

LINGUISTICS THEORIES, NIGERIAN LANGUAGES, AND THE QUEST FOR DEVELOPMENT

By

Kolawole Adeniyi¹ and Clement Odoje²

¹ Obafemi Awolowo University, Ile-Ife, kola.adeniyi@oauife.edu.ng;

² University of Ibadan, lekecclement2@gmail.com

Abstract

This work assesses the progression of African linguistics and the application of theories to the study of African languages in light of the developmental needs of the languages. It is observed that the vast majority of African languages are yet to be adequately described for human learning and generational transmission. However, as linguists continue to adopt theories driven by language computations, descriptions needed for human learning seem to become expendable, a reality amounting to doing linguistics for the machine at the expense of humans. With attention now shifting to computation-motivated approaches, further development that can make these languages suitable for science education, security, law, medical practices and other contemporary realities in Nigerian society tends to be less rigorously pursued. Apart from Hausa, Igbo and Yoruba, there is minimal progress being made in terms of the development of Nigerian languages. Standard dictionaries and other materials designed to aid the deployment of the languages to meet contemporary human needs within the speech communities are rarely produced in the languages. This is despite the quantum of linguistic work being done. When this is viewed in light of the endangered status of the languages, it projects a critical situation. It is concluded that there is a need for synergy between linguistics for human use and the machine. This can only be championed by African linguists with the developmental needs of African languages as objective.

Keywords: African linguistics; development; human learning; linguistics theories; Nigerian languages.

Introduction

Larry Hyman once asked these two questions: “What and how have African languages contributed to phonological theory?” and “What and how have linguistic theories contributed to the understanding of African phonology?” (Hyman, 2003a). In that paper, Hyman noted that African languages have made significant contributions to phonological theory in the areas of tone, vowel harmony, nasals and nasalisation, labial and palatal prosodies, slots and moras, reduplication and syntax-phonology interface.

In this study, we are reversing Hyman’s first question thus: “In which ways and how have linguistic theories contributed to the development of African languages?” We will begin with a review of the field of linguistics in Nigeria, followed by the state of Nigerian languages. We will then examine how theories have helped in the understanding of African languages. This will be followed by a review of how theory-based understanding has translated to the development of African languages and the roles of the application of theories by African linguists in all of these.

Language and Linguistics in Nigeria

Linguistic studies on Nigerian languages started before independence in 1960 (Elugbe, 2010). The primary objective then was to describe undocumented Nigerian languages and develop orthographies for those languages. Some major achievements of the field during this period include the Rivers Readers Project (RRP) (See Williamson 1976), the Ife Six Year Primary Project (ISYPP) (Fafunwa, 1975), and the publication of a practical guide for orthography development in Nigeria (Williamson, 1984). According to Elugbe (2010), the Nigerian languages being described during this period “began to contribute to the development and fine-tuning of linguistic theory”. These contributions included the fortis/lenis distinction, serial verb constructions, unvoiced implosives, and tone.

There are over 500 languages spoken in Nigeria, most at different stages of endangerment. Yoruba, which is one of the most extensively studied Nigerian languages is endangered (Fabunmi & Salawu, 2005; Adeniyi & Olaogun, 2020), and it is only recently that concerted efforts have been made to empower it to handle the teaching of science education (Adebayo 2022, p. 320-322). Many other Nigerian languages are already extinct. According to Simons and Fennig (2018), extinct Nigerian languages include Ajawa, Basa-Gumna, Holma, Luri, Kpati, Mawa, and O’chi’chi’. According to Akpan, Urua and Ekpenyong (2018, p. 14), many of the remaining living African languages “are yet to be described and documented”.

As a consequence of the state of the languages, it is now widely believed that African languages are either “stagnating or dying” as a result of lack of development and abandonment in favour of languages of wider communication, a situation occasioned by the bad state of language policies. The challenge of development appears to have been taken up by documentary linguistics, but even then there appears to be a wide gap between documentary linguistics, theoretical linguistics and the actual development of the languages.

How much help have theories been to the development of African languages?

That African languages have made significant contributions to linguistic theories has been comprehensively proven by Hyman (2003a). For instance, Hyman (2003a) notes that “The vast majority of African languages are tonal, and perhaps most also have vowel harmony (especially the type known as “ATR harmony”). “Not surprisingly, then, African languages have figured disproportionately in theoretical treatments of these two phenomena.” Other works have also highlighted the contributions of African languages to linguistic theories. These include Goldsmith (1976) on Autosegmental phonology, Adeniyi and Elugbe (2018) on crossing association lines in Autosegmental phonology, Olaogun (2018) on the Minimalist Approach to unifying explanation of the linguistic computation of question formation in Yoruba, Igala, and Ñjò-Kóo; Akinlabi (1985) on under-specification theory; Ajongolo (2005) on trace in Syntax, and many others.

The question, however, is whether the contributions of African languages to linguistic theories are appropriately reciprocated. In other words, how has theory-based understanding translated to the development of African languages? To serve as yardsticks are the landmark achievements of the early years of linguistics in Nigeria when the focus was on language description and orthography development, namely the RRP this brought about the development of many languages, and the ISYPP. The knock-on effect of the RRP was the increased participation of local languages in the

mass media and day-to-day governance of the people, while the ISYPP demonstrated the benefits of educating children in their native languages.

With the use of advanced theories in linguistic studies, it is expected that the field would make even more significant impacts on the country's linguistic landscape. In terms of research output, there are numerous theory-based insights that one may expect to have a significant impact on African languages. One such is Adeniyi (2015) who used an Auto segmental analysis of tonal characteristics to establish typological connections between some three-tone languages within the West Benue-Congo family. Seen in the light of Elugbe (2011), this sheds light on the ties between languages and their speakers, a reality that, if harnessed properly, should be applied to fostering peace in Nigerian society. To use Yoruba as a reference, the current orthography was developed in 1974 and despite extensive theory-based studies since then; there has not been any official input or review. Another question to ask is; what are the roles of the application of theories by African linguists in all of these? (Lieberman & Lieberman, 2000; Hyman, 2003b). This question can be amplified by asking "How many contemporary linguistic theories originated in Africa? Or originated primarily in response to an African language problem? How many of those theories are based on African language data such that non-African languages are studied in the light of such findings? In recognition of the validity of these questions, Hyman noted that "While it is easy to show how languages provide grist for theoretical mills, showing how theories benefit the study of languages is at best more complicated and perhaps even controversial" (Hyman 2003a, p. 10).

This is where the problem is; there are many theories in vogue, which are incapable of describing African languages as what they are. Rather, theories often come with basic assumptions about universals which are imposed on languages that may not conform to them. In a bid not to violate such assumptions, facts natural to many languages have been downplayed or even dispensed entirely. A ready example of this is the application of tone geometry in generative phonology to tone in African languages (Akinlabi, 1985; Hyman, 1993). To achieve this for a three-tone language, for instance, one of the tones has to be assumed to be underspecified. Under-specification is itself another theory. By assuming that one of three tones is underspecified, it becomes possible for the language to then accommodate a feature-based analysis by saying such a language has a "+" or "-" value for either of the other two tones. Yoruba is a good example of this, where the mid-tone is said to be underspecified (Akinlabi, 1985). Either high or low can then be chosen, where "+H" implies a high tone and "- H" implies a low tone, and vice versa. It is interesting that the idea failed and was eventually abandoned (for detailed discussions of these, see Clements, Michaud & Patin, 2011; Hyman, 2010; Adeniyi, 2015, 58 - 62).

Another problem with the application of theories to African languages relates to data. To draw from studies on Yoruba yet again, the mechanism for vowel hiatus resolution in Yoruba is yet to be satisfactorily described. The consequence is that it is still not understood how children acquiring the language master this aspect of the language. This is because there is no linguistic explanation of the choices of the elision of the first vowel, the elision of the second vowel, the coalescence of vowels in a sequence or the preservation of both. This is because, for every identified pattern, there are almost as many convincing counterexamples.

In examples 1-3 below, it is not known what informs the choices of which vowel is elided in (a) as against (b). Note that it has not been shown that tone plays any role in the choice as the choice of which tone to elide is independently determined based on the H>L>M hierarchy (Bamgbose, 1990, 57). Hence tone can be ignored in the examples 1 – 3. Example 2 (a–c) especially presents a three-way alternation between V1 elision, V2 elision and coalescence.

1. (a) Gbá ilẹ >gbálẹ “sweep” (V2 elision)
(b) Gbà ilẹ >gbilẹ “become established” (V1 elision)
2. (a) Pa irọ >purọ “lie” (Coalescence)
(b) Pa igi>pagi “kill a tree” (V2 elision)
(c) Gba ìmọ >gbìmọ “agree” (V1 elision)
3. (a) Sun ẹkún > sunkún “cry” (V2 elision)
(b) Sun ẹkún > sẹkún “cry” (V1 elision)
(c) Sun ẹkún > sọkún “cry” (coalescence)

In situations like this, one often finds only data that conforms to the position. This is more common in Optimality Theoretic accounts of the problem. The counterexamples do not come into the picture at all or are massively downplayed. This is not a problem only in phonology; it permeates other areas of linguistics. In the end, studies abound on the application of theories to languages, but knowledge is still not advanced, just as in the vowel hiatus problem in Yoruba. But the consequence is the abundance of research outputs that are neither capable of advancing knowledge of the languages nor able to develop their linguistic endowment and identity. We believe that it is based on such situations as these that Jibril notes that “the agenda for African linguists is still set by Western linguists and most research output in African linguistics is either a reaction to their works or attempts to refute or confirm their linguistic theories” (Jibril, 2007, p.281).

A seminal work that avoided Western theories entirely to advance knowledge is Akinlabi and Liberman (2000). In that work, Akinlabi and Liberman note as follows:

Our primary goal in writing this paper has been to establish and document certain complex descriptive generalizations, which will remain valid, subject to factual correction, across future changes in approaches to phonological modelling. The documentation of such descriptive generalizations is sometimes clearer and more accessible when expressed in terms of a detailed formal reconstruction, but only in the rare and happy case that the formalism fits the data so well that the resulting account is clearer and easier to understand than the list of categories of facts that it encodes. Our considered opinion is that the topic of this paper, given available morphophonemic formalisms, is not such a case.

In this less happy (and all too common) situation, subsequent scholars must often struggle to decode a description in an out-of-date formal framework to work back to a more superficial version of the facts, typically in terms of a list of paradigm-like displays, which they can then re-formalize in a new way. Having experienced this struggle often ourselves, we have decided to accommodate our successors by providing them directly with a plainer account.”

In this situation, Akinlabi and Liberman can be said to have avoided the application of theory merely for the sake of theory. Drawing directly from this position, Hyman (2003b, p. 6) argued that

describing African languages is much more than “plugging African data into a pre-existing framework.”

According to Jibril, the agenda for African linguistics in the 21st Century should be to rescue the field (linguistics) through “research...aimed at saving, preserving, developing and promoting African languages” (p. 285). But it seems that African linguists mostly continue in their reactionary enterprise, ostensibly to be seen as current. Another insightful perspective on this is Falola (2007: 34 - 35) who notes that, “... we have contributed more to data gathering than to theory building...we borrow ideas which we use to make sense of our data but we do not make enough use of our data to create an alternative body of knowledge that others can also take from us...”

Research output, policy and development

Objectively, it cannot be merely concluded that the application of theories has had no positive input to African linguistics. Some recent linguistic theories and their analytical vigour encapsulate a new dimension of looking at African languages. They help in analysing African languages better. However, these recent linguistic analyses have neither found their way to pedagogical procedures in schools nor brought about the desired development. Olaogun (2018) is a valid example of the analytic strength bringing to bear the linguistic computation of question formation in Yoruba, Igala, and Njò-Kóo. As good as his unifying explanation could be, to date; it has not found its way into school curricula. Teachers still stick to syntactic analyses of the likes of Bamgbose (1960, 1990), and Awobuluyi (1979, 2003, 2013).

This implies that the Minimalist Approach has benefited in terms of its development and adaptability to African languages than African languages benefitting from its syntactic analysis. This is just like Ajongolo (2005) observes that trace in syntax is not just a covert element; using Yoruba constituent focusing as an example, Ajongolo shows that it can be an overt element too. This is known to the speakers and assumed to be a mere pronoun, not minding whether it is a trace of a moved constituent. This idea modifies the notion of the trace as a covert element but has not developed Yoruba in that such narration that trace is a pronoun needs to change. This is therefore a case of theoretical output with zero implementation.

Machine learning, AI, and the place of language in contemporary societies: where are African/Nigerian languages?

In contemporary society, language, in addition to being a means of communication, has a socio-economic role similar to that of money in industrialised societies. While money is used to acquire material goods, language is used to acquire knowledge and intangible goods (Osborn 2010). This assertion foregrounds the fusion of technology and language. Osborn (2010, p. 5) explains that the information revolution worldwide is increasingly multilingual, and as the presence of the new ICTs in Africa extends to larger areas beyond the capital cities, there is a growing need to accommodate the use of diverse African languages and greater potential to tap the linguistic wealth of the continent for development and education. He then suggests two aspects of digital divides: European languages cannot meet all of Africa’s needs, and African languages have much to contribute. If African languages will contribute to ICTs, Natural Language Processing (NLP) will play a crucial role. For the NLP state of the arts to yield appropriate results, there is a need for a large quantum of data which to this end compared to European languages is not sufficiently available, hence the

categorization of African languages as low-resource/resource-scarce languages. The resource scarcity makes it difficult for machine learning to learn most of the features of African languages. Looking at it from a statistical machine learning point of view, resource-scarce languages (LRLs) are languages that have small user advantages concerning technology which could be ignored by the commercial world. LRLs can be understood as less studied, resource-scarce, less computerized, less privileged, less commonly taught, or low density, among other denominations. In an attempt to ensure the machine learns language features syntactically, part of speech tagging becomes a necessity.

Part of Speech tagging (PoS) is a process of associating grammatical properties to a word/token based on its functionality in a sentence that constitutes part of the corpus. This allows the machine to learn the patterns of a language easily even if the corpus is not much. The challenge, therefore, is the need to have a universal tagset that will cater for all languages of the world. Some grammatical properties of African languages are yet to be captured in the universal tagset. Yoruba for example is left without a definite tag for the infusion of an aspect and negation in a word like *Kii* as found in example (4) below.

4. *Kii jẹun*
 Neg + Asp V
 *He does eat

Kii in (4) above is an infusion of both the negation and habitual aspect. But since negation presently has no representation in the universal tagset (<https://universaldependencies.org/u/feat/index.html>), the consequence is the wrong gloss “*He does eat” rather than “*He does not eat.” The general implication of the absence of a universal tagset for negation is therefore inappropriate tagging of such words; another consequence is that the computer is given a wrong pattern to learn from. This in the long run will make the computer generate unacceptable expressions/output. Therefore, there is a need for a concerted effort to ensure that such important linguistic properties of African languages are incorporated and appropriately tagged for Artificial intelligence appliances to manipulate African languages correctly.

Another challenge is serial verbs as well as split verbs. African languages seem to have this unique syntactic structure where verbs can be stuck in what seems to be a simple sentence. When it comes to sentence alignment, this could be a challenge for machine translation, especially pairing with European languages that do not have similar sentence structures. Word and sentence alignments become a noticeable problem for translation. Consider Igbo (5) and Yoruba (6 – 7) examples below:

5. *Anyị bjararie, taa, n̄uọ, mañurja*
 We go eat chew drank and enjoyed
 “We go eat chew drank and enjoyed”

6. *Adé ra ẹran pa tà*
 Ade buy goat kill sell
 “Ade bought the goat so as to kill it for selling purpose”

7. Ebuka gba Ọlórún gbó
 Ebuka receive God hear
 “Ebuka believes in God”

In (5) and (6) above, there are several verbs used in the expressions that seem to be simple sentences. Translating such sentences to English for instance will warrant that the sentence is broken into three or four, otherwise, it will be unnecessarily longer compared to its equivalent from the source language. Also, the verbs *eat* and *chew* in (5) suggest that in African languages the semantic implication/sense of words may vary depending on context and other factors. Example (7) however, is an instance of a split verb where each component of the verb has a different meaning in isolation as against when they are used as a split verb. Each sentence poses both structural and semantic challenges when it comes to establishing translation parallels with European languages, which makes AI compliance with African languages problematic. Alabi et al (2020) seem to have provided a solution to the issues of ambiguity, polysemy, homophony and hyponymy through context word alignment in massive and curated word embedding but this is yet to solve structural issues.

An instance of this is LINGUSTICA meant to automatically analyse world languages morphologically but could not identify Yoruba prefixes when fed with Yoruba texts. It could however identify compound words using recall precession index. This justifies the need for a homegrown morphological analyzer if Africa wants to close the digital divide experienced today.

Conclusion

This work assessed the progression of African linguistics, with a focus on Nigeria. It noted that the aim was earlier to develop African languages, which was achieved in the early days when concerted efforts were made to describe languages. It was also the period when empirical studies capable of swaying policies were done. Practices then switched to being theoretical, at the expense of development and linguists began to adopt and apply theories to the study of African languages with progressively reducing emphasis on the developmental needs of the languages.

The consequence is that the vast majority of African languages are yet to be adequately described for human learning and generational transmission and there is still no language, especially Nigerian, that is developed enough for deployment in science education, security, law, medical practices and other contemporary realities in the Nigerian society tends to be less rigorously pursued. This has amounted to African languages contributing significantly to theory development, at the expense of the languages themselves. There is therefore a need for synergy between linguistics for human use and especially for the development of neglected African languages, and that for the theory and the machine. This can only be championed by African linguists with the developmental needs of African languages as objective.

References

- Adebayo, T. A. (2022). *Contact-induced variation and change in Yoruba: Optimality in grammar and its social dimensions*. Ph.D Dissertation, Tulane University.
- Adeniyi, K. & Olaogun, S. (2020). Lexical erosion in Yoruba. *Nordic Journal of African Studies*, 29(3), 1-19.

- Adeniyi, K. (2015). *Downstep in three-tone systems of West Benue-Congo languages*. Nigeria: University of Ibadan.
- Akinlabi, A., & Liberman, M. (2000). The tonal phonology of Yoruba clitics. In B. Gerlach, & J. Grijzenhout (Eds.), *Clitics in phonology, morphology and syntax* (pp. 31-62). Amsterdam: Benjamins.
- Akinlabi, A. M. (1985). Tonal under-specification and Yoruba Tone. Ph. D Thesis, Department of Linguistics and African languages, University of Ibadan, Ibadan, Nigeria, xv + 339pp.
- Akpan, A., Urua, E-A., & Ekpenyong, M. (2018). Bridging the gap between language documentation and description for effective linguistic research. *Journal of the Linguistic Association of Nigeria*, Supplement III.
- Alabi J., Amponsah-Kaakyire, K., Adelani, D., and España-Bonet, C. (2020). Massive vs. Curated Word Embeddings for Low-Resourced Languages. The case of Yoruba and Twi. *12th International Conference on Language Resources and Evaluation (LREC)*.
- Bamgbose, A. (1990). *Fonoloji ati Girama Yoruba*. Ibadan: University Press Plc.
- Clements, G.N., Michaud, A. & Patin, C. (2011). Do we need tone features? In J.A., Goldsmith, E. Hume, & L. Wetzels (Eds.), *Tones and features: phonetic and phonological perspectives* (pp. 75-100). Berlin: Walter de Gruyter,
- Elugbe, B. (2010). A golden jubilee without gold: 50 years of language and linguistic studies in Nigeria. *Paper presented at the 2010 Convocation of the Nigerian Academy of Letters*.
- Elugbe, B. (2011). *The scramble for Nigeria*. Ibadan: DB Martoy Books.
- Fabunmi, F. A., & Salawu, A. S. (2005). Is Yoruba an endangered language? *Nordic Journal of African Studies*, 14(3), 391-408.
- Fafunwa A.B. (1975). Education in the mother tongue, a Nigeria experience: The Six-Year (Yoruba Medium) Primary Education Project at the University of Ife, Nigeria. *West African Journal of Education*, 19, 213 - 228.
- Falola, T. (2007). Nationalising Africa, culturalizing the West, and reformulating the humanities in Africa. In S. Akinrinade et al. (Eds.), *Rethinking the humanities in Africa* (pp. 25 – 44). Ile-Ife: Faculty of Arts, Obafemi Awolowo University.
- Hyman, M. (2003a). African languages and phonological theory. *GLOT International*, 7(6), 153-163.
- Hyman, L. M. (1993). Register tones and tonal geometry. In K. Snider, & H. van der Hulst (Eds.), *The phonology of tone: the representation of tonal register* (75-100). Berlin: Walter de Gruyter.
- Hyman, L.M. (2003b). Why describe African Languages? *Proceedings of World Congress of African linguistics 4/Annual conferences on African linguistics* 34, Rutgers, USA, June 18.
- Hyman, L.M. (2008.) Issues in African language phonology. *UC Berkeley Phonology Lab Annual Report*.

- Jibril, M. (2007). New directions in African linguistics. In S. Akinrinade, D. Fashina, D. O. Ogungbile, & J. O. Famakinwa (Eds.), *Rethinking the Humanities in Africa* (pp. 281 -290). Ile-Ife: Faculty of Arts, Obafemi Awolowo University.
- Osborn, D. (2010). *African languages in a digital age: Challenges and opportunities for indigenous languages computing*. South Africa: HSRC Press.
- Simons, G. F., & Fennig, C. D. (2018). *Ethnologue: Languages of the world* (21st ed.). Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>
- Snider, K. (1999). *The geometry and features of tone*. Dallas: Summer Institute of Linguistics.
- Williamson, K. (1976). The Rivers Readers Project in Nigeria. In A. Bamgbose (Ed.), *Mother tongue education: The West African Experience*, (pp.135 – 153). Paris: UNESCO.
- Williamson, K. (1984). *Practical orthography in Nigeria*. Ibadan: Heinemann Educational Books.