5. Research Methodology

To achieve the objectives of the study, qualitative research methods were adopted. The identifying properties and classification criteria for Yorùbá nouns were extracted manually from the prose narratives of (Awóbùlúyì 1978; 2013), explicating the knowledge of Yorùbá nouns. Nine (09) lexical data among those that generated theoretical and dialectal controversies were purposively selected from the texts. The Qualia Structure (QS) frame of the generative lexicon was used as the tool of analysis in order to weigh the semantic load of those selected lexical items so as to determine their suitability as noun.

6. Literature Review

The issues surrounding the knowledge of Yorùbá nouns have been extensively treated by scholars. Some of these works include Johnson (1921), Bamgbose (1966), (1990) Awóyalé (1974), Awóbùlúyì (1978), (2011), (2013). The knowledge contained in these resources is essentially needed for the development of ontology-driven annotation. Johnson (1921) opines

¹ These models (YORNOA) and (YORNOB) have been collaboratively developed by Aina A. A. & Taiwo O. , accepted for publication in Journal of West African Languages and published in Journal of Linguistics Association of Nigeria, 22 (1), 2019 respectively, while (YORNO) was sole authored by Aina A.A. published in KIU Journal of Humanities, Volume 4(4), December 2019 .

that simple nouns are generally formed by prefixing the words ‘a’, ‘o’, ‘lo’, ‘e’, ‘e’, and ‘i’ to a verbal root e.g. a+dé = ‘adé’ (crown), ò+lù = ‘olù’ (a ginlet), o+lò = ‘olò’ (grinder), e+rù = ‘erù’ (a luggage), i+mò = ‘imò’ (knowledge). It admits that Yorùbá is non-inflective hence does not use terminal syllables to distinguish gender like English and other European languages. The work prefixes the words: ‘akò’ (male) and ‘abo’ (female) to make a noun-noun construction for the common term used to differentiate gender, for example ‘akò in’ (a stallion), ‘abo esin’ (a mare), or sometimes use “òkùnrin” or “obìnrin” when it involves human. For example, ‘omò òkùnrin’ (male child) or ‘omò obìnrin’ (female child). However, if we automate the simple rule mentioned here it will generate a lot of unacceptable noun phrases. Lexical annotation is done with attributes and properties to restrict the infinite items that could be generated.

Awóbùlúyì (1978:7) defines nouns as any word functioning as the subject of a verb or the object of a verb or preposition in a grammatical sentence in the language. Using this example:

Òjò ra ìwé
NP buy(past) book
(Ojo bought books)
Òjò is the subject, while ìwé is the object.

Although, Awóbùlúyì (2013) declares that the edition is not a translated version of Awóbùlúyì (1978), some earlier claims are held since there has not been a superior claim over many of these issues from other scholars and critics, even though the yardstick for measuring the superiority were not listed. There are other claims that he revised since he has been able to see new findings on these claims. For example, that ‘dà’ (a question marker) and ‘ńkò’ (a question marker) are verbs has been re-explained in this new version. However, thorough scrutiny revealed that many of the claims in the 1978 edition are still upheld and supported further with somewhat dialectic examples.

Awóbùlúyì (2013) expresses improvements in the criteria for classifying word classes more than in the previous edition. Instead of using function alone as the criteria, he mentioned forms, functions, and syntactic behaviours as additional criteria that must be necessarily observed. He pointed out five different features to identify Yorùbá nouns as follows:

(i) Initialising a word with a high front unrounded ‘i’ vowel as in the following examples:

1. **yàrá/yàrá** (room) **iyàrà/yàrà** (gallow)
iyèpè/yèpè (sand) **ilèkùn/lèkùn** (door)
ibùsùn/bùsùn (bed) **ìgbèsè/gbèsè** (debt)

According to him, all Yorùbá words beginning with vowels must be nouns except àti (and), àní (I mean), àimò (not knowing), àmò (but), àkíkà (loan word from Hausa Hákíkà), àfi (unless/except). The second set of those without initial vowels is commonly used today, but it was not like that at the beginning.

(ii) Underlying ‘i’-vowel insertion

Any word that accepts the underlying ‘i’ vowel insertion word initially is a noun. Consider this:

2 a. **Ìwà tútù dára** (Gentle behaviour is good)
 behaviour calm good
 b. **Ìwà a Tutù dára** (Tutù’s behaviour is good)
 behaviour i- vowel NP good

The Tutù in (b) is a noun because it accepts the underlying ‘i’ vowel.

(iii) Accepting Noun Qualifier

Any word that accepts the qualifier in a structure is a noun e.g.

3. **Ìwa rẹ, ìwá a rẹ** (His/hers behaviour)
 Tirẹ ti irẹ (tirẹ) (His/Hers)

If we accept ‘ìwà’ (behaviour) to be a noun because it takes a qualifier, then ‘ti’ must be a noun too. Consider also:

- | | |
|--|---|
| 4. Kíkọ (Written)
Ó bá rọra jókòó jéjé ẹ
(He/She sits gently) | Sísọ (Spoken)
ó bá rọra jókòó jéjé ẹ ẹ. |
|--|---|

The jéjé (gently) in this structure takes a qualifier 'ẹ', so it must be a noun.

(iv) Taking Noun Modifier/Noun Phrase

Any word, either disyllabic or polysyllabic that is contiguous to noun modifiers or a noun phrase must be a noun as shown in the following examples:

- | | |
|--|---|
| 5. Ìşẹ kọ; ìşẹ ní!
Mo bá Ọjó nìkan nìbè
Méta péré tó fún mi | (not only work but poverty)
(I met only Ọjó there)
(only three is enough for me) |
|--|---|

So ìşẹ (work), ìşẹ (poverty), Ọjó (a name), méta (three), are nouns because they precede the noun modifiers kọ, ní, nìkan, péré.

(v) Being Complement for preposition

Any word that occupies the complement position for the following preposition in any structure is a noun. The prepositions are: bá, fi, tí, fún, ní ati sí. They are underlined in the following examples:

- | | |
|---|--|
| 6. Mo bá <u>Olú</u> ra aşọ bò
Mo fi <u>sùúrù</u> yanjú ọ̀rọ̀ nàà
Mo ti <u>okò</u> lọ idálẹ̀
Mo pa àròwà fún <u>okùnrin</u> nàà
Mo ló, mo sì padà ní <u>kiákíá</u>
Mo şe àbèwò lọ sí <u>idíkò</u> | (I bought cloth for Olú)
(I settled the matter gently)
(I went from farm to traveling)
(I persuaded the man)
(I went, and came back quickly)
(I visited the park) |
|---|--|

The underlined words which are complements to the prepositions indicate them as nouns. However, the author shows that there are other structures that are nouns but which other scholars analysed to be prepositions. They are exemplified below:

- | | |
|---|--|
| 7. Èmi o ní lọ ní <u>tèmi</u> (ti èmi)
Òré mi ni <u>tòótó</u> (ti otító) | (I for myself will not go)
(my true friend) |
|---|--|

The fact that the underlined words i.e. ti and its qualifier 'èmi' stands in complement for the preposition 'ní' makes them nouns, not prepositions or emphatic structures as others analysed it.

(vi) Functioning as the subject or object: Any category that functions as a subject or object in a structure is also regarded as a noun. Examples are:

- | | |
|---|--|
| 8. Bòbì ra mótò, kò kọ ilé
Olẹ̀ jí kòkó
Olóhun kigbé | (Bòbì bought a motor, but he did not build a house)
(The thief stole cocoa)
(The owner shouted) |
|---|--|

Bòbì (name), olẹ̀ (thief), and olóhun (owner) function as subjects, in the structure, whereas ilé (house), kòkó (cocoa) and igbe (noise) are objects.

The author remarked conclusively that the question markers 'kí' (what) can be used to identify nouns but these question markers are not restricted to only nouns as claimed by other Yorùbá Scholars but they can co-occur with verbs.

7. Classification of Yorùbá Noun

The taxonomic hierarchy of Yorùbá noun classification with their instances according to Bamgbose (1990) followed by Awóbùlúyì (2013) is presented next. This is done to have an overview comparison of the two scholars' perspectives.

YORÙBÁ NOUNS

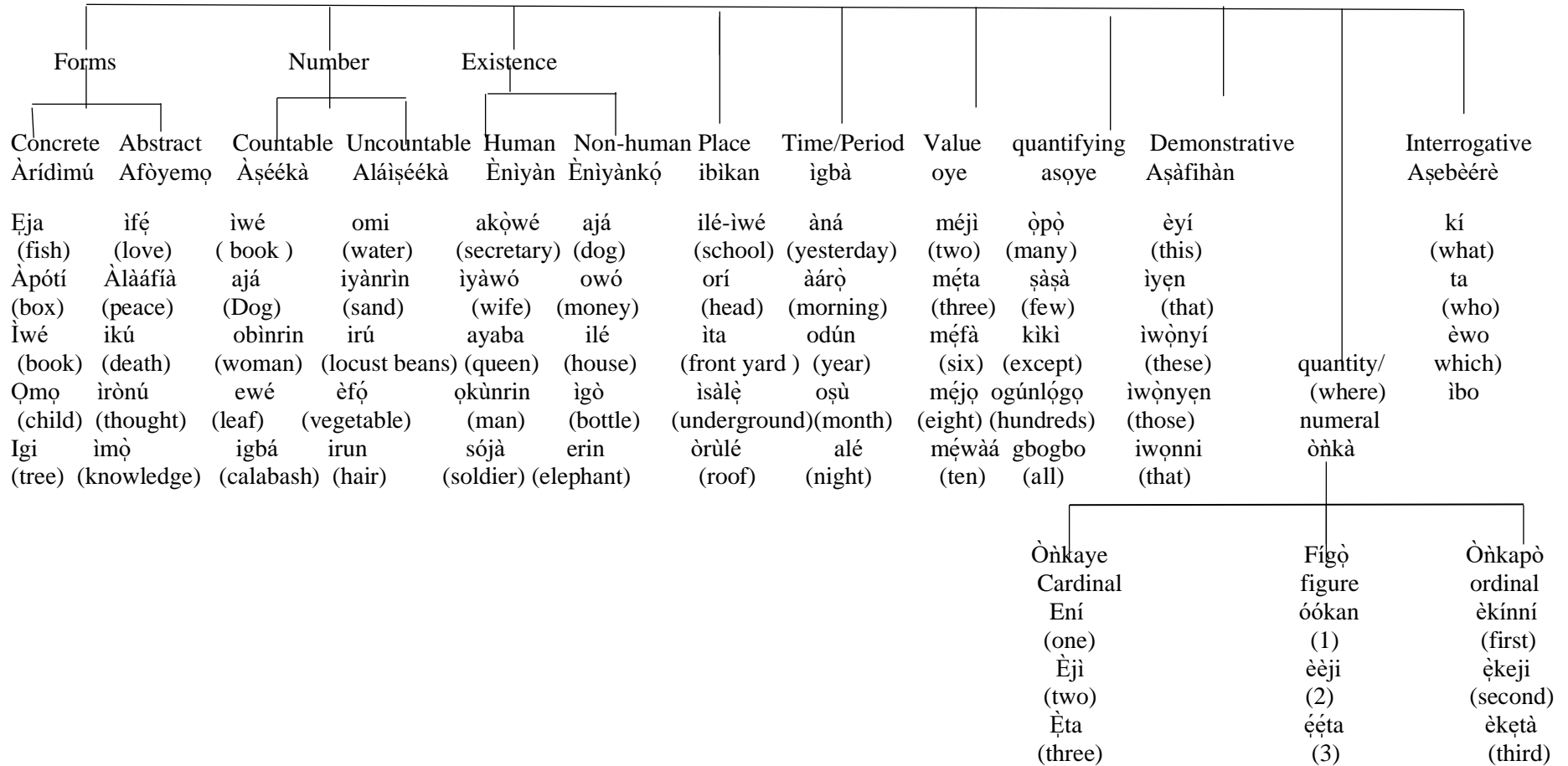


Table 1: Taxonomic Hierarchy of Yorùbá Nouns According to Attributing Features Bamgbose (1990)

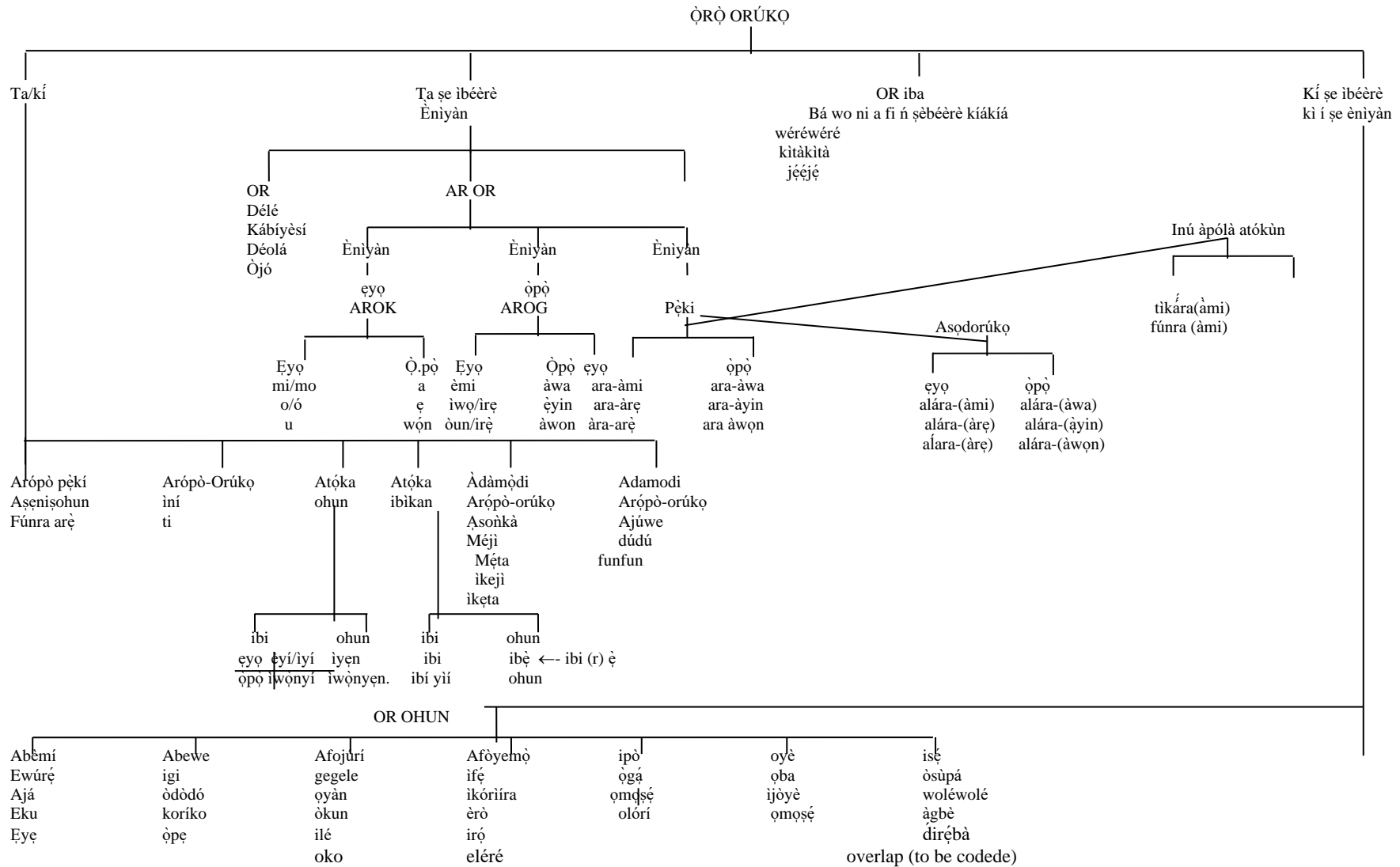


Table 2: Taxonomie Hierarchy of Yorùbá Nouns according to Attributing Features -Awóbùlúyì (2013)

8. The Generative lexicon

Generative Lexicon (henceforth, GL) is a theory of semantics that focuses on the distributed nature of compositionality in natural language. GL is concerned with spreading the semantic load of lexical items across all constituents of the utterance. Generative Lexicon introduces a knowledge representation framework that offers a rich and expressive vocabulary for lexical information. The motivations for this are twofold. Overall, GL is concerned with explaining the creative use of language. (Pustejovsky, 1995). We consider the lexicon to be the key repository holding much of the information underlying this phenomenon. GL was initially developed as a theoretical framework for encoding Selectional knowledge in natural language. This requires making some changes in the formal rules of representation and composition. Following standard assumptions in GL, the computational resources available to a lexical item consist of the following four levels:

- a. LEXICAL TYPING STRUCTURE: giving an explicit type for a word positioned within a type system for the language;
- b. ARGUMENT STRUCTURE: specifying the number and nature of the arguments to a predicate;
- c. EVENT STRUCTURE: defining the event type of the expression and any sub-eventual structure it may have; with subevents;
- d. QUALIA STRUCTURE (QS): Qualia Structure: is the “most lexical” part of GL. It contains prototypical information associated with entities, events, and properties denoted by the words. This information can be defined as the basic syntax of lexical meaning (structural indications allowing to interpret it), parameters that define and constrain lexical semantics, and a structural differentiation of the predicative force behind a lexical item. (Copestake, A., Sanfilippo, A., Briscoe, T., and de Paiva, V, 1991)

The qualia structure, inspired by Moravcsik’s (1975)’s interpretation of Aristotle’s *aitia*, is defined as the modes of explanation associated with a word or phrase in the language, and are defined according to (Pustejovsky, 1991b) as follows:

- a. FORMAL: the basic category which distinguishes the meaning of a word within a larger domain. Formal feature encodes distinctive features of entities (spatial orientation, size, form, dimensionality, color) and their relation with other elements within the same domain which includes taxonomic relation (IS A) ‘is a kind of’ i.e. a house is a kind of building): house: QS formal: the building of size x, of ... shape...
- b. CONSTITUTIVE: the relation between an object and its constituent parts. Constitutive encodes information about the internal constituency of:
 - i- entities: house: QS ...const.: roof, windows, basement...
 - ii-events: number and type of subevents
- c. TELIC: the purpose or function of the object, if there is one. In other words, telic: embraces data about the purpose and the function of entities and events
house: QS ...telic:
- d. AGENTIVE includes the factors involved in the object’s origins or “coming into being”. Precisely, agentive encodes the factors involved in the origin or creation of the entity or the event (creator, artifacts, natural class, causal chain, etc.).

Conventional interpretations of the GL semantic representation have been as feature structures (Pustejovsky, 1995). The feature representation shown in (9) below gives the basic template of argument and event variables, and the specification of the qualia structure.

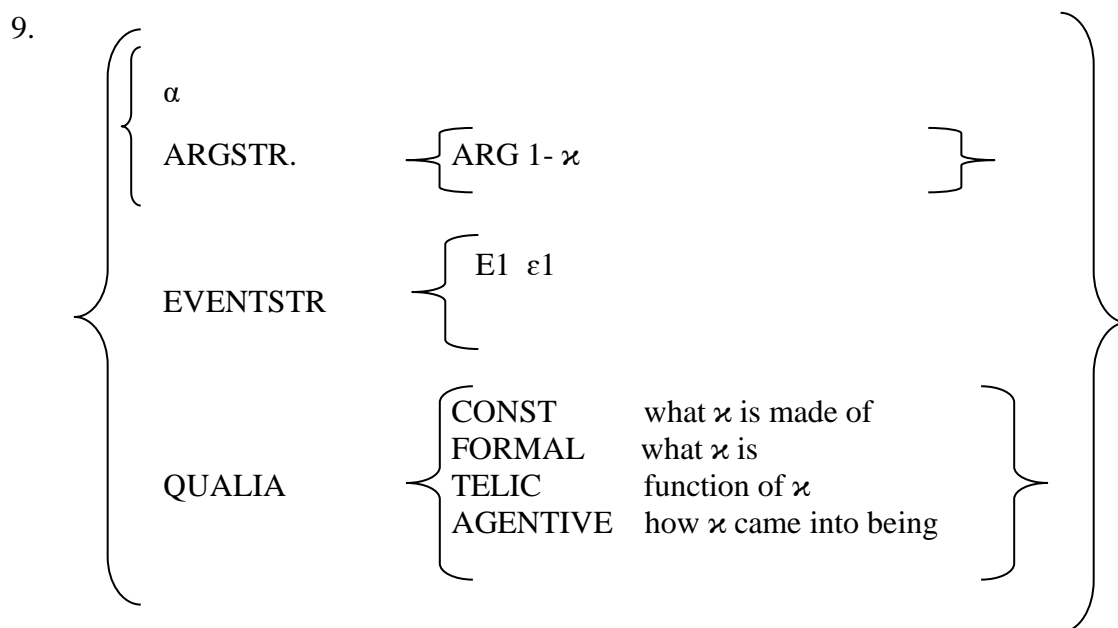
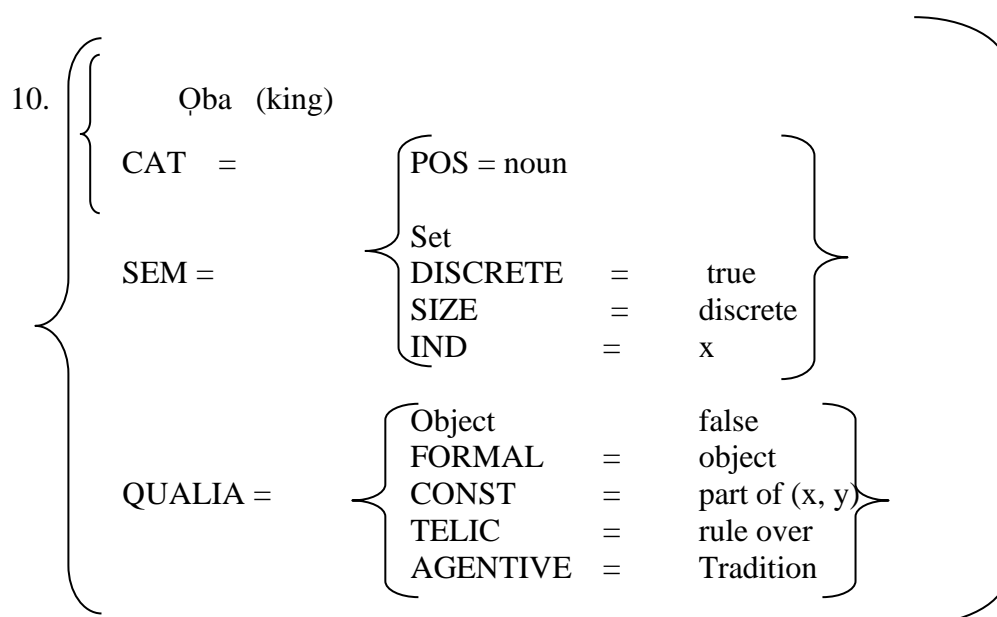


Figure 1: The Template of Qualia Structure (Pustejovsky, 1995)

9. Justification for Generative Lexicon

GL is a model based on the use of feature structures to analyse the semantic load of lexical items. The proliferation of lexical databases in multiple formats has led to growing concerns about the reusability of these lexical resources. However, a framework based on feature structures overcomes most of the problems inherent in classical database models, and it also enables accessing, manipulating or merging information structured in multiple ways. Veronis, J. & Ide, N. (1995). The widespread use in file representation of linguistic information, the applicability of feature structures to lexical databases seems natural, even though this has not yet been implemented so much in African languages. The use of feature structures in lexical databases opens up the possibility of compatibility with computational lexicons.

Having examined this, we shall now begin to examine Yorùbá nouns using GL schema as follows:



11. Omi (water)

CAT	=	[POS	=	noun	
			mass		(No mass)	
DISCRETE	=	[DISCRETE	=	false	
			SIZE	=	non-discrete	
			IND	=	x	Attribute of x
			Substance			
FORMAL	=	[FORMAL	=	Substance	
			CONST	=	part of (x, y)	
			TELIC	=	wash	
			AGENTIVE	=	Natural class	

12. ewúré (goat)

CAT	[POS	-	noun	noun
		Set			
SEM	[DISCREET	-	true	
		SIZE	-	discreet	
		IND	-	... x of kind hooved animal world	
QUALIA	[Object			
		FORMAL	=	fauna; x is a kind of same	
		CONST	=	part of (y: legs, head, hoof x)	
		TELIC	=	graze, meat, domestic	
AGENTIVE	=	natural class			

13.

CAT	ajá (Dog)		
	POS	-	noun
SEM	Set		
	DISCREET	-	true
	SIZE	-	discreet
	IND	-	... x of kind hooved animal world
QUALIA	Object		
	FORMAL	=	fauna; x is a kind of same
	CONST	=	part of (y: legs, head, hoof x)
	TELIC	=	graze, meat, domestic
	AGENTIVE	=	natural class

14.

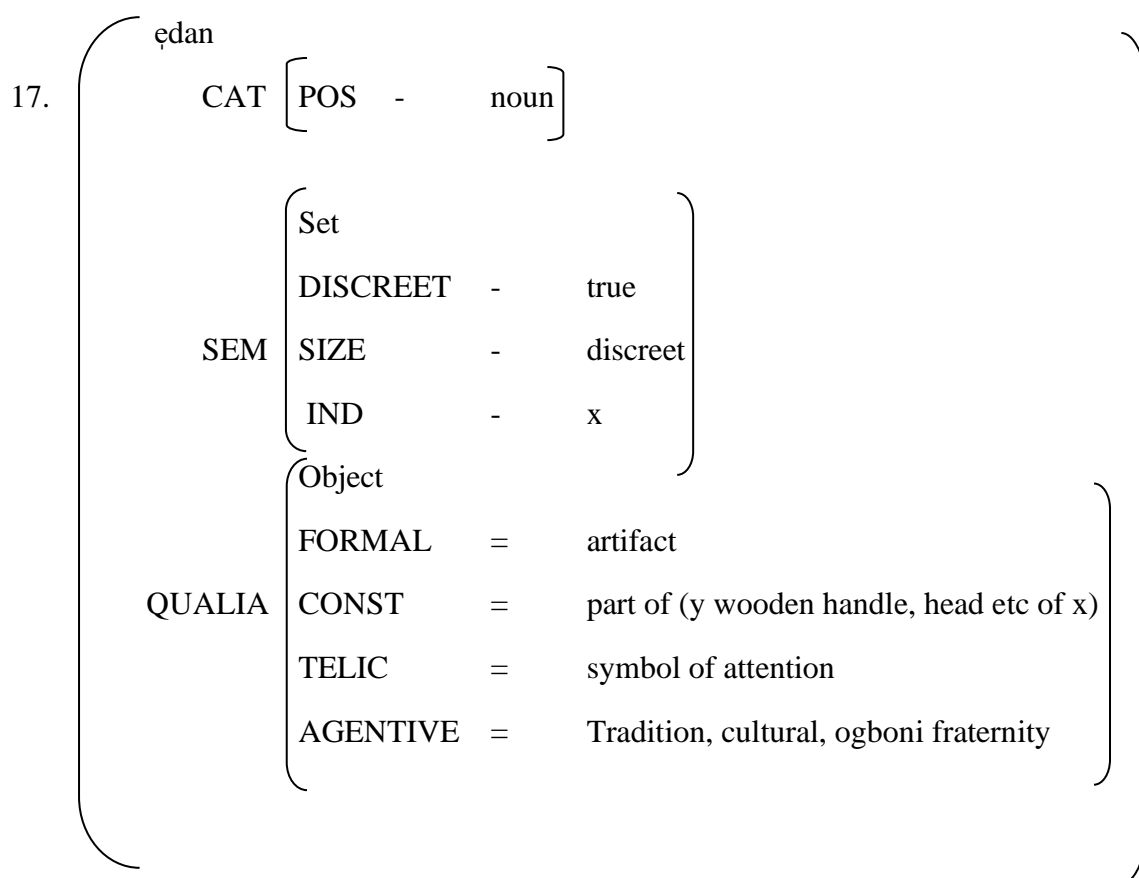
CAT	ẹkùn (leopard)		
	POS	-	noun
SEM	Set		
	DISCREET	-	true
	SIZE	-	discreet
	IND	-	... x of kind wild animal world
QUALIA	Object		
	FORMAL	=	fauna; x is a kind of same
	CONST	=	part of (y: legs, head, hoof x) beak
	TELIC	=	graze, meat, domestic/wild
	AGENTIVE	=	natural class

15. ẹyẹ (bird)

CAT	POS	-	noun	noun		
SEM	Set		DISCREET	-	true	
			SIZE	-	discreet	
			IND	-	... x of kind hooved animal world	Hypernyms
			Object			
QUALIA	Object		FORMAL	=	fauna; x is a kind of same leaves	
			CONST	=	part of (y: legs, head, hoof x) beak	
			TELIC	=	graze, meat, domestic/wild	
			AGENTIVE	=	natural class	

16. Igi (tree). Compare **igi** Arẹrẹ, Apẹpẹ, (hard wood) to **igi** ewédú (stalk of vegetable)

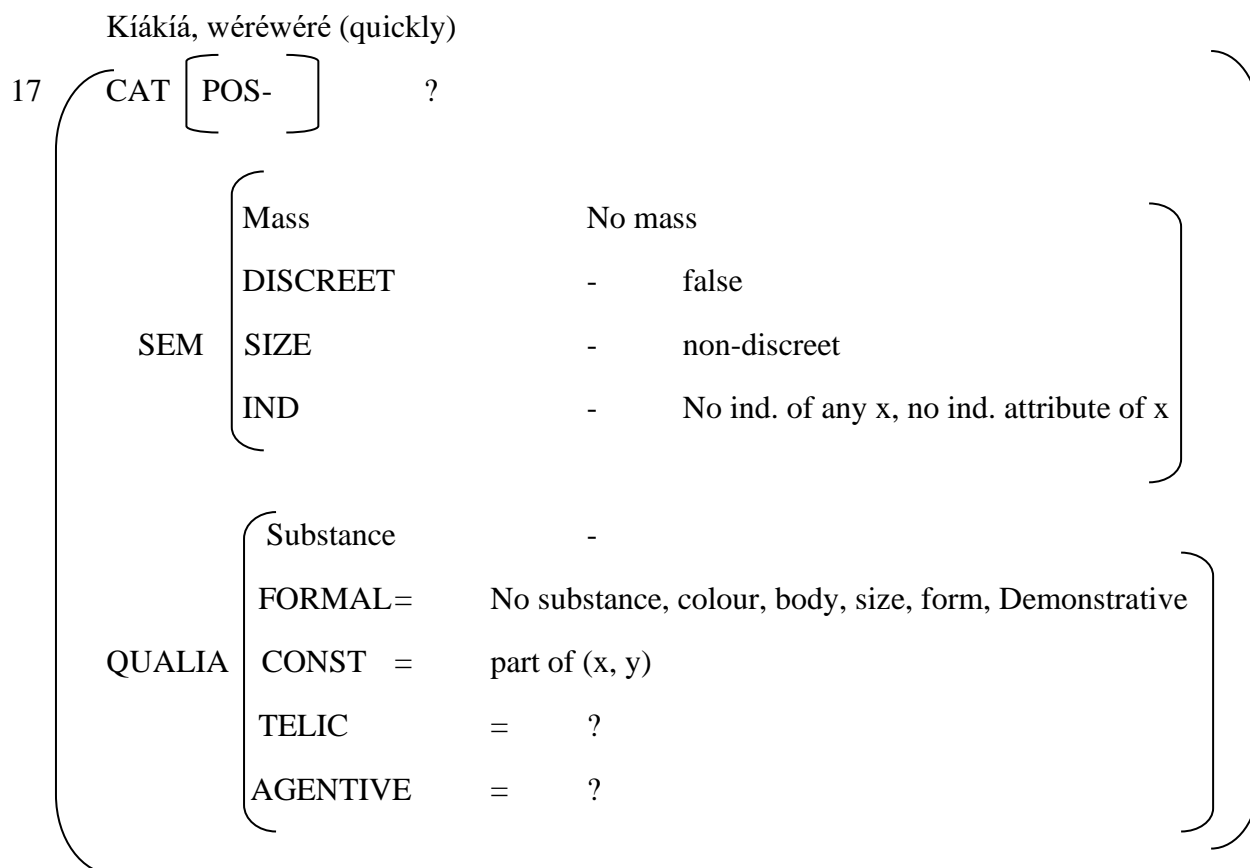
CAT	POS	-	noun		
SEM	Set		DISCREET	-	true
			SIZE	-	discreet
			IND	-	... x of kind
			Object		
QUALIA	Object		FORMAL	=	flora, x is a kind of plant
			CONST	=	part of (y leaves, branches, seed etc of x)
			TELIC	=	wood, shade, forestry
			AGENTIVE	=	natural, forestry etc



GL is similar to the subcategorization principle of syntactic theory but the difference between them is that the subcategorization principle involves the basic cognitive process of arranging items into categories. Verb determines the kind of noun it co-occurs with as well as its grammaticality based on the class and kind of that noun, but G L goes further in that it dwells on the semantic load and properties of each lexicon. These properties will now induce its usage selection.

10. On Awobulyi Manner nouns

Kíákíá, wéréwéré (quickly) are among the controversial lexical items which (Awóbùlúyì 1978; 2013) analysed as a noun because of the functions it performs. Whereas, Bamgbose (1990) classifies these items among the adverbial according to its structure and as it obtains in other languages of the world. The figure below reveals the attribute these items possess when subjected to the QS of lexical semantics.



Observe that the question mark (?) before CAT POS- indicates that the part of speech of the input is yet to be determined. From Yorùbá point of lexical semantics, this item does not have ‘mass’ property, false to ‘discreet’ property, no size, it is not an individual of any ‘X’ either human group or animal or plant, it does not carry any attribute property belonging to any individual of ‘X’ group. In the formal attribute of Qualia, the role of “demonstrative” is encoded for “kíákíá” wéréwéré (quickly). This is to say that we can only assign this attribute to another CAT ‘noun’ in another Q S to perform a modifying role or act as a complement for another CAT “noun”. This is also in addition to the ‘nil’ attribute filled for substance, colour, body size, and form. These attributes and properties are the features that must take a slot before annotation can be performed for any item that must be regarded as a noun. For example, let’s say the CAT ‘noun’ is “ìrìn” (walking), òrò (speech) as in:

ìrìn lèsòlèsò (walking slowly)	vs	ìrìn kíákíá (walking quickly)
ìrìn dẹ̀ngbẹ̀rẹ̀ (walking slowly)	vs	ìrìn wéréwéré (walking quickly)
òrò lèsòlèsò (talking slowly)	vs	òrò kíákíá (talking quickly)

We will have to extend the SEM – IND of the QS for “kíákíá” wéréwéré (quickly) in the above context to read ‘I attribute of x CAT “noun” (def)’. This means that the semantic property and representation of an individual attribute of an (x) noun category is defined in another Qualia structure. This is not at the same level of the category as the noun it modifies or provides information about. More so, annotation tools at present do not accommodate this kind of complex view, rather a straightforward grammar.

11. Polymorphic noun

The items regarded by (Awóbùlúyì 1978; 2013) as polymorphic nouns are: [*mo* (I), *mi* (me), *o* (you), *ó* (he/she), *irẹ* (you), *irẹ* (him/her), *won* (them)]. These items are so classified because they can occupy the subject position and take a direct object like every other noun, whereas other scholars classified these items as pronouns or pronominals.

18	{	<i>mo</i> (I)/ <i>mi</i> (me)/ <i>o</i> (you)/ <i>ó</i> (he/she)		
		CAT [POS-] ?		
		SEM	Mass mass (of reference) DISCREET - True (of reference) SIZE - discreet (of reference) IND - X, Attribute of X (of reference)	}
		QUALIA	Substance - (of reference) FORMAL= (of reference) CONST = part of (x, y) (of reference) TELIC = (of reference) AGENTIVE = (of reference)	}

19	{	<i>irẹ</i> (you), <i>irẹ</i> (him/her), <i>wọn</i> (them)		
		CAT [POS-] ?		
		SEM	Mass mass (of reference) DISCREET - True (of reference) SIZE - discreet (of reference) IND - X, Attribute of X (of reference)	}
		QUALIA	Substance - (of reference) FORMAL= (of reference) CONST = part of (x, y) (of reference) TELIC = (of reference) AGENTIVE = (of reference)	}

A close look at the QS for these lexical items shows that the items will still be problematic in being acceptable in QS. There are positive values for each of these properties, but we have to extend and specify in the 'Const'. as x ref QS that is 'of reference' to another QS which will

have all the QS properties. Why do we have to create too many rules when we can use fewer? The solution is to store the property for each of the nouns independently, after all, it is easier for a machine to read an 'ind' together with its properties and attached to the string or slot that it is meant for rather than to have an 'Ind' in a slot of QS and extends it to another QS.

12. Conclusion

Having analysed Yorùbá nouns on generative lexicon theory which is solely based on the semantics of the lexical items defined, it is observed that some items like kíákíá, wéréwéré, [mo (I), mi (me), o (you), ó (he/she), irẹ (you), irẹ (him/her), wọn (them) which Awóbùlúyì (1978;2013) classifies as nouns will not meet the basic requirements for encoding the lexical properties of noun class which is needed for abstraction representation that would apply for ontological annotations. These items being classified as a noun and polymorphic nouns are simpler and better regarded as Yorùbá adverbials, pronouns, and pronominal for an easy language engineering process. This study therefore concludes that we have to adhere to a simple and straightforward grammar, easily accessible as it obtains from other languages across the world.

References

- Aina, A. (2019). A Hybrid Yorùbá Noun Ontology. *KIU Journal of Humanities*, 4(4), 229-246.
- Aina, A., & Taiwo, P. (2019). Yorùbá Noun Ontology from Functional Perspectives. *Journal of the Linguistics Association of Nigeria*, 22(1), 189-211.
- Awóbùlúyì, O. (1978). *Essentials of Yorùbá Grammar*. University Press Ibadan.
- Awóbùlúyì, O. (2011). *Linguistics and Nation Building*. Ibadan.
- Awóbùlúyì, O. (2013). *Èkọ̀ Ìṣédá Ọ̀rọ̀ Yorùbá*. Akure. Montem Paperbacks.
- Awóyalé, Y. (1974). *Studies in the Syntax and Semantics of Yorùbá Nominalisation*. [Unpublished Ph.D Dissertation]. University of Illinois.
- Bamgbose, A. (1966). *A Grammar of Yorùbá*. Great Britain. Cambridge University Press
- Bamgbose, A. (1990). *Fonoloji ati Girama Yorùbá*. Ibadan University Press Limited.
- Copestake, A., Sanfilippo, A., Briscoe, T., & de Paiva, V. (1991). *The Acquilex LKB: An Introduction. Default Inheritance in Unification Based Approaches to the Lexicon*. Esprit BRA Acquilex Project (Action 3030):
- Veronis, J. & Ide, N. (1995). *A Feature-Based Model for Lexical Databases*. Proceedings of American Computational Linguistics.
- Johnson, S. (1921). *The History of Yorùbás*. Great Britain. Love & Brydone Printers Limited. Thetford Norfolk.
- Moravcsik, J. (1975). *Aitia as Generative Factor in Aristotle's Philosophy: Dialogue*, 14, 622-636
- Pustejovsky, J. (1991a). *The Generative Lexicon: Computational Linguistics* 17(4).
- Pustejovsky, J. (1991b). *The syntax of event structure: Cognition* 4, 47-81.
- Pustejovsky, J. (1995). *A Generative Lexicon*. MIT Press Cambridge, Massachuset.