Resolving Lexical Verb Ambiguity in Yoruba through Semantic Web Annotations

Aina A. Akindele
Department of Linguistics and Nigerian Languages
Ọlabisi Ọnabanjọ University,
Agọ-iwoye, Ogun State, Nigeria.
akindele.aina@oouagoiwoye.edu.ng

Abstract

Lexical verb ambiguity refers to the potential for multiple interpretations of units of speech in languages that renders it difficult or impossible to understand a lexical item without additional information. The goal of semantic web research is to allow the vast range of web-accessible information and services to be more effectively exploited by both humans and automated tools. Technological advancement has given rise to scholars developing automated translation systems from English to the Yorùbá language. However, the problem arises that these tools are not coping easily with semantic nuances found especially in Yorùbá verbs. To explore this problem, a number of Yorùbá lexical ambiguous verbs which may pose serious challenges to the existing machine translation tools were generated and analyzed. To facilitate a solution to the problem, the identified verbs are annotated following the Qualia Structure (QS) of the Generative Lexicon. Protégé 4.5 is employed as the editing tool to map the semantic interpretations of the QS to the semantic web. With the aid of the unique formal specification implemented for each of the lexically ambiguous verbs, the existing automated translation system can now utilize the Internationalized Resource Identifier (IRI) of the Unified Resource Locator (URL) found in the model developed to link the loop and thereby alleviate the problem of ambiguity for the automated translation system. This paper, therefore recommends that more efforts be geared towards ontological

annotation in order to foster more automated tools that will be built for the Yorùbá language.

Keywords: Lexical verb ambiguity, Semantic web annotation, Machine translation tools, Yorùbá Lexical verb.

Introduction

Human language use has changed drastically with the remarkable influence of technology around the world. The invention of information and communication technology has provided us with so many options in the light of making teaching, language learning and language use interesting, as much as making human communication easy and more productive. The role of artificial intelligent systems (AIS), Decision Support Systems (DSS), and Machine Translation (MT) will be one of the most significant drivers of social, political, and linguistic change in Nigerian society. At the moment, the role and status of the Yorùbá language are seeking to be upgraded in its social context, political function, socio-cultural use, education purposes, industries, media, agriculture, and communication across borders with the use of artificially intelligent systems of all sorts including machine translations. Even curriculum and language planners have sought to find a way to include in their documents the challenge of adapting, upgrading and applying technology in the use of our indigenous languages in all ramifications.

Lexical verb ambiguity, also known as semantic ambiguity is the presence of two or more possible meanings for a single word. It is also regarded as a homonym in the literary term. This concept differs from syntactic ambiguity, in that the former is located in a single word, whereas the latter involves the presence of two or more possible meanings within a sentence or sequence of words. Lexical verb ambiguity is sometimes used deliberately to create puns and other types of wordplay. For instance, the word “bank” has several distinct lexical definitions, including “financial institution” and “edge of a river” However; the case is different in the Yorùbá language. There are so many verbs that are polysemous in meaning even though the spelling outlook remains the same. This means that the same word can have several meanings even in a similar grammatical structure. For example, the verb with the mid-tone ‘pa’ can mean ‘kill’, ‘rub something on a surface’ ‘remove the bark of a tree’ ‘hit an item on a surface’, and so on. This phenomenon, therefore, poses serious problems to NLP and web searching agents and activities as we shall explore it and its solution in this paper.

Semantic web annotation /engineering otherwise known as ontological annotation refers to the repository that would attribute structure to the meaningful components of Web pages, providing an environment whereby software agents can move across pages to implement the different productive tasks. The semantic web is not exactly the same thing as the World Wide Web or the internet. Computers could scan Web pages for structure and formal processing but cannot in any reliable form process the semantics of documents. But Semantic Web achieves this in that information it is well-defined semantically; better - enabling humans and computers to work together co-operatively. (Berners-Lee, Hendler, and Lassila 2001),Pareja-Lora (2012). The web or the internet is a database of very versed electronic information. Search engines like Google, Yahoo and so on can access different information through the documents they contain. As veritable and enormous the pieces of information on the web pages, the contents of that information become a hard task for machines to access. Most of this information is human readable. They are in the format for only humans to process. And as much as it is important that an improved activities and web-meaning based processes should be explored, the need to create machine readable content arises. This need is what the semantic web technology fills.

Statement of the Problem

The importance of applying information technology to natural languages for the purpose of development in society, the concepts of Lexical verb ambiguity, and what semantic web engineering involve were discussed in the introduction section as the background need for the conception of this research. However, a number of positive developments have been witnessed despite the difficulty in employing these tools into African languages which are regarded by Adegbola 2009:154 as 'resource scarce languages. Some of the translation tools developed for Nigerian languages includes: Ezeanyeji, Ebinyasi and Mgbeafulike (2019, Eludiora, Agbeyangi and Fatusin (2005), Abiola, Adetunbi, Fasiku and Adenekan(2015), Akinwale, Obe, Adetunji and Adesuyi (2015) . Among other researches in Natural language Processing were (Iyanda 2014), (Adegbola 2006, 2008, and 2009), Hassan (2009), Odoje (2010, 2017), and an introductory work on Yorùbá Ontology development. Aina (2019).

These efforts have contributed in their own ways, to the call to save our indigenous language with the use of modern technology. In spite of these exciting responses, a serious problem arises as to what results when these automated tools are encountered with lexical entries, especially verbs that are polysemy or ambiguous in nature. The dictionary entries of African languages accounted for many of such entries explaining their different meaning and different contextual usages even though the phonetic outlook of the entry remains the same. The questions arising from this situation that begs for answers that this paper focuses on are: What problems of selection arise when

AIS and NLP especially rule-based tools are faced with lexically ambiguous words? What kind of component is present in ontological annotation which could access the root node of these tools so as to address the ambiguity? How can semantic web designs help to alleviate this problem of ambiguity? What implementation method can we adopt to be compatible with the selected samples for the implementation of this model? Lastly, what other applications can utilize this model?

Aims and Objective of this Study

The background needs as explained previously and the problems this paper seeks to address form the basis of the aim of this research that semantic web annotation is a means to an end in natural language processing and this gives credence to the driving objective to isolate some Yorùbá lexical verbs that are ambiguous, defined them within Generative Lexicon in order to implement semantic web formalisms using protégé 4.5 editor in resolving ambiguity challenge to the NLP tools built for Yorùbá language.

Scope and Limitation of this Research

Ambiguity exists in a different grammatical structure such as a single lemma or word, a phrase or a sentence. The scope of this work focuses on ambiguity as it exists in lexical verbs as a single lemma. The gathered data is limited to Yorùbá mono-syllabic lexical verbs as generated from alphabetic order and distribution.Semantic implementation of other forms of ambiguity in the polysyllabic verb, phrasal verbs and others that are concerned in NLP activities is reserved for further investigation.

Materials and Research Methods

To achieve the aim of this research, the following materials discussed in the following sections were employed to carry out the implementation of the model developed in this research.

Protégé

Protégé is an open source semantic web development environment which supports creating, uploading, modifying, and sharing of ontologies developed in the University of Manchester and The Centre for Biomedical Informatics Research, Stanford University. It is fully supported in Web Ontology Language (OWL). This semantic web technology editor possesses direct in-memory connections that enable description logic reasoners. With its customizable user interface, one can create and edit ontologies in any compatible single workspace. This is made possible through visualization tools that allow interactive navigation of ontology relationships. The highly configurable user interface creates the perfect environment for ontology beginners and professionals alike. Protégé 4.5 Desktop can be installed on various platforms like XMaCOSX, Windows, Linux, Unix Platform, Solaris, HPUX, AIX and another JAVA-enabled environment. It is distributed in the form of a ZIP file from the main Protege website, and includes the 64-bit Java Runtime Environment (JRE).

Web Ontology Language (OWL)

The Web Ontology Language (OWL) is an international standard and semantic language for encoding and exchanging ontologies on the World Wide Web. The emphasis on the Semantic Web is such that information could be uploaded for humans to search and make queries ordinarily, but beyond this, information should be given explicit meaning in some formal presentation so that machines can process it more intelligently. Instead of mere developing standard terms for concepts as is usually done in XML, the Semantic Web provides formal definitions for the standard terms developed. Machines can then use inference algorithms to reason about the terms. OWL is expressed in RDF graph and triples which occur in various syntactic forms. The weakness of RDF schema in strong semantic primitives leads to the development of OWL such that each of the important RDF schema terms are either included directly in OWL or they are superseded by new OWL terms.

Methods

The following outlined methods were adopted in order to achieve the aim of this research:

1. Generation of the some purposively selected Yorùbá monosyllabic Lexical Verb (YLV henceforth) in corroboration with lemma entries from Modern YorùbáDictionary (MYD) and Yorùbá Modern Practical Dictionary (YMPD) with the definitions and meaning.
2. The extracted entries were analyzed using the Qualia Structure (QS) frame elements of the generative lexicon which serves as a background scale for attributes in the concepts of each lexical entry therefore revealing the ambiguity in the semantic load of the verbs.
3. The tableau generated for the sampled lexical data is implemented using protégé 4.5 to make an annotated model which projects the attributes of the entries in a formal, machine readable format suitable for AIS and NLP activities.

Literature Review

The focus of this section is to review existing works that are related to the objective of this paper. The concepts of semantic web technology, Natural Language Processing (NLP), tools implemented for Nigerian languages including Yorùbá are different variables in this paper. Hence, a review of these related works establishes this study in its proper place and to carve out our niche among the existing others.

Semantic web engineering and semantic web technology are synonymous terms used to refer to ontological annotations. It aims to define and interconnect data in a structured formalism. The field of semantics in Linguistics has to do with the study of meaning. Kaplan and Bresnan (1989) illustrate the logical underlay of it that:

> The content of a sentence S, in a Language L, relative to a context C, is found by taking the semantic values of the parts of S and combining them in accordance with the semantic and syntactic composition of rules of L.

This implies that the context comprising the unit of words that makes up a sentence has to be taken into consideration before meaning can be derived from any corpus. Meaning is difficult to extract. The semantic content of terms, objects and strings with stated properties will connote what we think of as its semantic content values. Terms, such as semantic annotations, semantic tagging, and semantic mark-up and semantic labelling have existed in literature to refer to this activity of extraction of meaning in corpus. Two major kinds of semantic annotations exist namely: Semantic role labelling and Semantic sense tagging: Semantic role labeling is an activity whereby the semantic relationship held between the agents or patients in a given text is stated in a predicate-argument structure of clauses and sentences. Semantic sense tagging seeks to annotate the differences in the word senses. It involves the matching of open-class words in a corpus with their corresponding meaning being stored in an electronic data repository. The relationship that is commonly found in such annotations is the polysemy, homonymy, synonymy, ambiguity which makes the task fairly difficult to handle. The task includes Word Domain Disambiguation (WDD) or Domain Annotation which aims at tagging words in a corpus with a specialized field tag and sense label. The major benefit of it is that domain annotation reduces lexical ambiguity and this is the highest focus of this paper.

Pareja-Lora (2012:44) refers to Named Entity Recognition and Classification (NERC) as an application which performs more than sense tagging because there are varieties of terms found in texts which the meaning cannot be found in a common dictionary because they are ‘named entities’, but the application caters for these terms. Entities like a person, organization, place, dates, time quantities, monetary numerals, and percentages are uniquely identified by this labelling approach. These values are identified for the application to arrange with the aid of tagging them as a collection of named entities such that the data provided, enables a real-time semantic interpretation of texts. The work provided six layers which activities in semantic web technology and information extraction are divided to, these include Sense Tagging Layer, Concept Semantic Annotation Layer, Named Entity Annotation Layer, Semantic Domain Annotation Layer, Semantic Field Annotation Layer and Semantic Function Labelling Layer. An implementation which runs through all of these layers each with its activities is a major system requirement when a very large annotation corpus is involved. The scope of this paper explores the first two layers as this is the most fitting and appropriate for this work.

In addition to purpose and semantic sense tagging that can be performed for human to take the right decisions, how about the situation whereby the target user of annotations is a machine instead of a human? This is the exact area which SWT annotation addresses. The semantic web is sharply contrasted to the World Wide Web or the internet. The web or internet serves as a repository of huge information. It contains documents which search engines like Google, Yahoo etc can use to access different information. As good and enormous these pieces of information are, the contents of this information become a hard task for machines to access. In other words, most of this information are human readable. They are in a format for only humans to process. And as much as it is important that an improved activities and web-meaning based processes should be explored, the need to create machine readable content arises. Before machine can understand and process the annotations made, they must be encoded in the format, formal and clear vocabularies, syntax, schemes, and tag form meant for it to understand.

Berners-Lee, Hendler, and Lassila (2001) in the pursuit of meeting this need built a semantic whereby not only humans will read the web’s content, but computers’ set of instructions will also manipulate web content meaningfully. Computers, according to this work, could scan Web pages for structure and formal processing but cannot in any reliable form process the semantics of documents. So the semantic Web serves as a repository that would attribute structure to the meaningful components of Web pages, providing an environment whereby software agents can move across pages to implement different productive task. So, the Semantic Web is an extension of the existing one whereby information is well-defined semantically, better enabling human and computer to work together co-operatively and not another repository. This is exactly what this paper achieves in that the lexical verb annotated in the model reduces lexical ambiguity for machines and other web searching agent built for NLP.

With regards to the clarion calls to young scholars to stand up and learn computational tools in order to save our indigenous languages, some individuals have responded with great enthusiasm to these calls. Aina and Taiwo (2021) have highlighted some notable works in this regard. These are Hassan, Odejobi, Ogunfolakan and Adejuwon (2013), which used formal concept to analyse data in Yorùbá cultural domain. Eludiora (2012) adopts phrase structure grammar as the linguistic background to build English to Standard Yorùbá Machine Translation System. Odoje (2010) Rule-Based Machine Translation (RBMT) was complemented with Statistical Machine Translation (SMT) in Odoje (2017). Interestingly, one of the issues raised in Odoje (2017) was the difficulty for an SMT to capture some natural language intricacies, cultural nuances, differentiation in contextual meanings and semantic extension. Odoje (2017) cited examples which highlighted the verb *gbé* in the Yorùbá sentence " Ó gbé ìbọn fún ọdẹ" as being ambiguous. Should it be understood as "He gave the gun to the hunter" or "He shot the security guard". He agrees that issues like this pose a serious challenge to AI tasks including SMT and submits that *'there is the need for a corpus approach where the system would have come across gbé as to shoot in its training corpus whereby a weight would have been allotted to it'* This paper upholds the fact that a sort of semantic web annotation such as done in the model built in this research solves this problem arms down because all the semantic load of each of the meaning of the lexical verbs would have been defined an annotated each with its peculiar identifier and the annotated semantic weight which makes it easier for machine to forge ahead.

Other works in this direction includes Rayson, Uchechukwu and Hepple (2020). The authors used dataset from OPUS project and JW.ORG as a source for training data to build, maintain and publicly share a standard benchmark dataset for Igbo-English machine translation research. They concluded that an in-depth study of this work will contribute in the nearest future to the development of suitable translating devices in Igbo-English machine translation. Though this is an introductory work but the fact remains that if semantic web annotation is developed for the language concerned, this will support a good MT when it is later developed for Igbo language. Chinenyeze, Bennett and Taylor (2019) developed English to Igbo language translation system. The research used language model, translation model and decoder done in Microsoft Hub to generate the training of parallel document, and implemented a translation system compatible with Android studio environment which can be accessed through Android application in smartphones. As a backbone, Finite State Automata were used to determine the English and Igbo language tokens as valid or invalid transitions. The teething ambiguity challenge as it is faced with those other NLP systems will be resolved through ontological annotation such as it is implemented in this paper.

Eludiora, Agbeyangi and Fatusi (2005) is another work that should be reviewed to show the high importance of semantic web technology in NLP including Machine translation. This work adopted a name: 'tone change verbs' for Yorùbá lexical verbs that possesses a low tone in its orthography form but changes to mid tone if it is preceded by a noun phrase in a sentence. For example (buy) 'rà' →(Ade buys fish) 'Adé **ra** eja', (sell) 'tà' → (Ade sells fish) 'Adé **ta**ẹja'. (Note that the tone on 'rà' and 'tà' changes to mid. (Bamgbose 1990, Yusuf 2006). According to this work, 'these tone changing verbs do pose some challenges in English to Yorùbá machine translation'. In order to solve these challenges, twenty of these verbs were captured and re-write rules were formulated for the two languages The system accepts English sentence, then discover the pattern for the sentence and translate it in a word- for -word order, and if it discovers in the word for word patterns any of the stored tone changing verbs, the system will supply the alternative mid tone verb. The paper provided finite state automata for the English and Yorùbá sentence structures provided the software designs in UML and used python toolkits for parsing and implementation. The result of the comparison between Google Translate and IFEMT2 shows that IFEMT2 performs better in the translation of those sentences. However, as good as this attempt is, especially in fostering the development of African languages, the rule-based approach adopted has so many peculiar challenges too, especially in capturing the semantic nuances found in African languages but is better to have a system with its challenge than not having any at all. Interestingly the tone-changing challenge as experienced in this system can be captured and formally specified in both human and machine-readable format as a means to solve the problem as premised and demonstrated in this paper.

Abiola, Adetumbi, Fasiku and Olatunji (2010) falls in line with the NLP works for African languages. This work formulated twenty rules for the Yorùbá noun phrase translation which were specified using context-free grammar. It modelled and recognized the grammar of the language using finite state automata whose operations was based on the first set techniques. The first set technique allows the parser to choose the production to apply based on the first input word for an input phase with the aid of PHP hypertext pre-processor (PHP) and My Structured Query Language (MySQL). The work developed a bilingual lexicon made up of words in English and their corresponding translation in the Yorùbá language. The beauty of our work is still established that semantic web technology is a hand in glove to solve ambiguity problems that may be encountered in NLP work like this too.

Other works in this direction too were Agbeyangi, Eludiora and Adenekan (2005), Akinwale, Obe, Adetumbi and Adesuyi (2015) and a host of others. It is interesting to note that these scholars attempted these efforts as a result of the dire need to start work on modelling African languages to meet with the future challenge, Therefore, we have also chosen the semantic web alternatives as a means to an end in providing backbone to NLP activities. Other works in this direction is Aina (2019a, 2019b), Aina and Taiwo (2019 and 2021). All these works are demonstration of the use of semantic web annotation to address different semantic challenge for web searching agents.

Data for this Research

Lexical verbs in the language are very enormous and many of them are root morphemes that affixes could be attached to them to form compounded words. They can also be mono syllabic as well as being bi-syllabic, Taiwo (2006). The space and scope will not allow us to deal with all possible lexical verbs in Yorùbá language. For these reasons, Sixty (60) of the different types of verbs are randomly but purposively generated. The data were generated in alphabetical order. We followed the automatic consonant arrangement in the alphabet inventory of the language, i.e. (b, d, f…), and attached the vowel in a e ẹ i o ọ u order.Awobuluyi (1978). It is arranged in the progression of low, mid, and high (dò, re, mí) tone patterns. Each verb generated was checked based on Yorùbá semantic plausibility. Those that have no meaning were labeled with an asterisk (*) and they were seventeen (17) in number, those with only one meaning were labeled (-) and they were also eighteen (18), because our interest is on those that are ambiguous that is, those with more than one meaning, whichis the focus of this paper, were labeled (+), and they were twenty five (25) in number.They are presented in the following section:

+bà	i. *hit*	òkúta **ba** ọmọ náà (a stone hit the child)
	ii. *laid on top*	ẹyẹ **bà** léorùlé(the bird landed on the roof)
+bá	i.*meet*	ó**bá** mi ní ilé (he/she met me at home)
	ii. *overtake*	ó**bá** mi ní ọ̀nà (he/she met me on the road)
+ba	i.*laid ambush*	ọ̀darànnáà **ba**sí ẹ̀gbè igbó (the criminal laid ambush)
	ii. *weave*	mo ti **ba** irun kíkẹ́ tán(I have finished weaving kíkẹ́'s hair)
iii.	*nurse seed*	moti **ba** kúrú tomato tan (I have finished nursing the tomato seed)
-bẹ̀	i. plead	Olú **bẹ** Ade

+bẹ	i. peel	Mo **bẹ** iṣu lánàá (I peeled yam yesterday)
	ii. scratch (dialectal)	Ara ń **bẹ** mí (Body is scratching me)
	iii. excessively smart	Ọmọ náà **bẹ** (The child is excessively smart)
+bẹ́	i. burst	Ọ̀pá eponáà **bẹ́**(the pipeline bursted)
	ii. jump	Ó **bẹ́** lé òrùlégìjà (He jumped on top of the roof)
	iii. too forward	O máń **bẹ́**jù (You are always too forward)
+bì	i. vomit	Ọmọ náà **bì**(the child vomited)
	ii. reflective cloth	Àwọnaṣọ náa ń **bì**(the clothes are reflective)
-bi	i. ask	Ṣadé**bi**Ọlá ní ìbéère (Ṣadé asked Ọlá a question)
-bí	i. gave birth	Obìrìnnáà **bí**ọmọ okùnrin (The woman gave birth to a male child)
+bò	i. cover	Mo **bo**araàmi (I covered my body)
	ii. keep (secret)	**Bò** mí ní àsírí (keep my secret)
	iii. being shady	Ewe oriiginaa**bò**dáadáa (The leaves of the tree is shady)
-*bo	No semantic meaning	
-bó /bẹ́	i. peel	**bó** èpo ara igi náà kí o tó gée
+bọ̀	i. arrive	Mo tí ń **bọ̀** (I will soon arrive)
	ii. boil	Jọ̀wọ́ **bọ** ẹran náà dáadáa (Please boil the meet very well)
-bọ	i. worship	Wọ́n **bọ** ọṣun (They worshipped ọ̀sun)
-bọ́	i. fall	O **bọ́** lati òkè (It fell from up)
+bù	i. take/cut	Jọ́wọ́ **bu** ẹran fún mi (pease cut meat for me)
	ii. part of a thing	ebi**bù** iṣu (Yam pieces)
	iii. tear	Ọmọ naa subú, ó sì **bù** lara
yámayàma		(The child fell aǹ got seriously wounded)
-bu	i. dirty	ẹsẹ̀ Olú **bù** nígbà tí ó dé (Olu’s feet were very dirty at his arrival)
-bú	i. abuses	Ayọ̀ **bú** ọ̀rẹ́ rẹ̀ (Ayọ̀ abuses his friend)
+dà	i. lead sheep/livestock	Okùnrinnáà **da**erankọjá(The Herder led his flock)
	ii. ink spread/pour	Yíǹkì gègé náà ań **dà**
*da	No semantic meaning	
-dá	i. create	Ta ni ó dá ọ? (Who created you?)

	ii. break	Ìjì atẹ́gùn dá iginá à (The windstorm broke the tree)
-dè	i. tie	Wọ́ṇ fi okùn de igi (They tied wood with rope)
*de	No semantic meaning	
+dé	i. arrive	Wọ́n ti**dé**(They have arrived)
	ii. cover	Mo **dé**agolo (I covered the can)
+dẹ̀	i. sluggish	Ọmọ náà **dẹ̀**gan-an (ọ̀+dẹ̀) (The child is very sluggish)
	ii. soft	kí Ọ́lọ́run **dẹ**ilẹ̀ fún ẹnirere (May the lord soften ground for a good man: metaphoric)
	iii. release	Jọ̀wọ̀ **dẹ**okùnọrùnyìí náà (Please release this tie for now)
+dẹ	i. hunt	Wọ́n ń **dẹ**ìgbẹ́ lọ́wọ́ (They are hunting at the moment)
	ii. keep an eye	Gbogbo ọ̀nạ̀ nimo fí n **dẹ** ọ́ (I am keeping eyes on you in all ways)
*dẹ́	No semantic meaning	
+dì	i. tie	Fi okùn**di**igi (Tie the wood with rope)
dì	ii. change into	Ó ti**di**olówó (He has Changed to become a rich man)
-dí	i. block	Jọ̀wọ́ má **di** ọ̀nà yẹn (Please don't block the road)
*dò	symbol for low tone (̀)	
*do	No semantic meaning	
+dó	i. settle	Íbití a **dó**sí (The place we settled)
	ii. sex	Ó sì **dó**ọmọ náà (He had sexual intercourse with the child)
*dọ̀	No semantic meaning	
*dọ	No semantic meaning	
*dọ́	No semantic meaning	
+dù	i. struggle for	Wọ́n **du** Oyè (They struggled for chieftaincy title)
	ii. denied of something	Wọ́n fi ẹ̀tọ́ mi **dù** mí (They denied me of my right)
-du	i. run (dialectal <Àwórì>)	Mo **du** lọ (I ran away)
-dú	i. black	Ó **dú**bíikóró ịṣin (He is black like ịṣin seed)
+fà	i. pull	Ó **fa**ilẹ̀kùn (He pulled the door)
	ii. cause	Kí ló **fa**ìjà (What caused fight)
	iii. move slowly	Mo **fà** bí ìgbín (I moved slowly like snail)
*fa	No semantic meaning	

+fá i. pack **Fa**́ìgbẹ́ ọmọ náà (Pack the child's faeces)
ii. cut out from Àìsàn jẹjẹrẹ jẹ́ kí woṇ́ **fá**oyànobìrinnáà (They cut out the woman's breast because of cancer)

*fè No semantic meaning
*fe No semantic meaning
*fé No semantic meaning
+fẹ̀ i. widenfẹ okùnyẹndíẹ̀
ii. to distort the face ó fẹ ojú mọ́ ọmọ́ náà (He distorted his face towards the child)
*fẹ No semantic meaning

+fẹ́ i. want Mo **fẹ́** má a lọ(I want to start going)
ii. marry Mo **fẹ́** fẹ́ ìyàwó (I want to marry a wife)
iii. desire Mo fẹ́ ẹja tútù (I desire fresh fish
iv. blow Mo **fẹ́** eérú sí lára (I blew ashes to his body)

+fì i. swing O **fì** sọ̀tùn-ún, sósì(It swung leftward and rightward)
ii. Àìsàn náà **fi** mí díẹ̀(The illness battle with me a little)

-fi i. prep
*fí No semantic meaning
+fò i. jump Mo **fo** odò kojá (jumped over the stream)
ii. dry/disappear Eewo náà **fò** kúrò(the boil disappeared)

*fo No semantic meaning
-fó i. float Aṣọ náà **fó** sórí omi(The cloth floated on top of water)
+fọ̀ ii.wash Mo **fọ** aṣọ (I washed cloth)

*fọ No semantic meaning
+fọ́ i. surpriseMo **fọ́**nigbatiolùkọ́ wolé (I got surprised when teacher came in)
ii. like (colloquial) Mo **fọ́**fúnọmọ yẹ̀n (I like the girl)

+gà i. open **Ga** agbòrùn yẹn (open the umbrella)
ii. deceive Ma **gà** mí mọ́ (Don't deceive me)
iii. hang **Ga** ẹran yẹn sí orí iná (Hang the meat on fire)

-ga i. tall Igi náà <u>ga</u> fíofío
-gá ii.stop (can't move forward) Mo ti <u>gá</u> si bí o (I have stopped here)

8. Data Analysis

The data gathered for this research work is analyzed within the framework of Lexical Semantics (LS). This is done with the assumption that when a semantic load of individual verbs is defined and specified, there can be no two verbs that are exactly alike. The differentiating components formally specified become the basis whereby NLP and web searching agents accessed for utilization and convenience operation. Lexical semantics is basically designed to represent the meaning of predicative elements and the semantics of propositions. It should be considered both as a semantic model providing a representational framework and a language of primitives on the one hand and as a methodology on the other hand, allowing for the introduction of new primitives to the language, whenever justified. Another important characteristic of LS is the close relations it has with syntax, allowing the implementation of a comprehensive system of semantic composition rules that could be operated within conceptual categories, semantic fields and primitives. Other elements are conceptual variables, and semantic features similar to selection restrictions. One of the basic goals of lexical semantic theory is to provide a specification of word meanings in terms of semantic components and combinatory relations among them. Different works in lexical semantics converge now on the hypothesis that the meaning of every lexeme can be analyzed in terms of a set of more general meaning components, some or all of which are common to groups of lexemes in a language or cross-linguistically. In other words, meaning components can be identified which may or may not be lexicalized in particular languages. Talmy(1985).The individuation of the meaning components characterising classes of words in a language and of the possible combinations of such components within word roots leads to the identification of lexicalization patterns varying across languages. Moreover there is a strong correlation between each combination of meaning components and the syntactic constructions allowed by the words displaying them. (Jacobs and Rau 1990). Scientist in the attempt to account for semantic activities on concepts and word categories proposed Generative lexicon (GL). The data generated for this research is described within the GL framework.

The Generative lexicon

Generative Lexicon is a theory of semantics. It deals with the distributed nature of compositionality in natural language. GL is concerned with spreading the semantic load of lexical item across all constituents of the utterance. Generative Lexicon introduces a knowledge representation framework which offers a rich and expressive vocabulary for lexical information. Above all, 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:

(1) The lexical typing structure which provides an explicit type for a word positioned within a type system for the language. The explicit typing structure adopted in this paper follows Chafe (2007) and it identified the following typing structure: Agent, Patient, Experiencer, Beneficiary, Complement, Locative and Instrument.(2.) The argument structure specifies the number and nature of the arguments to a predicate. The arguments modelled are drawn according to Chafe (2007). (3) The event structure defines the event type of the expression and any subeventual structure it may have with subevents. Cook (2008) identified basic event structures as experiencer, benefactive and locative verb matrix with each having state, process, action and action process value.(4)Finally, the Qualia structure (QS) 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, a structural differentiation of the predicative force behind a lexical item. The qualia structure is defined as the modes of explanation associated with a word or phrase in natural language, and are defined with the formal, constitutive, telic and agentive properties.Pustejovsky(1991):

The formal property distinguishes the meaning of a word within a larger domain. Formal feature encodes distinctive features of entities (e.g. transitivity and supra segmental features like tone as adapted to African languages) and their relation with other elements within the same domain. Equally, the constitutive structure defines the relation between a verb and its lexical entailment, encoding information about an action, state, process that will have to necessarily occur before the actual verb.Telic structure states the purpose or function of the verb if there is one. In other words, telic: embraces data about the purpose and the function of action and events and the agentive structure includes the factors involved in the lexicon's origins or "coming into being". Precisely, agentive encodes the factors involved in the origin or creation of the entity or the event (creator, artefacts, natural class, causal chain, etc.)

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

Having discussed the analysis frame for coding lexical verbs, it must be noted that the GL is justified on the basis that it is a feature based approach to analysis of lexical item. The proliferation of lexical databases in multiple formats has given birth to the growing concern over the reusability of lexical resources. A model 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. It also opens up the possibility of compatibility with computational lexicons. Because of the widespread use in file representation of linguistic information, the applicability of feature structures to lexical databases seems natural. More efforts in the recent days are geared towards implementing so much of these lexical concepts for African languages for future machine use. Some few samples of the generated data are now presented in the GL frame forthwith.

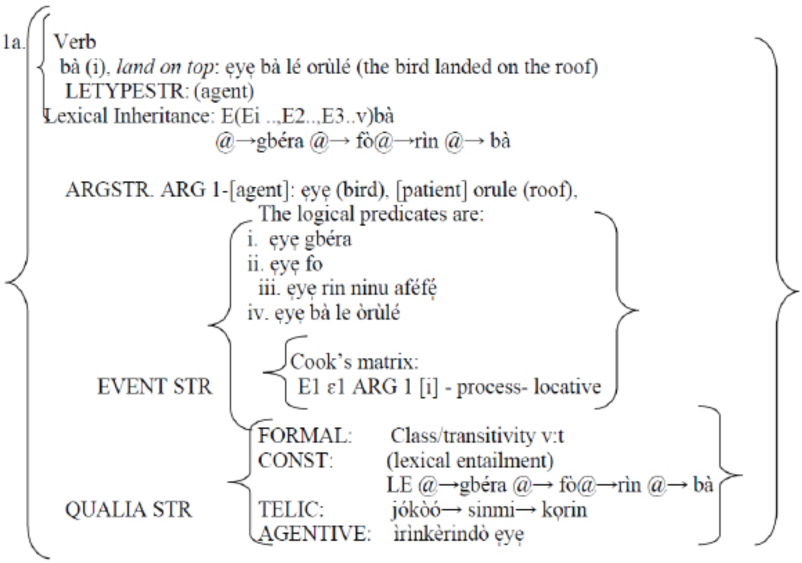

Chart 1a: Generative Lexicon Frame for Analysis of the verb 'bà' (*land on top*)

1b. α Verb
bà (ii), *hit* : Ọ̀kúta ba omo náà (A stone hit the child)
LETYPESTR: (Instrument)
Lexical Inheritance: E{Ei ..,E2..,E3..v}bà
@→ mu @→sọ @→gbéra @→ rin @→ bà

ARGSTR. ARG 1-[agent]: enì kan (someone), [patient] omo (a child),
The logical predicates are:
i Ẹnì kan mú òkúta
ii. Ẹnì náà sọ òkúta
iii. òkúta gbéra
iv. òkúta rìn nínú afẹ́fẹ́
v. òkúta ba ọmọ náà

EVENTSTR Cook's matrix
E1 ε1 ARG 1 [i] - process- benefactive
FORMAL what ϰ is Class/transitivity v:t, Case frame
[S_Tc], where S= Subject, Tc= tone change
(low to mid) bà→ba
QUALIA STR CONST what ϰ is made of (lexical entailment)
@→mu@→sọ@→gbéra @→rin@→ bà
TELIC function of ϰ ba/ kan/ kolu/fa ipalara/irora
AGENTIVE ìgbòkègbodò ẹnì kan

Chart 1b: Generative Lexicon Frame for Analysis of the verb 'bà' (*hit*)

2a α Verb
ba (i), *laid ambush* : Ọ̀daràn náà **ba** sí ẹ̀gbẹ́ igbó (The criminal laid ambush)
LETYPESTR: (locative)
Lexical Inheritance: E{Ei ..,E2..,E3..v}ba
gbéra @→ rìn @→ ba
ARGSTR. ARG 1-[agent]: Ọ̀daràn (a criminal), [patient] ẹ̀gbẹ́ igbó (bush side),
The logical predicates are:
i Ọ̀daràn kan gbéra
ii. Ọ̀daràn náà rin lọ sí ibì kan
iii. Ọ̀daràn náà ba

EVENTSTR Cook's matrix
E1 ε1 ARG 1 [i] Action-process experiential verbs
FORMAL what ϰ is Class/transitivity v:i mid tone
CONST what ϰ is made of (lexical entailment)
gbéra @→rin@→ ba
QUALIA STR TELIC function of ϰ ṣe ìjàmbá
AGENTIVE ìgbòkègbodò ọ̀daràn

Chart 2a: Generative Lexicon Frame for Analysis of the verb 'ba' (laid ambush)

2b.
α Verb
ba (ii), *weave*: Mo ti **ba** irun kíkẹ́ tan (I have finished weaving kíkẹ́'s hair)
LETYPESTR: (agent)
Lexical Inheritance: E{Ei ..,E2..,E3..v}ba
tú irun @→ ya irun @→ ba irun

ARGSTR. ARG 1-[agent]: ẹnì kan (someone), [patient] irun (hair)
The logical predicates are
i. ẹnì kan tú irun
ii. ẹnì kan ya irun
iii. ẹnì kan ba irun

Cook's matrix
EVENTSTR E1 ε1 ARG 1 [i] Action-process verb
FORMAL what ϰ is Class/transitivity v:t mid tone
CONST what ϰ is made of (lexical entailment)
tú irun @→ ya irun @→ ba irun
QUALIA STR TELIC function of ϰ ṣe aájò ẹwà
AGENTIVE àṣà, oge ṣíṣe

Chart 2b: Generative Lexicon Frame for Analysis of the verb 'bà' (*weave*)

2c.
α Verb
ba (iii), *nurse seed* : Mo ti **ba** kúrú tomato tan (I have finished nursing the tomato seed)
LETYPESTR: (agent)

Lexical Inheritance: E{Ei ..,E2..,E3..v}ba
ṣá ilẹ̀ @→ tú ilẹ̀ @→ ba kúrú

ARGSTR. ARG 1-[agent]: ẹnì kan (someone), [patient] ilẹ̀ (ground)
The logical predicates are:
i. ẹnì kan ṣá ilẹ̀
ii. ẹni náà tú ilẹ̀
iii. ẹni náà ba kúrú tòmátì

Cook's matrix
EVENTSTR E1 ε1 ARG 1 [i] Action-process verb
FORMAL what ϰ is Class/transitivity v:t mid tone

CONST what ϰ is made of (lexical entailment)
ṣá ilẹ̀ @→ tú ilẹ̀ @→ ba kúrú
QUALIA STR TELIC function of ϰ mú èso jáde
AGENTIVE iṣẹ́ àgbẹ̀

Chart 2c: Generative Lexicon Frame for Analysis of the verb 'bà' (*nurse seed*)

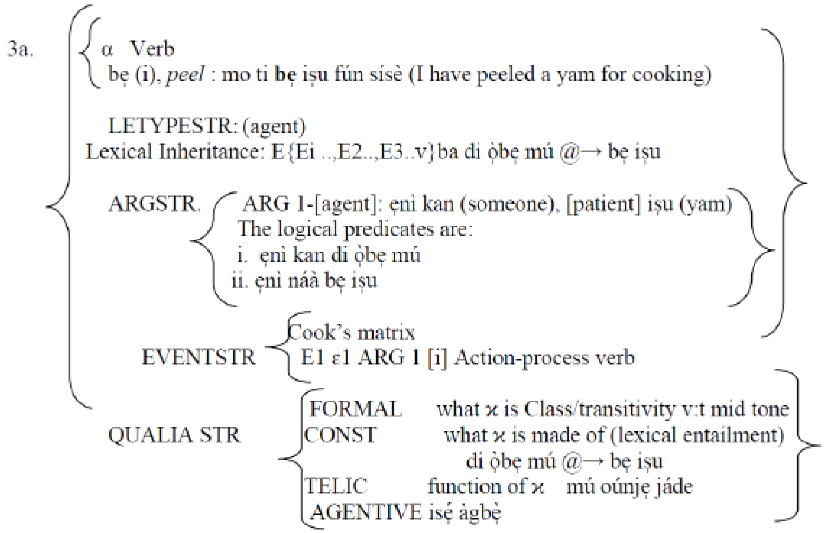

Chart 3a: Generative Lexicon Frame for Analysis of the verb 'bẹ' (*peel*)

Implementation of Yorùbá Lexical Verb

The procedures we follow are outlined below:

1. I Started protégé A welcome to protégé banner is dialled
2. After the Welcome to Protege dialog box appears, the next step is to press the `Create New OWL Ontology' button.
3. A `Create Ontology URI Wizard will appear'. EveryOntology is named using a Unique Resource Identifier (URI). I replaced the default URI with the name of my model http://www.yorlexverb@yswt.com/ontologies/yorvoa.owl and press `Next'.
4. I save the Ontology to a file on my PC, browsing to the hard disk to save the ontology to a new file, named `yorvoa.owl'. Once a file is chosen then one can press `Finish'.

The implementation of the Yorùbá mono syllabic lexical verb ontology proceeds. The step by step implementation procedure may be too enormous for the scope of this paper. However, a first glance of the whole implemented model is displayed in Figure 1. Any existing phenomena or concept is believed to be a product of 'Thing'. The Yoruba lexical verb in question is a sub-class of 'Thing' as shown in the figure, while each of the lexical verbs along with each component and properties is a sub-class of the concept 'Yoruba lexical verb'.Figure 2 shows the components of a section in the implemented ontology. The frame presented in 1-3 under section 8 is modelled in Figure 3.

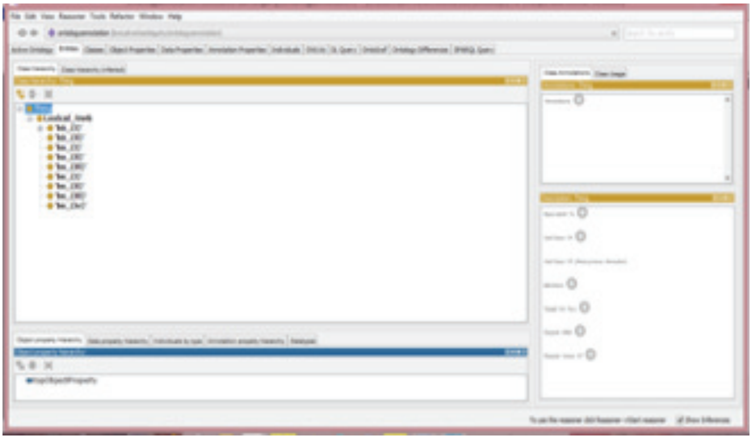

Figure 1: An overview of the Implemented Model

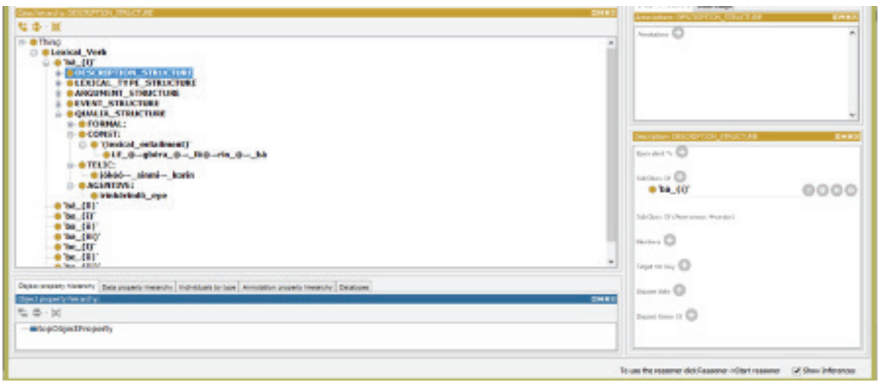

Figure 2: The View of a Section in YORLEVE Implementation

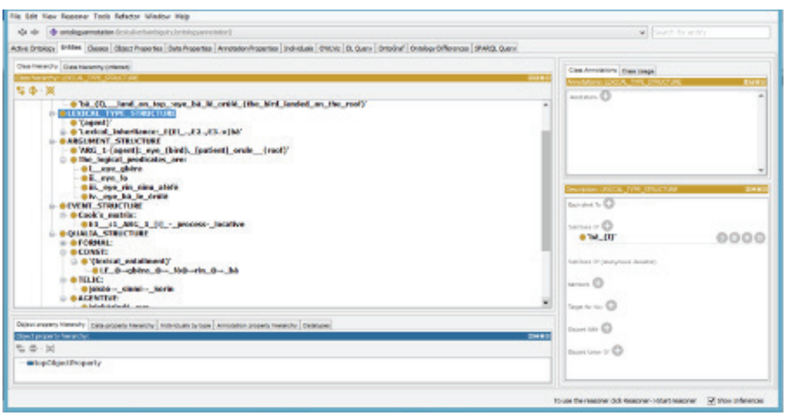

Figure 3: The implementation of YORLEVE frame in section 8

OWL Properties Implemented for YORLEXVE

OWL Properties represent relationships between classes. There are two main types of properties: Object properties and Datatype properties. Object properties are relationships between two individuals. Object properties link an individual to an individual. OWL also has a third type of property called Annotation properties. Annotation properties can be used to add information (metadata that is data about data) to classes, individuals and object/datatype properties.

Properties may be created using the `Object Properties' tab inProtégé. Figure 4 shows the buttons located in the top left hand corner of the `Object Properties' tab that are used for creatingOWL properties. As can be seen from this illustration, there are buttons for creating Datatype properties, Object properties and Annotation properties. All properties created in this implementation

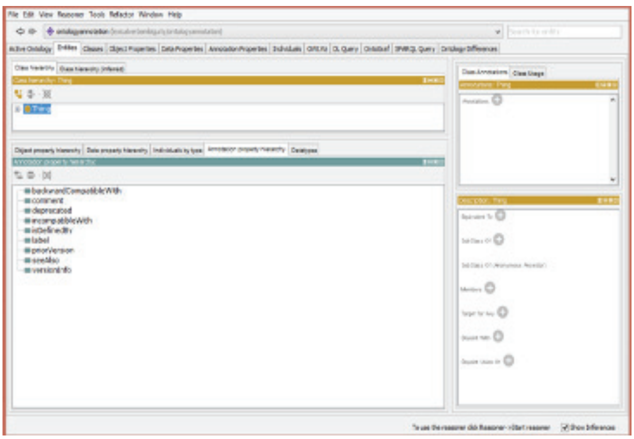

Figure 4: An overview of Annotated properties

are object properties.

Projection of NLP Systems incorporating YORLEXVE

As explained previously, Ontology leads to the sharing of knowledge between systems and people. It plays a key role in the Semantic Web supporting information exchange across distributed environments. The Semantic Web represents data in such a way that it can be machine-processed. The following applications are some NLP applications that will need to utilise YORLEXVE:

1. Artificial Intelligent systems

2. Decision Support Systems
3. Semantic Web Service Discovery
4. Multi agent Systems
5. Search engines
6. E- Commerce

However, the subject of how any of the above listed applications can utilise YORLEXVE is a subject of further and independent research. Depending on the focus of the implementation of the different applications, system development activities had to be generated to meet with the inter-operability needs. Furthermore, some NLP applications in Yorùbá language like Odoje (2010, 2017), Eludiora (2012), Aina (2018) may utilise YORLEXVE in their source codes. Importantly, the requirements for ontology utility as espoused in Berners-Lee, Hendler and Lassila (2001) must be followed. These are highlighted briefly as follows:

(1) The ontology must be available on the web with an open licence.
(2) It must be machine-readable structured data (e.g., CycL instead of image scan of a table).
(3) It must have non-proprietary format (e.g., OBO instead of CycL).
(4) It must comply with the open standards from the W3C (RDF Schema and OWL).
(5) It must be reusable in other people's ontologies.

The guidelines and procedures in Garijo (2013) must be strictly followed to load YORLEXVE into the utilising applications. Let us briefly demonstrate this hypothetically, that Odoje (2010) will need to utilise YORLEXVE in its lexicon module which comprises lexical items representation of the Yorùbá parts of speech, let us say lexical verbs in more specific sense. The YORLEXVE annotations URL accessibility must be approved from the admin of its hosting site located in http://purl.org .[1]

1 The URI for our semantic web annotation models including the one developed in this paper is targeted at: http://purl.org/net/yorno and yorlexverb.

Having explained this, the next task for Odoje (2010) to utilise YORLEXVE is to dereference ontological annotations following the outlined tasks in vocabulary dereferencing according to Berners-Lee, Hendler and Lassila (2001). Some of the steps include:

1. Set the purl redirection for Semantic Web resources (add a 303 redirection instead of 302) and add the target URL[2] of YORLEXVE into the utilising application loop. The particular loop this time is Odoje (2010). The redirection is looped into the lexicon module of its E-Y MT (English to Yorùbá Machine Translation). One can only dereference a resource, only if one is in control of the server where the resources are going to be delivered.
2 Adhere to W3C documents standard by using the link 'http://www.w3.org/TR/swbp-vocab-pub/#recipe3',] which is a simple redirection for vocabularies with a hash namespace and create an htaccess file for the link. For YORLEXVE, the index.html file has the documentation of the ontology, while yorlexverb-ontology1.1.owl contains the rdf/xml encoding. All the files are located in a folder called yorlexverb.content. In order to avoid an infinite loop when dealing with the redirections of the vocabulary it must be tuned in this 'middle way extract' of the set of instruction and employed:

```
# Turn off MultiViews
Options -MultiViews
# Directive to ensure *.rdf files served as appropriate content type,
# if not present in main apache config
AddType application/rdf+xml .rdf
```

However, it has been said that the focus of each application implementation will determine how it will be utilised but the above procedure is a demonstration of YORLEXVE utility, more so that approaches towards system development are many, depending on the tools and functions of the system.

Conclusion

This paper has introduced us to semantic web technology as a formal method of expressing relations in a specific domain. We have generated the Yorùbá lexical verb which is restricted to mono syllabic verb only, and in compliance with automatic alphabetic arrangement extractable in MYD. We have presented the attributing properties for its sub classification, using cooks

2 To link the URI, place the cursor on the ontology IRI of the ontology header, the code is displayed automatically and can be retrieved for linking the loop.

frame to compose the lexical type structure, argument type structure and the event type structure of GL. The elicited knowledge has been defined for its entities, classes and object properties, axiomatic expression using protégé 4.5 which is expressed in OWL as the implementation tool. The justification of the procedure for utilising our model into other NLP system is demonstrated, using the proposed lexicon module of Odoje (2010). This paper demonstrated and recommends that a huge of works in semantic web technology should be encouraged as it is an aid to reach an end in AIS, NLP and other web utility agents and activities. The formal specification as done in the paper becomes useful in sharing, interoperating and reuse in knowledge-based systems.

REFERENCES

Abiola, O.B., Adetunmbi, A. O., Fasiku, A.I. and Olatunji K. O..A Web-Based English to Yoruba Noun Phrase Machine Translation System. *International Journal of English and Literature.*5.3: 71-78. Retrieved July 12, 2020 from www.academicjournals.org/EUL (2014).

Adegbola, T. Globalization: Colonizing the Space Flows? *Globalization and the Future of African Languages*. F.O.Egbokhare and C. Kolawole. Ibadan Cultural Studies Group, Ibadan. 2 – 11 (2006)

--------------- 2008. Indigenizing Human Language Technology for National Development. A Lecture delivered as First ARCIS Distinguished Guest Lecture, U.S.A.

--------------- 2009. Building Capacities in Human Language Technology for African Languages. EACL 2009 Workshop on Language Technologies for African Languages – AFLaT. 53–58. Athens: Association for Computational Linguistics.

Agbeyangi A. O.,Eludiora S.I. and Adenekan O. A. English to Yoruba Machine Translation System using Rule Based Approach.*Journal of Multidisciplinary Engineering Science and Technology.(JMEST).* 2.8: 2275-2280. Retrieved July 12 2020 from www.jmest.org (2005.)

Aina, A. A.. A Hybrid Yoruba Noun Ontology.*KIU Journal of Humanities*. 4.4: 229-246.College of Humanities and Social Sciences, Kampala International University, Uganda (2019a).

………….. 2019b.Development and Preservation of Yoruba Cultural Identity: An Ontology–Based Approach.*Ọ̀páǹbàtà- Jọ́nà ìmọ̀ Áfíríka: LASU Journal of African Studies*, 7: 19-29. Department of African Languages, Literatures and Communication Arts, Lagos State University, Ọ̀jó, Lagos.

Aina, A.A. and Taiwo P. O. Yoruba Noun Ontology from Functional Perspectives.*Journal of the Linguistics Association of Nigeria*, 22.1: 189-21

Linguistic Association of Nigeria.www.jolan.org.ng/access/downloader.php. PDF (2019.)

…………...2021. Ontological Annotation for Natural Language Development: A Yorùbá Noun Preliminary Model.*Journal of West AfricanLanguages*, 48.1 www.journalofwestafricanlanguages.org. PDF

Akinwale O.I.,Adetunmbi A. O., Obe O. O. and Adesuyi T.. Web- Based English to Yoruba Machine Translation.*International Journal of Language and Linguistics*.3.3:154-159 (2005). Retrieved on July 20 2021 from http://www.sciencepublishinggroup.com/j/ijll

Awobuluyi, O*Essentials of Yorùbá Grammar.*University Press Ibadan.

--------------- 2006. *Linguistics and Corporate Social Responsibility*. A Faculty Lecture of University of Ibadan. Ibadan. DB Martoy Book.

Bamgbose, A. *FonolojiatiGirama Yorùbá*: Ibadan: UPL(1990.)

Berners-Lee, T., Hendler, J. and Lassila, O.. The Semantic Web: Scientific American. May, 2001. Retrieved July 12 2014 from http://www.sciam.com/ (2001)

Chinenyeze, C. E., Bennet, E. O. and Taylor, O. E. A Natural Language Processing System for English to Igbo Language Translation in Android. *International Journal of Computer Science and Mathematical Theory.* 5.1:64-75 International Institute for Academic Research and Development (2019) Retrieved May 16 2020

Eludiora S.I., Agbeyangi A. O. and Fatusin O. I.Development of English to Yoruba Machine Translation System for Yoruba Verbs' Tone Changing. *International Journal of Computer Application.*129.10:12-17 (2005). Retrieved July 12 2020 from www.ijconline.org.

Emuoyibofarhe, N., Adebayo, S., Lala, G.O., and Aremu, R.O. A Yorùbá Cultural Tradition Repository Knowledge Based System. *International Journal of Emerging Trends in Science and Technology* 2.7: 2830-2841 (2015)

EAGLES,. EAGLES: Recommendations for the Morphosyntactic Annotation of Corpora (1996a) EUROPEAN PROJECT DELIVERABLE: EAGLES Document EAG--TCWG—MAC/R, EAGLES Consortium, 1996. Retrieved online Dec. 28, 2015 from http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html.

EAGLES, EAGLES: Recommendations for the Syntactic Annotation of Corpora. EUROPEAN PROJECT DELIVERABLE: EAGLES Document EAG–TCWG–SASG/1.8, EAGLES Consortium, 1996. Retrieved online Dec. 28, 2015 from http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html (1996b.)

EAGLES. EAGLES: Preliminary Recommendations on Semantic Encoding, Final Report. EUROPEAN PROJECT DELIVERABLE: EAGLES LE3-4244, The EAGLES Lexicon Interest OntoTag: A Linguistic and

Ontological Annotation Model Suitable for the Semantic Web 350 Group, 1999. Retrieved online Feb. 15, 2016 from http://www.ilc.cnr.it/EAGLES/EAGLESLE.PDF. (1999)

Eludiora, S.I..Development of English to Yorùbá machine translation.PhD Thesis Dept. of Computer Science and Engineering.O.A.U. Ife.xi+206pp (2012).

Ezeanyeji, P.C., Ebinyasi, E.D. and Mgbeafulike, I. J. Development of an Enhanced Bi-lingual English- to- Igbo Translator. *COOU Journal of Physical Science.* 2.8:140-144 (2019.). Retrieved July 2020.

Garijo, D., 2013. How to (Properly) Publish a Vocabulary or Ontology in the Web.Ontology Engineering Group, Universidad Politécnica de Madrid.Retrieved online September 14, 2017 fromhttps://figshare.com/articles/881824.

Hassan, H. 2009. A Lexical Syntax for Statistical Machine Translation.PhD Thesis.School of Computing, Dublin CityUniversity.xvi+218pp.http://doras.dcu.ie/2320/1/Hany -Final-Thesis.pdf. Retrieved on 04/05/2013.

Hassan, J.A., Odejobi, O.A., Ogunfolakan, B.A. and Adejuwon, A, 2013.Ontology Engineering in Yorùbá cultural Heritage Domain.*African Journal of Comp. & ICT*s IEEE. 6.5:181-198. Retrieved online Oct. 20, 2015 from http/www.ajocict.

Jacobs, P.S. and Rau, L.F., 1990.Scisor: Extracting information from on-line news. *Communications of the ACM*, 33.11:88-97.

Odejobi, O.A. 2015. Self-knowledge and intellectualism: A Perspective from Yorùbá Cosmology. *Nationalism and Economic Justice in Nigeria.*Salami Y.K., Famakinwa J.O. and FasikuG.eds. O.A.U. Press. 747-784.

Odoje, C. O. 2010. English-Yorùbá Rule Based Syntax in Machine Translation. M. A. Project. Department of Linguistics and African Languages.The University of Ibadan, Ibadan.

---------------2013. Language Inequality: Machine Translation as the Bridging Bridge for African Languages. *Ago-Iwoye Journal of Languages and Literatures.* 4:22-49

---------------2017. Statistical machine translation. PhD. Thesis.Dept. of Linguistics and Africa Languages.University of Ibadan.xvi+218pp.

Pareja – Lora, A. OntoTag: A Linguistic and Ontological Annotation Model Suitable for the Semantic Web. A Ph.DThesis.Facultad De Informática Universidad Politécnica De Madrid.viii+365 (2012.). Retrieved online on Aug. 28, 2017from https:// www.researchgate.net.publication.

Pustejovsky, J. The Generative Lexicon. *Computational Linguistics*, 17. 4 (1991)

Talmy, L. Lexicalization Patterns: Semantic Structure in Lexical Form. *Language Typology and Syntactic Description: Grammatical Categories* *and the Lexicon*. Shopen, T. (ed.). Cambridge, Cambridge University Press(1985).

Taiwo, O. *Mofoloji*. Ibadan: Layof Ventures Publishing Educational Service (2006.

Yusuff, A.*Ojúlówó Gírámà Yorùbá.*Joytal Printing Press. Lagos (2006.)